Agricultural & Forestry Economics and Management (2024)
Clausius Scientific Press, Canada
DOI: 10.23977/agrfem.2024.070212
ISSN 2616-2202 Vol. 7 Num. 2

Optimization of Crop Planting Strategy Based on Interior Point Method

Zihan Liu#, Xiaoyan Tan#, Yinan Li#

School of Business, Xi'an International Studies University, Xi'an, 710128, China
#These authors contributed equally.

Keywords: Interior Point Method, Multi-Stage Stochastic Programming Model, Multiple Linear Regression Model, K-Means Clustering

Abstract: For the problem of optimal planting strategy of crops in a village in North China for the next seven years, the expected sales volume is calculated based on the expected value method, and a linear programming model is constructed to formulate the optimal planting scheme of crops. In order to enhance the risk-resistant ability of the planting strategy, a multi-stage stochastic planning model of crop planting under uncertainty conditions is established and further programs are obtained. Finally, a multiple linear regression model is constructed, and the K-means clustering algorithm is used to investigate the correlation between the expected sales volume of each crop, the sales price, the planting cost, and the substitutability and complementarity between different crops. Compared with other literatures in recent years, this paper comprehensively uses Monte Carlo method and interior point method to enhance the universality and reasonableness of the optimal crop planting strategy, and to help the countryside to produce income and realize the sustainable economic and social development.

1. Introduction

The optimization of crop planting strategies is an essential requirement for improving crop production efficiency in the context of limited arable land resources in rural China. In addressing the optimization problem of crop planting strategies, the objective should be to maximize benefits while considering relevant constraints, mitigating soil continuity challenges, and enhancing efficiency[1]. This approach helps reduce planting risks associated with uncertainty factors, improves production efficiency, and promotes the sustainable development of the rural economy.

At present, certain milestones have been achieved on the issue of crop planting strategies. Chu Yanjie, Wang Siyuan and Lin Renheng starting from the mathematical planning method of solving the optimization model, the application characteristics of dynamic planning are further summarized, and the planting strategy planning in a certain place is taken as an example, which is combined with dynamic planning to determine the optimal strategy of crop planting[2]. Sun Liquan divides the objective to be planned into several stages for consideration, and the dynamic programming (DP) model is applied in order to determine an optimal planting plan for crops and achieve the purpose of optimizing things[3].

Although significant progress has been made in the study of crop planting strategy optimization problems, much of the existing literature is outdated and does not adequately address the advancements in modern planting technologies and concepts. Additionally, there is a need to focus more on incorporating multifaceted constraints. The interior point method, a numerical approach for solving linear or nonlinear programming problems, can efficiently find global optimal solutions and handle larger-scale problems. However, its application to crop planting strategies has not yet been explored in existing research.

In response to the above analysis, this paper employs the interior point method to solve a planning model aimed at determining the optimal crop planting strategy under multiple constraints. Furthermore, cluster analysis is used to investigate the correlations between different crops, thereby enhancing the universality of the planting strategy. Finally, this paper proposes practical and effective strategic recommendations for optimizing crop planting strategies. This work holds significant practical value in addressing the serious shortage of land resources in China and provides a robust foundation for the sustainable development of agricultural society.

2. System Description

2.1 Background Knowledge

Definition 1 The expected value of a decision variable g is the sum of its gain and loss values in different states of nature multiplied by the corresponding probabilities of occurrence, a method known as the expected value method.

$$A(g_i) = \sum_{j=1}^{n} p(\theta_j) g_{ij} \tag{1}$$

where:$A(g_i)$ is the variableg_i the expected value of the variable;g_{ij} is the variableg_i the value of the gain or loss in the state of nature, and$p(\theta_j)$ is the state of natureθ_j probability of occurrence.

Definition 2 The basic principle of the interior point method is to transform the original problem into an unconstrained problem by converting the constraints of the feasible domain into a penalty function constructed at the boundary of the feasible domain. The objective function can be rewritten as:

$$minF(X,\delta) = -[f(X) + \delta P(X)] \tag{2}$$

where $\delta P(X)$ is the penalty term, andδ is the penalty factor, and$P(X)$ is the penalty function.

Definition 3 A multiple linear regression model can be used to derive a confidence interval for the mean and a prediction interval for individual values of the dependent variable, based on a given independent variable. Y by deriving confidence intervals for the means of the dependent variable and prediction intervals for individual values.

The regression equation established using the multiple linear regression model is as follows:

$$Y = C + \alpha_1 X_1 + \alpha_2 X_2 \tag{3}$$

Among them.X_1 , andX_2 denote the independent variables, and$Y$ denotes the dependent variable, and$C$ is a constant.

Definition 4 The K-means clustering algorithm is a statistical method used to classify data based on the correlation between individual objects, enabling the reasonable classification of the original sample data. First, the entire sample space is numbered, and KK sample points or points outside the sample space are randomly selected as the nn initial clustering centers required for the initial clustering. Next, the Euclidean distance of each crop from the nn initial clustering centers identified in step 1 is calculated. The specific formula is:

$$D_{mn} = \sqrt{\sum_{k=1}^{2}(X_{mk} - K_{nk})^2} ,n=1,\dots,n;m=1,\dots,m \tag{4}$$

Second, the distance of each crop to each of the nn initial clustering centers is compared, and each crop is assigned to the cluster corresponding to the initial clustering center with the shortest distance. Based on the clusters defined in this step, the new clustering centers are recalculated using the following formula:

$$G = \sum_{m=1}^{k} \sum_{x_q \varepsilon A_m} \left(x_q - a_m\right)^2 \tag{5}$$

Finally, loop steps 2, 3, and 4 until it is found that the center of clustering no longer changes or the number of iterations is reached before stopping the loop and obtaining the final clustering results.

For the purpose of this analysis, it is assumed that the system satisfies the following conditions:

Hypothesis 1 When crop sales are good, sales are close to total production;

Hypothesis 2 2024~ The unit price of crops sold in the countryside in 2030 is the same as in 2023;

Scenario 3 2024~ No extreme weather events occur in the countryside in 2030 and crops can grow normally;

Scenario 4 2024~ In 2030, the market economy develops normally and there are no special events that disrupt the market order and crops can be sold normally;

2.2 System Description

2.2.1 Analyzing Crop Planting Schemes under Different Over-Treatment Based on Linear Programming Models

To explore the optimal crop planting plan for the next seven years in a village in North China, we analyzed crop planting statistics from 2023. Starting from two scenarios where total production exceeds expected sales volume, we further optimized the planting strategy by ensuring that the planting plots for each crop in each season are not overly dispersed. The countryside includes four types of land plots—flat dry land, terraced land, hillside land, and irrigated land—and two types of greenhouses: ordinary greenhouses and intelligent greenhouses. This paper uses flat dry land, terraced land, and hillside land as examples to illustrate the agricultural planting optimization model, which is established with the objective of maximizing the profit from crop planting and sales, i.e., economic efficiency.

(1) Introduction of decision variables

In order to obtain a specific planting scheme for each crop in each plot for each year, the decision variables are first introduced[4] :

$$X_{ijl} = \begin{cases} 1, & \text{Plant the -th crop at the j-th plot in the l-th year} \\ 0, & \text{Not planting the i-th crop on the j-th plot in the l-th year} \end{cases} \tag{6}$$

where $i = 1, \dots, 15$.$j = 1, \dots, 26$， $l = 1, \dots, 7$;

Considering the convenience of field management, the planting area of each crop within a single plot should not be too small. Using the crop planting data from 2023 as a reference, this paper assumes that the same crop is planted in the same plot every season, fully utilizing the planting area of the plot. Based on this assumption, the objective function and constraints are further determined.

(2) Establishment of the Objective Function

To maximize the profitability of crops, the cost-benefit theory of economics is used as the foundation for constructing the formula to calculate the profit of crop cultivation in the countryside [5]. Based on the economic cost-benefit theory, the profit formula for crop cultivation in the village is defined as the product of the total expected sales volume of the crop and the unit price, minus the total cost of cultivation. This provides the annual profit for each crop under the two scenarios considered:

When the crop's actual total production is less than the total expected sales volume.

In this case, there is no surplus production exceeding the corresponding expected sales volume.

$$f_i(X) = \sum_{j}^{26} \left(e_{ij} \cdot b_i - c_{ij}\right) \cdot d_{ij} \cdot X_{ijl} \tag{7}$$

When the crop$i$ Actual total production is greater than total expected sales:

Scenario 1: Crops$i$ Total production in excess of the corresponding expected sales volume is left unsold, resulting in wastage.

$$f_i(X) = a_i \cdot b_i - \sum_{j=1}^{26} X_{ijl} \cdot d_j \cdot c_{ij} \tag{8}$$

Scenario 2: The portion of total crop production that exceeds the corresponding expected sales volume is assumed to be sold at a reduced price of 50% of the sales price in 2023.

In Case 2, compared to Case 1, since the total production exceeds the corresponding sales volume, the strategy of selling the surplus at a reduced price is adopted instead of halting sales entirely. This approach generates additional expected sales revenue, resulting in the crop profit being calculated as follows:

$$f_i(X) = a_i \cdot b_i - \sum_{j=1}^{26} X_{ijl} \cdot d_j \cdot c_{ij} + \frac{b_i}{2}\left(\sum_{j=1}^{26} X_{ijl} \cdot d_i \cdot e_{ij} - a_i\right) \tag{9}$$

In summary the objective function is obtained as follows:

$$f(X) = \sum_{l=1}^{7}\sum_{i=1}^{15} f_i\,(X) \tag{10}$$

where.a_i denotes the expected sales volume of the *ith* crop, andb_i denotes the unit price at which *the ith* crop is sold, andc_{ij} denotes the cost of planting *the ith* crop at *the jth* plot, thed_j denotes the planted area at the *jth* plot, wheree_{ij} denotes the acre yield of crop i at plot j, and$f_{i(x)}$ denotes the profit of crop i in a year.$f_{(x)}$ denotes the total profit of the selected plots in 2024-2030.

(3) Establishment of constraints

① Plot crop type constraints

Considering the conditions of topographic features, land characteristics, climatic features, etc., each plot is suitable for growing vegetables for only one season per year, i.e:

$$\sum_{i}^{15} X_{ijl} = 1, j = 1, \dots, 26; l = 1, \dots, 7 \tag{11}$$

② Natural growth law constraints

Considering the natural growth pattern of crops, it is not suitable for the same crop to be planted in the same plot (including greenhouses) with heavy cropping in adjacent years, i.e.:

$$X_{i_1 j l_1} + X_{i_1 j (l_1+1)} < 2, j = 1, \dots, 26; l = 1, \dots, 6 \tag{12}$$

③ Soil sustainability constraints

Considering the sustainable use of the soil, it is advisable to plant legumes at least once in three years to make full use of the soil for legume rhizobacteria in order to protect and improve the environment for soil cultivation, viz:

$$X_{i_1 j l} + X_{i_2 j (l+1)} + X_{i_3 j (l+2)} \geq 1, \quad i_1, i_2, i_3 = 1, \dots 5, \quad l = 1, \dots 5 \tag{13}$$

(4) Optimize the integrated presentation of the model

The optimization model is developed as follows:

$$\max f(X)$$
$$\text{s.t.} \begin{cases} \sum_{i}^{15} X_{ijl} = 1, j = 1, \dots, 26; l = 1, \dots, 7 \\ X_{i_1 j l_1} + X_{i_1 j (l_1+1)} < 2, j = 1, \dots, 26; l = 1, \dots, 6 \\ X_{i_1 j l} + X_{i_2 j (l+1)} + X_{i_3 j (l+2)} \geq 1, \quad i_1, i_2, i_3 = 1, \dots 5, \quad l = 1, \dots 5 \end{cases} \tag{14}$$

2.2.2 Exploring Crop Planting Schemes under Uncertainty Based on Multi-Stage Stochastic Planning

On the basis of the above crop planting scheme, further considering that the expected sales volume, mu yield, planting cost and sales price of crops are subject to uncertainty by climatic factors, market conditions and other influences, the deterministic planning model in the traditional sense cannot fully reflect the basic situation of crop planting. Compared to traditional two-stage stochastic optimization, multi-stage stochastic optimization accounts for stochastic variables associated with cross-multi-period constraints. This approach involves re-planning the planting strategy at each stage to meet established constraints while also maintaining a certain correlation with the planting strategy of the subsequent stage [6].

Therefore, a multi-stage stochastic planning model for crop planting under uncertainty is considered.

Assumption 5 The year-to-year growth rates of expected sales, acreage, planting costs, and sales prices of various crops obey a uniform distribution over a range.

By assuming that their variations obey a uniform distribution within a certain range, the possible realizations of each random variable can be represented, and a multi-stage stochastic planning model with uniform distribution of expected sales volume, acreage, planting cost and sales price is established. The specific uncertainty factors are modeled as follows:

(1) Expected sales volume

① Wheat and corn

$$a_{i(l+1)} = a_{il} \cdot (1 + \alpha_l), \quad \alpha_l \sim U[5\%,10\%], \; l=1,2,\ldots,6, \quad i = 6,7 \tag{15}$$

② Crops other than wheat and corn

$$a_{i(l+1)} = a_{il} \cdot (1 + \beta_l), \beta_l \sim U[-5\%, +5\%], \; l=1,2,\ldots,6, \; i = 1,2,\ldots,5,8,\ldots,41 \tag{16}$$

(2) Sales price

① Food crops

$$b_{i(l+1)} = b_{il}, \quad l=1,2,\ldots,6, \; i = 1,2,\ldots,16 \tag{17}$$

②Vegetable crops

$$b_{i(l+1)} = b_{il} \cdot (1 + 5\%), \quad l=1,2,\ldots,6, i = 17,18,\ldots,37 \tag{18}$$

③Edible mushrooms (except morel mushrooms)

$$b_{i(l+1)} = b_{il} \cdot (1 + \gamma_l), \quad \gamma_l \sim U[-5\%, -1\%], \quad l=1,2,\ldots,6, i = 38,39,40 \tag{19}$$

④ Sheepshead Mushroom

$$b_{i(l+1)} = b_{il} \cdot (1 - 5\%), \quad l=1,2,\ldots,6, i = 41 \tag{20}$$

(3) Cultivation costs

$$c_{i(l+1)} = c_{il} \cdot (1 + 5\%), \quad l=1,2,\ldots,6, i = 1,2,\ldots,41 \tag{21}$$

(4) Yield per acre

$$e_{i(l+1)} = e_{il} \cdot (1 + \mu_l), \mu_l \sim U[-10\%, +10\%], l=1,2,\ldots,6, i = 1,2,\ldots,41 \tag{22}$$

In view of this program only in the initial planting program on the basis of the introduction of the expected sales of crops, acres of production, planting costs and sales price of a total of four random variables, the objective function has not been substantially changed, and will not be repeated listed here.

It should be noted that in terms of the constraints, in order to enhance the risk resistance of the final crop planting program and reduce the interference and influence of uncertainty factors on it, the assumption that a single plot is planted with only a single type of crop is eliminated from the previous one, and taking into account that the planting area of each crop in a single plot (including greenhouses) should not be too small. Therefore, the constraint that a single plot grows two crops per season is added to make it possible to maximize the optimal solution of the crop planting scheme for the next seven years in the village while satisfying the constraints. The specific new constraints are as follows:

$$\begin{gathered} X_{ijl} = \left[X_{1jl}, \ldots, X_{15jl}\right]^T, j = 1,\ldots,26, l = 1,\ldots,7 \\ \left|\left|X_{ijl}\right|\right|_0 = 2, j = 1,\ldots,26, l = 1,\ldots,7 \end{gathered} \tag{23}$$

Since the planting strategy of a certain crop on different plots of land is basically the same idea, the integration of multi-stage stochastic planning model is presented in detail here as an example in the case of flat dry land, terraced land and hillside land where the total production exceeds the expected sales volume partially sold at a reduced price of 50%:

$$
\begin{gathered}
maxf(X) \\
a_{i(l+1)} = a_{il} \cdot (1+\alpha_l), \quad l=1,\dots,6, i = 6,7 \\
a_{i(l+1)} = a_{il} \cdot (1+\beta_l), \quad l=1,\dots,6, i = 1,\dots,5,8,\dots,15 \\
b_{i(l+1)} = b_{il}, \quad l=1,\dots,6, i = 1,\dots,15 \\
c_{i(l+1)} = c_{il} \cdot (1+5\%), \quad l=1,\dots,6, i = 1,\dots,15 \\
e_{i(l+1)} = e_{il} \cdot (1+\mu_l), \quad l=1,\dots,6, i = 1,\dots,15
\end{gathered} \tag{24}
$$

$$
s.t.\begin{cases}
\sum_{i}^{15} X_{ijl} = 1, \quad l=1,\dots,7, i = 1,\dots,15 \\
X_{ijl} = [X_{1jl},\dots,X_{15jl}]^T, j = 1,\dots,26, l = 1,\dots,7 \\
\left|\left|X_{ijl}\right|\right|_0 = 2, j = 1,\dots,26, l = 1,\dots,7 \\
X_{ijl} \epsilon [0,1], \quad l=1,\dots,7, i = 1,\dots,15 \\
X_{i_1jl_1} \cdot X_{i_1j(l_1+1)} = 0, \quad l_1=1,\dots,6 \\
X_{i_1jl_2} + X_{i_2j(l_2+1)} + X_{i_3j(l_2+2)} \geq 1, \quad i=1,\dots,15, l_2=1,\dots,6
\end{cases} \tag{25}
$$

2.2.3 Study of Correlation Based on Multiple Linear Regression and K-Means Clustering

Simulated data is generated in the absence of real data representation, based on historical trends, model solving, in this context, the simulation data has a certain degree of rationality. And the expected value method is used in the previous paper to calculate the expected sales volume, which has a certain degree of subjectivity. Therefore, this paper here focuses on the correlation between the expected sales volume and sales price and planting cost in real life, to further optimize the crop planting strategy. Using the expected sales volume estimated by the expected value method and the existing sales price and planting cost data, a multiple linear regression model is established to determine the fit between the regression line obtained from the solution and the linear model, and based on the regression line, recalculate the value of the expected sales volume, and generate simulated data to solve the problem.

In addition, in view of the fact that substitutability and complementarity among crops are also important influencing factors of planting strategy, K-means clustering model was established to further adjust the optimal planting strategy of crops. The value of K, the number of clusters divided, is set to be 4, and the classification results are obtained to explore the substitutability and complementarity among crops.

3. Main Results

3.1 Linear Programming Model

Crops are subject to losses at the storage or marketing stage after harvest, and a review of the literature shows that the average storage loss rate for food crops is 2.17% and the marketing loss rate is 0.61%. The loss rate of vegetable crops is 25%; the loss rate of edible mushroom products is 5%. When sales are good, the wastage rate is low and the expected sales volume is close to the total crop production, the probability of this event is 0.7 When the sales situation is poor, the crop loss rate is high, at this time, combined with the total crop production and the loss rate can be calculated the expected sales volume of crops, the probability of the event is 0.3[7].

Expected sales = total production×(1-loss rate)×0.3+total output×0.7

After performing the calculations, the expected sales volume for each type of crop in 2023 can be derived.

From the data, it is evident that crops with high daily demand, such as wheat, corn, and cereals, have high expected sales volumes. In contrast, vegetable crops like lettuce, celery, and spinach show lower expected sales volumes, which may be attributed to the limited cultivation area. Perennially low temperatures in the region constrain large-scale vegetable cultivation. Therefore, when developing the planting program for 2024 – 2030, it is crucial to ensure that the total production of major food crops meets the expected sales volume, while the planting areas for vegetable crops are carefully planned.

According to the optimization objective, based on the interior point method, when the variables approach the boundary, the objective function steeply increases due to the penalty function. This mechanism prevents the variables from crossing the boundary and forces the minima of the unconstrained problem to move infinitely closer to, or within, the feasible domain DD, until it converges to the minima of the original constrained optimization problem. The maximum allowable number of iterations for the interior point method is set to be 1000, the termination tolerance for the function value is10^{-10} and the termination tolerance of X10^{-10} . Then use Matlab to solve the linear programming model obtained from the establishment of the above model, to solve the maximization problem of linear programming, and to obtain the final optimal planting scheme, take case one as an example, solve and obtain the final optimal planting scheme of crops: that is, A1 planting corn, A2 planting cereal grains, A3 planting sorghum, A4 planting black beans, A5 planting soybeans, A6 planting red beans, B1 planting sweet potatoes B2 planting climbing beans, B3 planting pumpkin, B4 planting buckwheat, B5 planting millet, B6 planting wheat, B7 planting barley, B8 planting Shinola, B9 planting corn, B10 planting soybeans, B11 planting sweet potatoes, B12 planting wheat, B13 planting soybeans, B14 planting cereals, C1 planting pumpkin, C2 planting buckwheat, C3 planting wheat, C4 planting green beans, C5 planting black beans, C6 planting sorghum. The specific data is shown in the following Figure 1:

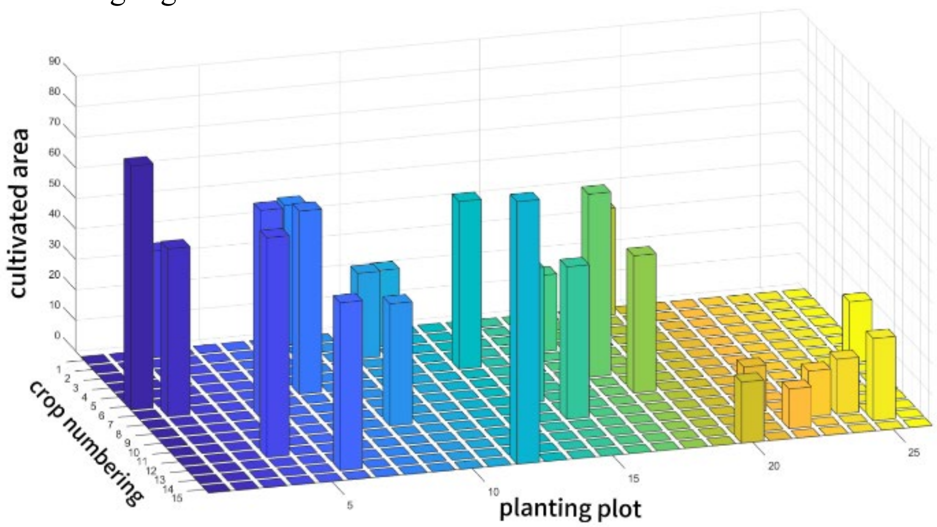

Figure 1: Graph of the results of the solution for case 1

3.2 Multi-Stage Stochastic Planning Models

The study of crop planting scheme under multi-stage stochastic planning model is still an optimization model, so the interior point ethod is still used and Matlab is used to solve this multi-stage stochastic planning model. In addition, in order to analyze the effect of crop planting strategy on profit, 100 Monte Carlo experiments were conducted, and the total profit from 2024 to 2030 was obtained after averaging them, and the final optimal crop planting scheme was obtained, i.e., A1 planting soybeans and sweet potatoes, A2 planting soybeans and wheat, A3 planting buckwheat and shinola, A4 planting shinola and barley, A5 planting red beans and corn, A6 planting sweet potatoes and barley, B1 planting green beans and pumpkins, B2 planting wheat and pumpkins, B3 planting soybeans and grains, B4 planting soybeans and Shinomai, B5 planting corn and pumpkins, B6 planting soybeans and Shinomai, B7 planting soybean mu and black beans, B8 planting red beans and crawler beans, B9 planting soybeans and buckwheat, B10 planting mungbeans and buckwheat, B11 planting mungbeans and corn, B12 planting climbing beans and sorghum, B13 planting soybeans and red beans, B14 planting soybeans and climbing beans, C1 planting Shinola and barley, C2 planting sweet potatoes and barley, C3 planting red beans and wheat, C4 planting wheat and grain, C5 planting buckwheat and Shinola, and C6 planting climbing beans and millet. The specific data is shown in Figure 2 below:

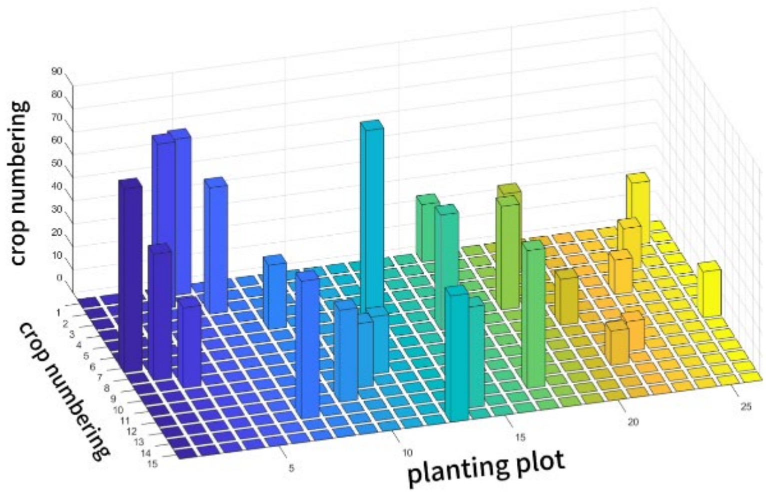

Figure 2: Three-dimensional plot of the solution result

3.3 Multiple Linear Regression and K-means Clustering

The multivariate linear regression model of expected sales volume, sales price and planting cost was solved using SPSS software and the following result Table 1 was obtained:

Table 1: Table of coefficient results

	a constant (math.)	unit price of goods sold	Cost of cultivation
ratio	6404.77865	-523.9945314	1.452541179

From the results of spss, the results of F-test show a significance P-value of 0.000***(***represents a significance level of 1%)indicating that the regression results show significance at the level, rejecting the original hypothesis that the regression coefficient is 0, and the simulated data basically satisfy the basic assumptions of the multiple linear regression model; in terms of the goodness-of-fitR^2 is 0.765, the value is close to the value of 1, indicating that the model fit results are better, and there is a higher degree of linear relationship between the expected sales volume of a certain crop of the dependent variable and the independent variables of sales price and planting cost; in addition, the parts of the multiple covariance tests, theX_1 andX_2 of VIF are all 1, both less than 10, indicating that the regression model does not have the problem of multicollinearity and the model is well constructed;

The regression effect is shown in Figure 3 below:

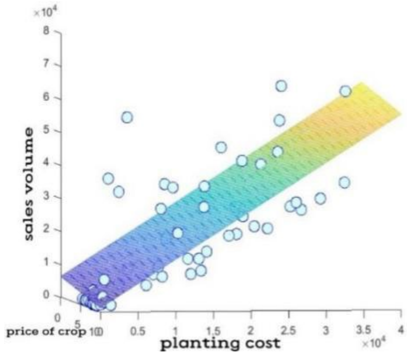

Figure 3: Return effect diagram

and obtain the final multiple linear regression equation:

$$Y = 6404.779 - 523.995 * X_1 + 1.453 * X_2 \tag{26}$$

Based on the above multiple linear regression equations, the expected sales volume data required for the planning model was recalculated to be used in the subsequent model solution.

The cost of cultivation and sales revenue of each crop were used as clustering indices. The K-means clustering algorithm was applied using SPSS software, yielding the preliminary clustering results as follows: soybeans, black beans, red beans, green beans, sorghum, rice, cowpeas, kidney beans, tomatoes, eggplants, baby bok choy, white radish, shiitake mushrooms, and shiitake mushrooms belonged to the same category, and reptile beans, millet, buckwheat, pumpkin, sweet potatoes, oat, barley kidney beans, potatoes, spinach, peppers, cauliflower, cabbage, oatmeal, cucumbers, lettuce, peppers, cabbage, yellow cabbage, celery and carrots belong to the same category, wheat, cereals and elm mushrooms belong to the same category, and maize, Chinese cabbage and morel mushrooms belong to the same category.

The crops are categorized as follows, based on the fact that they show a more pronounced degree of differentiation in terms of economic characteristics:

1) **High-return value-added** medium and high-value-added crops usually contain high-cost inputs while enjoying high yields. This suggests that there may be reasons, such as technological innovations, that make it necessary to make large capital investments while delivering efficient returns on cropping income.

2) **Cost-optimized** in Cost-optimized crops mainly refers to economic entities that control costs to enhance profits. It can be seen that the cultivation of this crop focuses on fine field management and cost control to achieve profitability.

3) **High-end value type** Middle- and high-end value crops have higher incomes, but also higher cost inputs. This reflects the fact that this type of crop mainly relies on market advantages and technological research and development, etc. for economic returns.

4) **The low-yield type of** crops in the low-yield category have low economic returns. Their inputs on costs are similar to those of category 3, but the lack of management and control of cost inputs results in inefficient utilization of funds and low levels of income, making the overall economic benefits low.

The cluster center coordinates and scatter plot of the distribution of crops in different categories are shown in Figure 4 below:

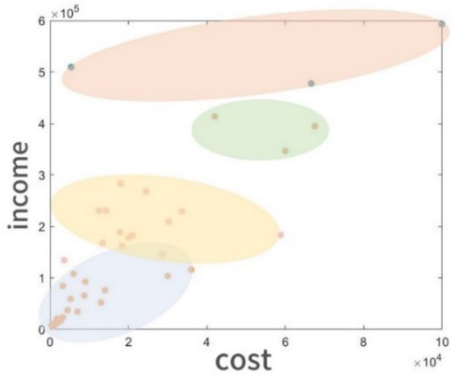

Figure 4: Scattered map of crop distribution

As shown in the figure above, the high-end value type and low-yield type exhibit greater internal differences compared to the high-return value-added type and the cost-optimization type. Additionally, the substitutability between crops within the high-end value type and low-yield type is weak, and therefore, they are not considered substitutes. Furthermore, given the specific limitations of crops suitable for planting in each type of plot, these constraints are applied to further screen the clustered data, resulting in the final substitutability characteristics between crops. This paper uses the crops in Cluster I as an example, with the following findings: soybean, black bean, red bean and mung bean can be substituted for each other in the grain (beans); tomato, eggplant, baby bok choy and white radish can be substituted for each other in vegetables; high value-added type and low-yield type in the grain category can be substituted for each other; and high yield type is not considered as a substitute for each other.

As can be seen from the above table, the richness of crops that can be optionally planted has been greatly improved by studying the relationship between substitutability and complementarity of crops. The optimized crop planting strategies are shown in Table 2 below.

Table 2: Table of planting strategy

Parcel name	crop name	area	crop name	area
A1	soya bean	36.3792	sweet potato	43.6208
A2	soya bean	27.4229	maize	27.5771
A3	buckwheat	17.137	Shinomai	17.863
A4	Shinomai	34.7989	barley	37.2011
A5	azuki bean	37.5157	corn	30.4843
A6	sweet potato	27.0017	barley	27.9983
B1	mung bean	30.532	cucurbit	29.468
B2	maize	21.3579	cucurbit	29.6421
B3	soya bean	20.3084	millet	19.6916
B4	soya bean	14.178	Shinomai	13.822
B5	corn	12.2505	cucurbit	12.7495
B6	soya bean	42.7088	Shinomai	43.2912
B7	soya bean	27.5991	black soybean	27.4009
B8	azuki bean	21.5626	refried beans	22.4374
B9	soya bean	26.3541	buckwheat	23.6459
B10	mung bean	11.6064	buckwheat	13.3936
B11	mung bean	29.4827	broomcorn millet	30.5163
B12	refried beans	22.2738	common sorghum	22.7262

B13	soya bean	16.9716	azuki bean	18.0284
B14	soya bean	11.0387	refried beans	11.0387
C1	Shinomai	8.3104	barley	6.6898
C2	sweet potato	6.505	barley	6.495
C3	azuki bean	7.6172	maize	7.3828
C4	maize	9.0279	millet	8.9721
C5	buckwheat	13.6718	Shinomai	13.3282
C6	refried beans	8.9378	broomcorn millet	11.0622

4. Conclusion

This paper provides a comprehensive analysis of the data and an in-depth explanation of the patterns and principles presented. The model established objectively and comprehensively addresses the real-world problem, taking into account many critical factors, such as crop loss rates, and develops a suitable planting plan for villages in North China. This plan contributes to improving the profitability of agricultural production. The model is also highly adaptable, allowing for quick optimization under different real-world conditions, and it can efficiently solve planning-related real-world problems. However, the quantification of potential planting risks is not sufficiently addressed, and uncertainties, such as planting risks, should be further analyzed and quantified. In summary, the model is versatile and can be applied to the development of planting programs in other regions, such as Northeast and Northwest China. Additionally, it can be adapted for other types of program development, such as procurement programs and marketing strategies.

References

[1] Wang Leiming. Research on optimization of agricultural planting structure under water resource constraints[D] . Yangling: Northwest Agriculture and Forestry University, 2017

[2] Yanjie Chu, Siyuan Wang, Renheng Lin, Qi Wang, Wanlan Zhang. Research on optimal density and fertilization program of compact corn Nonghua 101[J]. Agricultural Science and Technology Newsletter, 2015, 46(05):103-107+223.

[3] Sun Li-Quan. Application of dynamic planning model in determining the optimal planting scheme of crops[J]. Jilin Agriculture, 2015, 22(08):84.

[4] Yang Wei, Zhou Letian, Yang Peng, et al. Research on resource allocation of EV charging station based on Multi-objective programming model [J]. Journal of Tangshan Normal University, 2024, 46(03):8-12.

[5] Zhong Lihua, Pan Feng, Yang Yuyao, et al. Research on the influence mechanism of green power trading on the behavioral decisions of carbon market players under the perspective of carbon measurement [J/OL]. Electricity Measurement and Instrumentation, 2024, 1-16.

[6] Xuenan Zhao, Kaiyue Duan, Meng Li, et al. Multi-stage stochastic planning scheduling strategy considering the unanticipated nature of scheduling decisions [J]. Electric Power Construction, 2022, 43 (10): 87-97.

[7] Lin Lulu. Research on Marketing Strategy of Lidar Products of Company G [D]. Zhejiang Gongshang University, 2020.