Optimization study of crop planting strategies based on dynamic linear and correlation

Jinghan Wang [1, #], Deling Cui [1, #], Jiajing Liu [2, *, #]

[1] School of Management Engineering, Qingdao University of Technology, Qingdao, China, 266520

[2] School of Information and Control Engineering, Qingdao University of Technology, Qingdao, China, 266520

* Corresponding Author Email: Liujiajing_Rain@163.com

[#]These authors are contributed equally.

Abstract. The purpose of this paper is to study the crop planting strategy of a rural village in North China from 2024 to 2030 to cope with market fluctuations and uncertainty of planting risks. Firstly, a dynamic linear programming model is constructed and combined with Gaussian white noise to simulate random fluctuations, and Lingo software is used to solve the optimization model of crop revenue under the consideration of uncertainty. Then, on the one hand, we analyze the substitutability and complementarity between crops by introducing Spearman's correlation coefficient; on the other hand, we conduct regression analysis on the expected sales volume and sales price, planting cost, introduce the correlation coefficient to correct the existing model, and finally, we use Monte Carlo simulation to generate virtual data to solve the corrected model. The results of the study show that under the dynamic model that takes uncertainty into account, the total planting return in the village is 3,268,000 yuan, the total planting return is adjusted to 3,185,000 yuan after introducing the correlation constraint, and the total return of the grain category is increased by substituting sweet potatoes with pumpkins. This study provides a decision-making framework for agricultural planting strategies that consider economy and risk control, which is of realization significance for rural revitalization and agricultural modernization.

Keywords: Dynamic Linear Programming, Gaussian White Noise, Correlation Analysis, Monte Carlo Simulation.

1. Introduction

As an important food production base in China, the development of agricultural production in the mountainous regions of North China is of great significance to national food security. Recent studies on crop planting optimization predominantly rely on static historical data, lacking dynamic constraints. For instance, Liu et al. analyzed China's crop planting patterns over 30 years using spatiotemporal clustering based on historical parameters, while Li et al. proposed a decision tree model using 60-year climate data from the North China Plain for winter wheat and maize cultivation [1-2]. These approaches remain dependent on retrospective analyses without real-time adaptability. Although Li, J., & Chen, W. dealt with temporal dimension (seasonal planting decisions) through dynamic programming combined with linear programming to optimize spatial layout, it did not take into account the uncertainty factors such as market fluctuation and climate change, which resulted in limited practical application [3-4]. In recent years, correlation analysis has increasingly been applied in agricultural research. For instance, Meng et al. identified key drivers of China's agricultural output through integrated data processing using multivariate linear regression and time series analysis, while Aidoo et al. employed Pearson correlation and regression models to quantify the impact of climatic variables (e.g., temperature, precipitation) on maize yields in sub-Saharan Africa [5-6]. In addition, there is little research on planting strategies based on substitutability and complementarity among crops, and there is still a lack of systematic research on how to effectively integrate multifactorial dynamics and crop substitutability. In this study, an optimization framework integrating dynamic linear programming and stochastic simulation is proposed for crop cultivation planning in a village in the mountainous region of North China for the next seven years. Based on the regional cultivation characteristics and geographic environment, the model construction achieves a double breakthrough: first, by integrating the spatio-temporal dynamic constraint system (including crop cultivation arable land limitation, crop rotation system cycle, etc.), the optimization model is built according to the local conditions to fit the region; second, Monte Carlo stochastic simulation technology is introduced to quantify the degree of influence of climate fluctuations, market changes and other uncertain factors on the cultivation efficiency, which significantly improves the model robustness. improve the model robustness. Different from the traditional one-dimensional optimization paradigm, this framework not only integrates the external dynamic variables (e.g., planting constraints, market uncertainty, etc.), but also deeply analyzes the internal mechanism of the model - revealing the substitution and complementary effects among crops through correlation analysis, and regressing the multiple factors of the objective function, which ultimately forms a dynamic optimization system with multidimensional dynamic optimization system. This study can provide scientific planting programs for villages to adapt to market fluctuations and reduce planting risks, and at the same time provide theoretical references for the decision-making knowledge system of smart agriculture.

2. Research methods

2.1. Introduction to the methodology

2.1.1. Optimization analysis of crop returns considering uncertainties

The main objective of this study is to establish a crop revenue optimization model, so it is necessary to clarify the effects of uncertainty factors on the expected sales volume, planting cost, mu yield, and sales price. Firstly, the changes of planting cost, mu yield and sales price are taken as necessary constraints, and the degree of uncertainty is simulated by using Gaussian white noise; at the same time, additional constraints are identified in terms of heavy cropping, crop rotation, and convenience of field cultivation and management from the unique geographic conditions of the countryside as well as spatial and temporal features of crop planting and growth adapted to it, and the dynamic crop yield optimization model is established. The general optimization model is then used to generate the expected sales volume data of various crops in the next 7 years using Gaussian white noise, and then according to the classification basis of crops, the corresponding constraints under different crop types are selected to solve the respective yield optimization models. Finally, the results are summed to derive the overall return.

2.1.2. Impact of correlation on crop returns

Firstly, the Spearman correlation coefficient between crops was solved by using SPSS software, and the substitutability and complementarity between crops were analyzed, and then regression analysis was carried out on the expected sales volume with the sales price and the planting cost to derive the correlation between the three components. Finally, on the basis of the established optimization model, the constraints are adjusted by introducing the correlation coefficient, and the three components are solved separately using Monte Carlo simulation virtual data, and the optimal benefits of each component are summed up to arrive at the overall optimal benefits.

2.2. Data acquisition and preprocessing

The main source of data for the village is the official website of the National Student Mathematical Modeling Contest (https://www.mcm.edu.cn/). The types of arable land and crops cultivated in the village are as follows: there are 54 pieces of arable land in total, including 34 pieces of open arable land and 20 pieces of arable land in greenhouses. Open cultivated land includes flat dry land, terraced land, hillside land and watered land, among which flat dry land, terraced land and hillside land can only grow food crops in one season; watered land can grow rice in one season or vegetables in two seasons, and can only grow cabbages, carrots and radishes in the second season; greenhouse cultivation is divided into ordinary greenhouses and intelligent greenhouses, ordinary greenhouses can grow vegetables in one season and edible fungi in one season, and can only plant the edible fungi in the second season, and intelligent greenhouses can grow vegetables in two seasons. In addition, the official website also provides us with data on the area planted with different crops, planting cost, mu yield and sales price in 2023 in this village. Because the data are too complicated, this study has not only carried out the necessary cleaning operation on the existing data, but also carried out the following processing:

First of all, considering that the relationship between crop types and cultivated land is relatively complex, in order to simplify the solution process of the optimization model, this study first analyzed the correlation between crop types and cultivated land, and then classified the crop types as follows: the remaining food crops (except rice) are only grown in a single season on flat drylands, terraces and hillsides, with no intersection with other crop types and cultivated land; edible mushrooms are only grown in the second season in ordinary greenhouses, with no intersection with other crop types and cultivated land; and for the remaining crops, the relationship between each crop type and cultivated land is so complicated that the calculation of the crops can only be done in part to achieve partial optimization, but not in total, so the crops are regarded as an integrated whole. It can be seen that there is revenue independence between the three parts of crop division, so when calculating the overall benefit maximization can be calculated separately, that is, as long as the three parts of the above achieve their own benefit maximization, the sum of the three benefits is the overall benefit maximization. The specific classification is shown in Table 1 below:

Table 1. Classification of different crop types

Split Categories	Crop Types
Part I	Remaining food crops
Part II	Edible mushroom
Part III	Rice, remaining vegetable crops, cabbage - carrot - white radish

Second, because the subsequent model involves a lot of numbered data, to facilitate the understanding of this study, the crops and plots were numbered, and by the crops and plots of the number of the alternation to represent the alternation of the year: there are a total of 59 crops in the countryside, so that i denotes the number of the crops, $i = 1 \ldots 59$, then when $i = 60 \ldots 118$ then represents the second year, and so on. Similarly, there are 54 plots in the village, let j denote the number of crops, $j = 1 \ldots 54$, then when $j = 55 \ldots 108$ then it represents the second year.

Finally, considering the specificity of certain crops, special attention needs to be paid to 16 (rice), 35-37 (cabbage, white radish, red radish) and 38-41 (edible mushrooms). And for the special legume crops, this paper notates the legume array $u[i'] = [1, 2, 3, 4, 5, 17, 18, 19]$, where $i' = 0 \ldots 7$. Remember the legume cropland array $v[j'] = [1, 2 \ldots 54]$, where $j' = 0 \ldots 53$.

3. Modeling and Solving

3.1. An optimization model for crop planting returns considering uncertainty

3.1.1. Expected sales volume under uncertainty

According to the latest information released by the agricultural departments of crops grown in the mountainous regions of North China and the relevant industry associations (official website of the Ministry of Agriculture and Rural Development: Agricultural Products Market Monitoring Report in North China), the stagnant sales rate of crops was obtained to be around 5%; thus, 95% of the total crop production in 2023 (stagnant sales rate of 5%) was used to represent the expected sales volume of each crop in 2024 [7]. Based on this expected sales volume, Gaussian white noise was used to generate the expected sales volume data for the next 7 years to simulate the effect of uncertainty.

Gaussian noise is a class of noise whose probability density function obeys a Gaussian distribution (i.e., normal distribution), which can simulate small changes by adding specific mean and standard deviation to the two-dimensional data [8]. Using MATLAB to apply Gaussian noise to the expected sales of some crops, the specific results are shown in Table 2 below:

Table 2. Data on expected sales volume of selected crops before and after taking uncertainties into account (in pounds)

Crop Type	Steady State	Consideration of Uncertainties						
	2024~2030	2024	2025	2026	2027	2028	2029	2030
maize	162298	170413	178934	191459	202946	219182	230141	232443
corn	126113	133679	140363	154340	162120	171847	183876	185715
soya bean	54150	53609	53609	54145	55227	54123	55205	55757
rice	19950	19751	20343	20140	20140	20542	20132	20333
mushroom	6864	6795	6999	6929	6929	7068	6926	6996

3.1.2. A dynamic optimization model for crop returns under uncertainty

1) Establishment of the objective function

For the sake of simplicity, it is stipulated that the stagnation strategy is applied directly to the portion of the total production in each crop that exceeds the expected sales volume. The objective function of the crop planting returns taking into account the uncertainty factor is shown in equation (1) below:

$$\max W=\begin{cases}\sum_{i=1}^{7n} p_i\cdot S_i-\sum_{i=1}^{7n}\sum_{j=1}^{7m} C_{ij}\cdot x_{ij}\cdot a_{ij}, & P_i>S_i\\ \sum_{i=1}^{7n}\sum_{j=1}^{7m} p_i\cdot P_i-\sum_{i=1}^{7n}\sum_{j=1}^{7m} C_{ij}\cdot x_{ij}\cdot a_{ij}, & 0\le P_i\le S_i\end{cases} \tag{1}$$

Where i is the number of the crop, $i=1...7n$, at this time $n=59$; S_i is the expected sales volume of the $i-th$ crop; j is the number of the plot, $j=1...7m$, at this time $m=54$; C_{ij} is the cost of planting the $i-th$ crop in the $j-th$ plot; W is the total return in the next 6 years when stagnation is used; x_{ij} is the number of acres planted by the $i-th$ crop in the $j-th$ plot; p_i is the selling unit price of the $i-th$ crop; P_i is the total production of the $i-th$ crop in a year, a_{ij} is a 0-1 variable, as shown in equation (2) below:

$$a_{ij}=\begin{cases}1, & \text{The i}^{\text{th}}\text{ crop is planted on the j}^{\text{th}}\text{ plot}\\ 0, & \text{The i}^{\text{th}}\text{ crop is not planted on the j}^{\text{th}}\text{ plot}\end{cases} \tag{2}$$

2) Establishment of constraints

(1) Acre yield constraints

The acre yield of crops is affected by climate, soil, topography, and other factors and varies ±10% per year [9]. The acre yield constraints are shown in equation (3) below, where i denotes the $i-th$ crop, $i=1...n$; j denotes the $j-th$ plot, $j=1...m$; k denotes the year, $k=2024...2030$; and r_1 denotes the magnitude of the effect of the uncertainty factor on the acre yield, which takes a value between ±10%.

$$P_{i+n(k-2024),j+m(k-2024)}=P_{i+n(k-2024),j+m(k-2024)}\cdot\left(1+r_1\right)^{k-2023} \tag{3}$$

(2) Planting cost constraints

Due to the influence of market conditions, the planting cost of crops increases by about 5% per year on average, and the constraints on planting cost are shown in equation (4) below, where i denotes the $i-th$ crop, $i=1...n$; j denotes the $j-th$ plot, $j=1...m$; k denotes the year, $k=2024...2030$;

and r_2 denotes the size of the impact of uncertainty factors on planting cost, which takes the value of ±10%.

$$C_{i+n(k-2024),j+m(k-2024)} = C_{i+n(k-2024),j+m(k-2024)} \cdot \left(1+r_2\right)^{k-2023} \tag{4}$$

(3) Sales price constraints

Food is the basic need of human life, its price fluctuation will directly affect people's life, so food crops remain basically stable; vegetable crops have a tendency to increase in sales price, on average, about 5% per year. Therefore, the constraint formula for sales price is as follows:

$$p_{i+n(k-2024),j+m(k-2024)} = p_{i+n(k-2024),j+m(k-2024)} \cdot \left(1+r_3\right)^{k-2023} \tag{5}$$

For grain crops ($i=$ 1...16), r_3 takes the value of about 0%; for vegetable crops ($i=$ 17...37), r_3 takes the value of about 5%; and for edible mushroom crops (38~40), r_3 takes the value of 1~5%.

(4) Successive cropping constraints

Growing the same crop on the same plot for a long period results in soil elemental imbalance, deterioration of physico-chemical properties, and vulnerability of the crop to a single pest or disease. Therefore, there is a need to regulate continuous cropping.

$$a_{ij} + a_{i+n,j+m} \le 1, \quad i = 1..6n, j = 1..6m \tag{6}$$

From equation (2), a_{ij} indicates whether the $i-th$ crop is planted on the $j-th$ plot or not, and equation (6) is expressed as not planting the same crop in two consecutive years.

(5) Crop rotation constraints

Crop rotation is the regular rotation of different crops on the same piece of land for the purpose of equalizing soil nutrients. The present study focuses on constraining this situation by planting legume crops at least once in every three years in the village.

$$a_{u\left[i'\right]+nk,v\left[j'\right]+mk} + a_{u\left[i'\right]+n(k+1),v\left[j'\right]+m(k+1)} + a_{u\left[i'\right]+n(k+2),v\left[j'\right]+m(k+2)} \ge 1 \tag{7}$$

According to the previous section, $u\left[i'\right]$ is the legume array, $v\left[j'\right]$ is the legume cropland array, and k is noted as the year iteration variable. Where $i' = 0 \ldots 7$, $j' = 0 \ldots 53$, and $k = 0 \ldots 5$.

(6) Farming and management constraints

The following constraints have been established because farmers generally do not allow crops to be spread out too thinly or individual plots (including greenhouses) to be too small in order to facilitate farming operations and field management:

First, for rice #16, if one season of rice is chosen to be planted on a watered plot, the plot is to be planted with all rice, at which point the number of acres with rice planted on the $j-th$ plot is equal to the acreage planted on that plot.

$$x_{16+nk,j} = m_j \cdot a_{16+nk,j} \tag{8}$$

Where $x_{16+nk,j}$ denotes the number of acres of rice planted on the $j-th$ plot in the $k-th$ year, m_j denotes the acreage planted on the $j-th$ plot, and $a_{16+nk,j}$ denotes whether or not rice was planted on the $j-th$ plot in the $k-th$ year. $j = 1 \ldots 7m$, $k = 1 \ldots 7$.

Second, for all legume crops, if legumes are selected for planting in plot j, then all legumes are to be planted in that plot, at which point the number of acres with legumes planted in plot j equals the acreage of that plot.

$$x_{u[i']+nk,j} = m_j \cdot a_{u[i']+nk,j} \tag{9}$$

The interpretation of equation (9) is similar to that of equation (8) except that it is replaced by legume crops. $i' = 0\ldots7$, $j = 1\ldots7m$, $k = 1\ldots7$.

Finally, for other types of crops, it was found that a maximum of two crops were being co-cropped in the same plot. There are:

$$x_{ij} \geq \frac{m_j}{2}, \quad i \neq 16, u[i'] \tag{10}$$

Where x_{ij} denotes the number of acres of the $i-th$ crop planted in the $j-th$ plot and m_j denotes the number of acres in the $j-th$ plot.

(7) Watered land planting constraints

From the previously mentioned cultivation law in the countryside: the watered land can be planted with one season of rice or two seasons of vegetables, and the second season can only be planted with cabbages, white radishes, and carrots, which needs to be considered separately here due to the "or" relationship between crops planted in the watered land:

$$\sum_{j=27k}^{34k} a_{16j} \cdot a_{ij} = 0, \quad i = 17k..34k, \quad k = 1..7 \tag{11}$$

Watered land is numbered 27 to 34, and other vegetable crops that can be grown on watered land are numbered 17 to 34.

(8) Productive capacity constraints

The production capacity of a cropland is related to the area of that cropland, and the decision variable in this model is x_{ij}, i.e., the number of acres of the $i-th$ crop planted on the $j-th$ plot, such that the maximum acreage planted is M acres, and the minimum acreage planted is m acres, and thus there is:

$$m \leq x_{ij} \leq M, \quad i = 1..7n, j = 1..7m \tag{12}$$

3.1.3. Model aggregation

1) Aggregation of generic model

Objective function:

$$\max W = \begin{cases} \sum_{i=1}^{7n} p_i \cdot S_i - \sum_{i=1}^{7n}\sum_{j=1}^{7m} C_{ij} \cdot x_{ij} \cdot a_{ij}, & P_i > S_i \\ \sum_{i=1}^{7n}\sum_{j=1}^{7m} p_i \cdot P_i - \sum_{i=1}^{7n}\sum_{j=1}^{7m} C_{ij} \cdot x_{ij} \cdot a_{ij}, & 0 \leq P_i \leq S_i \end{cases} \tag{13}$$

Constraints:

$$\text{s.t.}\begin{cases} P_{i+n(k-2024),j+m(k-2024)} = P_{i+n(k-2024),j+m(k-2024)} \cdot (1+r_1)^{k-2023} \\ C_{i+n(k-2024),j+m(k-2024)} = C_{i+n(k-2024),j+m(k-2024)} \cdot (1+r_2)^{k-2023} \\ p_{i+n(k-2024),j+m(k-2024)} = p_{i+n(k-2024),j+m(k-2024)} \cdot (1+r_3)^{k-2023} \\ x_{16+nk,j} = m_j \cdot a_{16+nk,j} \\ x_{u[i']+nk,j} = m_j \cdot a_{u[ii']+nk,j} \\ x_{ij} \geq \dfrac{m_j}{2}, \quad i \neq 16, u[i'] \\ \sum\limits_{j=27k}^{34k} a_{16j} \cdot a_{ij} = 0, \quad i = 17k..34k, k = 1..7 \\ m \leq x_{ij} \leq M, \quad i = 1..7n, j = 1..7m \\ a_{ij} + a_{i+n,j+m} \leq 1, \quad i = 1..6n, j = 1..6m \\ a_{u[i']+nk,v[j']+mk} + a_{u[i']+n(k+1),v[j']+m(k+1)} + a_{u[i']+n(k+2),v[j']+m(k+2)} \geq 1 \end{cases} \tag{14}$$

2) Aggregation of the chunking models

(1) Part I

From Table 1, the first part is the rest of the food crops, with a total of 15 different types of crops, i.e., $n = 15$; and a total of 26 different plots, i.e., $m = 26$. Therefore, the following adjustments are made on top of equation (1):

$$\max W_1 = \begin{cases} \sum\limits_{i=1}^{7\times15} p_i \cdot S_i - \sum\limits_{i=1}^{7\times15}\sum\limits_{j=1}^{7\times26} C_{ij} \cdot x_{ij} \cdot a_{ij}, \quad P_i > S_i \\ \sum\limits_{i=1}^{7\times15}\sum\limits_{j=1}^{7\times26} p_i \cdot P_i - \sum\limits_{i=1}^{7\times15}\sum\limits_{j=1}^{7\times26} C_{ij} \cdot x_{ij} \cdot a_{ij}, \quad 0 \leq P_i \leq S_i \end{cases} \tag{15}$$

For the first part, the continuous crop constraints, legume rotation constraints, tillage and management constraints, production capacity constraints, and three necessary constraints should be satisfied. Substituting $n = 15$, $m = 26$ into equations (3) ~ (5), (6) ~ (7), (9) ~ (10) and (12):

$$\text{s.t.}\begin{cases} P_{i+15(k-2024),j+26(k-2024)} = P_{i+15(k-2024),j+26(k-2024)} \cdot (1+r_1)^{k-2023} \\ C_{i+15(k-2024),j+26(k-2024)} = C_{i+15(k-2024),j+26(k-2024)} \cdot (1+r_2)^{k-2023} \\ p_{i+15(k-2024),j+26(k-2024)} = p_{i+15(k-2024),j+26(k-2024)} \cdot (1+r_3)^{k-2023} \\ a_{ij} + a_{i+15,j+26} \leq 1, \quad i = 1..15\times6, j = 1..26\times6 \\ x_{u[i']+15k,j} = m_j \cdot a_{u[i']+15k,j} \\ x_{ij} \geq \dfrac{m_j}{2}, \quad i \neq u[i'] \\ 13 \leq x_{ij} \leq 80 \\ a_{u[i']+15k,v[j']+26k} + a_{u[i']+15(k+1),v[j']+26(k+1)} + a_{u[i']+15(k+2),v[j']+26(k+2)} \geq 1 \end{cases} \tag{16}$$

(2) Part II

From Table 1, the second part is edible mushrooms. There are a total of 4 different types of crops, i.e., $n = 4$; and there are a total of 16 different plots, i.e., $m = 16$. The following adjustments were made on top of equation (1):

$$\max W_2 = \begin{cases} \sum_{i=1}^{7\times4} p_i \cdot S_i - \sum_{i=1}^{7\times4}\sum_{j=1}^{7\times16} C_{ij}\cdot x_{ij}\cdot a_{ij}, & P_i > S_i \\ \sum_{i=1}^{7\times4}\sum_{j=1}^{7\times16} p_i \cdot P_i - \sum_{i=1}^{7\times4}\sum_{j=1}^{7\times16} C_{ij}\cdot x_{ij}\cdot a_{ij}, & 0 \le P_i \le S_i \end{cases} \tag{17}$$

For the second part, farming and management constraints, production capacity constraints, and three necessary constraints should be satisfied. Substituting n=4, m=16 into equations (3) ~ (5), (10), (12):

$$\text{s.t.} \begin{cases} P_{i+4(k-2024),j+16(k-2024)} = P_{i+4(k-2024),j+16(k-2024)}\cdot\left(1+r_1\right)^{k-2023} \\ C_{i+4(k-2024),j+16(k-2024)} = C_{i+4(k-2024),j+16(k-2024)}\cdot\left(1+r_2\right)^{k-2023} \\ p_{i+4(k-2024),j+16(k-2024)} = p_{i+4(k-2024),j+16(k-2024)}\cdot\left(1+r_3\right)^{k-2023} \\ x_{ij} \ge \frac{m_j}{2} \\ 0 \le x_{ij} \le 0.6 \end{cases} \tag{18}$$

(3) Part III

From Table 1, the third component includes rice, remaining vegetable crops, cabbage-white radish-red radish. There are a total of 39 different types of crops, i.e., $n = 39$; and a total of 28 different plots, i.e., $m = 28$. The following adjustments are made on top of equation (1):

$$\max W_3 = \begin{cases} \sum_{i=1}^{7\times39} p_i \cdot S_i - \sum_{i=1}^{7\times39}\sum_{j=1}^{7\times28} C_{ij}\cdot x_{ij}\cdot a_{ij}, & P_i > S_i \\ \sum_{i=1}^{7\times39}\sum_{j=1}^{7\times28} p_i \cdot P_i - \sum_{i=1}^{7\times39}\sum_{j=1}^{7\times28} C_{ij}\cdot x_{ij}\cdot a_{ij}, & 0 \le P_i \le S_i \end{cases} \tag{19}$$

For the third component, continuous crop constraints, crop rotation constraints, tillage and management constraints, watered land crop constraints, production capacity constraints, and three necessary constraints should be satisfied. Substituting $n = 39$ and $m = 28$ into equations (3) ~ (5), (6) ~ (7), (9) ~ (12):

$$\text{s.t.} \begin{cases} P_{i+39(k-2024),j+28(k-2024)} = P_{i+39(k-2024),j+28(k-2024)}\cdot\left(1+r_1\right)^{k-2023} \\ C_{i+39(k-2024),j+28(k-2024)} = C_{i+39(k-2024),j+28(k-2024)}\cdot\left(1+r_2\right)^{k-2023} \\ p_{i+39(k-2024),j+28(k-2024)} = p_{i+39(k-2024),j+28(k-2024)}\cdot\left(1+r_3\right)^{k-2023} \\ a_{ij} + a_{i+39,j+28} \le 1, \quad i = 1..39\times6, j = 1..28\times6 \\ x_{u\left[i'\right]+39k,j} = m_j \cdot a_{u\left[i'\right]+39k,j} \\ \sum_{j=17k}^{34k} a_{16j}\cdot a_{ij} = 0, \quad i = 17k..34k, \quad k = 1..7 \\ 0.6 \le x_{ij} \le 22 \\ x_{ij} \ge \frac{m_j}{2}, \quad i \ne u\left[i'\right] \\ a_{u\left[i'\right]+39k,v\left[j'\right]+28k} + a_{u\left[i'\right]+39(k+1),\left[j'j'+28(k+1)\right.} + a_{u\left[i'\right]+39(k+2),v\left[j'\right]+28(k+2)} \ge 1 \end{cases} \tag{20}$$

3.1.4. Model solution

Solving the problem with the help of Lingo software, we obtained the benefit of vegetable crop as 2350220 Yuan, edible mushroom crop as 204139 Yuan, grain crop as 330742 Yuan, rice crop as 383124 Yuan and the overall optimal benefit as 3268225 Yuan.

3.2. Modification of crop yield optimization model after considering correlation

3.2.1. Substitutability and complementarity among different crops

Spearman correlation coefficient is a non-parametric measure of correlation that grades the correlation between variables and is used to analyze the correlation between two continuous variables [10]. SPSS software was used to solve the correlation coefficients among crops and the significance level was tested and the results showed that Spearman correlation coefficient among different crops satisfied the significance level.

(1) Food crops

The results of the solution are shown in Table 3 below:

Table 3. Spearman's correlation coefficient between food crops

Types	soybean	black beans	red beans	mung beans	wheat	corn	millet	sorghum	miliaceum	mustard	pumpkin	sweet potato	shinohta
soybean	1	0.87	0.87	1	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
black bean	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1
red bean	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1
mung bean	1	0.87	0.87	1	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
wheat	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1
corn	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1
millet	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1
sorghum	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1
miliaceum	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1
mustard	0.87	0.5	0.5	0.87	0.5	0.5	0.5	0.5	0.5	1	0.5	0.5	0.5
pumpkin	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1
sweet potato	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1
shinohta	0.87	1	1	0.87	1	1	1	1	1	0.5	1	1	1

As can be seen in Table 3 above, the correlation between mustard wheat and other grain crops is moderate and the Spearman correlation coefficient between other grain crops is greater than even 0.8, i.e., it shows a strong positive correlation, which can be considered as substitutability between most of the grain vegetables and there is a certain degree of complementarity between mustard wheat and other grain crops.

(2) Selected vegetable crops

The results of the solution are shown in Table 4 below:

Table 4. Spearman's correlation coefficients between selected vegetable crops

Types	cowpea	sword bean	kidney bean	potato	tomato	eggplant	spinach	green pepper	cabbage	oily lettuce	small greens	cucumber	lettuce
cowpea	1	1	1	0.87	1	1	1	1	1	1	1	1	1
sword bean	1	1	1	0.87	1	1	1	1	1	1	1	1	1
kidney bean	1	1	1	0.87	1	1	1	1	1	1	1	1	1
potato	0.87	0.87	0.87	1	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
tomato	1	1	1	0.87	1	1	1	1	1	1	1	1	1
eggplant	1	1	1	0.87	1	1	1	1	1	1	1	1	1
spinach	1	1	1	0.87	1	1	1	1	1	1	1	1	1
green pepper	1	1	1	0.87	1	1	1	1	1	1	1	1	1
cabbage	1	1	1	0.87	1	1	1	1	1	1	1	1	1
oily lettuce	1	1	1	0.87	1	1	1	1	1	1	1	1	1
small greens	1	1	1	0.87	1	1	1	1	1	1	1	1	1
cucumber	1	1	1	0.87	1	1	1	1	1	1	1	1	1
lettuce	1	1	1	0.87	1	1	1	1	1	1	1	1	1

As can be seen in Table 4 above, the Spearman's correlation coefficients between the vegetables studied were all greater than 0.8, i.e., they showed a strong positive correlation, whereas the correlation between potatoes and the other vegetables was comparatively weaker, but in general, it can be assumed that there is substitutability rather than complementarity between these vegetables.

(3) Chinese cabbage-white radish-carrot

The results of the solution are shown in Table 5 below:

Table 5. Spearman's correlation coefficient between cabbage, white radish, carrot

Types	Chinese cabbage	white radish	carrot
Chinese cabbage	1	1	1
white radish	1	1	1
carrot	1	1	1

As can be seen in Table 5 above, the Spearman correlation coefficient between cabbage, white radish and carrot is 1, i.e., it shows a very strong positive correlation, and there is only substitution, not complementarity, between these three.

(4) Edible mushrooms

The results of the solution are shown in Table 6 below:

Table 6. Spearman's correlation coefficients within edible mushroom

Types	shaggy ink capsule	shiitake	grade of shiitake mushroom	morel mushrooms
shaggy ink capsule	1	1	0.87	0.5
shiitake	1	1	0.87	0.5
grade of shiitake mushroom	0.87	0.87	1	0.87
morel mushrooms	0.5	0.5	0.87	1

As can be seen in Table 6 above, the Spearman correlation coefficient between elm yellow mushrooms and morel mushrooms, shiitake mushrooms and morel mushrooms is 0.5, i.e., it shows a medium positive correlation, while the rest of the edibles show strong positive correlation with each other, which can be regarded as the existence of a certain degree of complementarities between elm yellow mushrooms and morel mushrooms and shiitake mushrooms and morel mushrooms and the existence of a substitutability relationship between the other crops.

3.2.2. Correlation between expected sales volume and sales price, planting costs

SPSS software was used to perform regression analysis between expected sales volume and sales price and planting cost [11]. The results of the solution are shown in Table 7 below:

Table 7. Relationship between expected sales volume and sales price, planting costs

Model		Unstandardized coefficient		Standardized coefficient	t	Significance
		B	Standard Errors	beta		
Expected sales volume and sales price	Constant	13.412	3.431		3.909	0.000
	Expected sales volume	0.000	0.000	-0.224	-1.437	0.159
Expected sales volume and. growing costs	Constant	2140.876	413.378		5.179	0.000
	Expected sales volume	-0.012	0.009	-0.226	-1.450	0.155

As can be seen from Table 7 above, both expected sales volume and sales price and planting cost are negatively correlated, and with the growth of expected sales volume, both sales price and planting cost gradually decrease.

3.2.3. Optimal planting strategy after taking relevant factors into account

(1) Consideration of substitutability and complementarity between crops

Sweet potatoes and pumpkin as an example (both correlation coefficient of 1, with the substitutability), the substitutability of the relationship between the existing constraints to be adjusted, due to sweet potatoes and pumpkin belongs to the same food crops, and food crops sales revenue has independence, thus only the food crop revenue to analyze the solution can be.

Using Lingo software to solve the problem, the gain of the food category is 355434 yuan, compared with the overall gain of the unadjusted optimization model (330742 yuan), which indicates that the planting scheme of expanding the sweet potato planting area and replacing pumpkin with sweet potato is scientific, and that this model has high sensitivity to the change of the planting scheme of the crops that have an alternative relationship, which further illustrates that this planting model is rationality and practicality is better.

(2) Consider the correlation between expected sales volume and sales price and planting costs

The existing constraints are adjusted by introducing the correlation factor k in this study. k is the result of the standardized coefficients between expected sales and sales volume, and between expected sales and planting costs in Table 7.

Monte Carlo methods, also known as random sampling techniques or statistical test methods, are non-deterministic (probabilistic statistical or stochastic) numerical methods used to solve mathematical and physical problems [12]. In this study, based on the range of uncertainty of the relevant data given in the previous section, the Monte Carlo method was used to simulate the data to generate the expected sales volume, acre yield, planting cost, and unit price of sales for the years 2024-2030, which were substituted into the optimization model to solve the problem.

Solving the problem with the help of Lingo software, we obtained the benefit of vegetable crop as 17097452 Yuan, edible mushroom crop as 246733 Yuan, grain crop as 283324 Yuan, rice crop as 373535 Yuan and the overall optimal benefit as 3184644 Yuan.

4. Conclusion

This study proposes a dynamic optimization framework that integrates Gaussian noise simulation and Monte Carlo methods to effectively deal with uncertainty and correlation in agricultural cultivation. The dynamic yield optimization model can satisfy yield maximization while weighing risk control, while correlation studies can further rationally allocate resources, which provides a scalable decision-making tool for the modernization of mountainous agriculture, and has a wide range of application potential. The model still has deficiencies: the current calibration of model parameters still relies on historical static datasets, and the response mechanism to sudden climate anomalies has not yet been fully considered; the ecological sustainability indicators exist only in the form of basic constraints, and the role of soil microbial activity, carbon and nitrogen cycling, and other deep ecological processes on the planting strategy has not been quantified. In future work, machine learning can be combined to construct a dynamic parameter prediction system with spatio-temporal continuity to further improve the accuracy and robustness of the model, and an ecosystem service valuation model can be introduced to transform the biodiversity maintenance and soil and water conservation benefits into a computable multi-objective optimization function.

References

[1] Liu Z H, Yang P, Wu W B, Li Z G, You L Z. Spatiotemporal changes in crop planting structure in China over the past 30 years [J]. Acta Geographica Sinica, 2016, 71 (5): 840 - 851.

[2] ZHAO J C, WANG C, SHI X Y, et al. Modeling climatically suitable areas for soybean and their shifts across China [J]. Agricultural Systems, 2021, 192: 103205.

[3] ZHANG C L, YANG G Q, WANG C Z, HUO Z L. Linking agricultural water-food-environment nexus with crop area planning: A fuzzy credibility-based multi-objective linear fractional programming approach [J]. Agricultural Water Management, 2023, 277: 108135.

[4] DONG Z. Research on agricultural development and rational use of rural land [J]. Heilongjiang Science and Technology Information, 2013 (14): 292.

[5] MENG L. Research and prediction of agricultural output value based on SPSS software [J]. Advances in Applied Mathematics, 2024, 13 (5): 2155 - 2159.

[6] BIUDES M S, VOURLITIS G L, VELASQUE M C S, et al. Gross primary productivity of Brazilian Savanna (Cerrado) estimated by different remote sensing-based models [J]. Agricultural and Forest Meteorology, 2021, 307: 108456.

[7] National Statistical Office. (2024). Statistical bulletin on national economic and social development in 2023 [EB/OL].

[8] HUANG J Y, WANG H X, LU F P, et al. Impact of Gaussian white noise on time-frequency distribution of PSWFs signals [J]. Journal of China Academy of Electronics and Information Technology, 2020, 15 (11): 1042 - 1049.

[9] ZHAO Y X, XIAO D P, BAI H Z, TAO F L. Research progress on the response and adaptation of crop phenology to climate change in China [J]. Progress in Geography, 2019, 38 (2): 224 - 235.

[10] CHEN X S. Analysis of water quality change trend in Wuliangsuhai using Spearman rank correlation coefficient method [J]. Leather Manufacturing and Environmental Protection Technology, 2022, 3 (7): 128 - 130.

[11] FANG T. Research on water requirement of jujube trees based on linear regression analysis and BP neural network [J]. Southern Agricultural Machinery, 2025, 56 (5): 44 - 46, 50.

[12] ZHOU J. Reliability analysis of flood control facilities based on Monte Carlo simulation [J]. Value Engineering, 2024, 43 (10): 150 - 153.