Eurasia Journal of Science and Technology
ISSN: 2663-1016
DOI: https://doi.org/10.61784/ejst3066

OPTIMIZATION OF CROP PLANTING STRATEGIES BASED ON SPEARMAN-NORMAL STOCHASTIC LINEAR PROGRAMMING

ZhuoFan Yang[1*], JinTao Hu[2], XiaoHan Yang[1], XinYu Wang[3]
[1]Computer College, Guangzhou Maritime University, Guangzhou 510725, Guangdong, China.
[2]Electronic and Electrical Engineering College, Zhaoqing University, Zhaoqing 526061, Guangdong, China.
[3]Navigation College, Guangzhou Maritime University, Guangzhou 510725, Guangdong, China.
Corresponding Author: ZhuoFan Yang, Email: yangzhuofan@gzmtu.edu.cn

Abstract: This study considers a variety of planting conditions and uncertainties, and comprehensively analyzes and optimizes crop planting strategies to achieve profit maximization. This study is based on a detailed data set of crop planting information of a village in a mountainous area of North China in 2023, including expected sales volume, planting cost, per mu yield, and selling price. In the first phase of the study, assuming that the expected sales volume, planting cost, acre yield, and selling price of the crop remained stable as in the dataset, two scenarios were considered: excess sales were unsold and excess was sold at a 50% discount. The optimal planting scheme was solved by a simple linear programming model. In the second stage of the study, in order to solve the uncertainty and potential risk of expected sales volume, yield per mu, planting cost and selling price, a fluctuation factor is introduced, and a new objective function containing the fluctuation factor is constructed by using normal curve probability random fluctuation method. In the third stage of the study, taking into account the fungibility and complementarity among crops, as well as the correlation between expected sales volume, planting cost and selling price, Spearman correlation analysis was used to define and solve the substitution complementarity coefficient and correlation coefficient. These coefficients are introduced into the new constraints of the linear programming model to further optimize the model, and finally a more scientific and practical Spearman-Normal stochastic linear programming model is proposed to optimize crop planting strategies, which is conducive to facilitate field management, improve production efficiency, and reduce planting risks caused by various uncertainties.
Keywords: Linear programming model; Normal curve probability random fluctuation method; Spearman correlation analysis; Spearman-Normal stochastic linear programming; Optimize crop planting strategies

1 INTRODUCTION

Based on the actual situation of rural areas, making full use of the limited arable land resources and adapting to local conditions to develop organic planting industries is of great practical significance for the sustainable development of rural economies. A village in the mountainous area of North China has complex geographical conditions, low temperatures, limited arable land resources, and can only grow one crop per year [1-2]. Now, in response to the planting needs of this village, it is necessary to establish a mathematical model to optimize the crop planting plan. The main goal is to rationally arrange the planting time and plots of different crops to increase production efficiency, reduce the planting risks caused by overstocking or price fluctuations, and at the same time take into account the requirements of soil rotation and the convenience of field management [3]. In addition, future climate change, fluctuations in market demand, changes in planting costs and sales prices, and other uncertain factors also need to be considered to ensure the sustainability of crop planting and maximize economic benefits.

Carlo Filippi employed integer programming to identify the crop combination that maximizes farmers' expected profits, defining market-related factors as random variables [4]. Hengtian Ma utilized linear programming to optimize crop planting schemes under a series of constraints and employed Monte Carlo simulation to model market scenarios, thereby accounting for market uncertainties [5]. JE Annetts proposed a multi-objective linear programming model aimed at considering agricultural conditions such as market prices, potential crop yields, soil and weather characteristics to optimize profits or environmental outcomes, or both [6]. From the existing literature, it is evident that optimizing crop planting schemes under multiple factors has long been a hot topic of interest for experts and scholars. The primary approach has been to establish linear programming models or combine them with other models to consider fluctuating factors [7]. However, these research methods generally fail to take into account fluctuating factors or simulate them within the fluctuation range with equal random probabilities, lacking realistic and scientific analysis.

This paper proposes a linear programming model that constructs a new objective function containing a stochastic fluctuation factor in the normal curve probability and a constraint condition including the Spearman correlation coefficient. The Spearman-Normal Stochastic LP model considers the correlation between variables more objectively by introducing the Spearman correlation coefficient, better capturing the interdependence among variables; by introducing the normal curve probability stochastic fluctuation factor, the random probability of the fluctuation factor within the fluctuation range becomes more realistic and scientific, thereby enhancing the robustness and adaptability of the model; it more realistically reflects the complexity and uncertainty in the real world, improving the practicality of the model. This model combines the efficiency and certainty of linear programming with the methods for handling correlation and randomness, making it more practical and effective in the planning of crop planting schemes.

2 METHODOLOGY

2.1 Data Analysis

2.1.1 Description of the dataset

This study is based on the specific conditions of cultivated land in a village in the mountainous area of North China and the relevant statistics of crop planting in 2023. The dataset is available at https://www.mcm.edu.cn/html_cn/node/a0c1fb5c31d43551f08cd8ad16870444.html. The first data table provides the existing cultivated land data of the village, including the land type and area of each plot, as well as explanations of the land types, such as: flat dry land, terraced fields, and hillside land can only grow one crop per year. The second data table presents the data of crops planted in the village, including the crop name, type, and cultivated land corresponding to each crop number, as well as explanations of the crops grown on each plot, such as: flat dry land, terraced fields, and hillside land are suitable for growing one season of grain crops (except rice) per year. The third data table shows the crop planting situation in 2023, including the crops planted on each plot, the corresponding planting area, and the planting season. The fourth data table provides the relevant statistics for 2023, including the acre yield, planting cost, and selling price of each crop in the corresponding land type and planting season.

2.1.2 Data preprocessing

Using the planting and sales data of 2023 as the basic input data for the model, preprocess the per mu yield, planting cost, expected sales volume and price of different crops, as shown in Figure 1 Data preprocessing below.

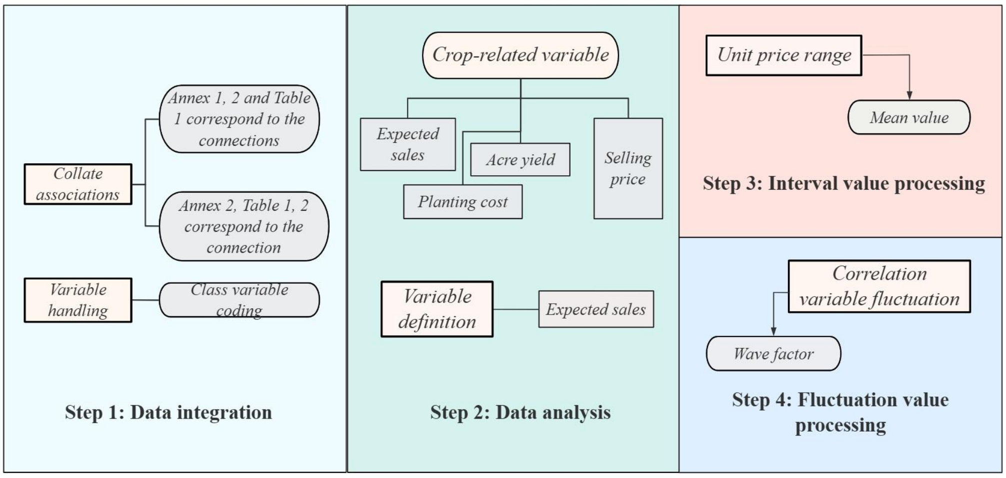

Figure 1 Data Preprocessing

By analyzing the attached table in the dataset, a data table of crop-related variables for 2023 was obtained. This table details the specific data of planting costs, acre yield, and selling prices of various crops in different plots and different seasons in 2023, for further analysis and modeling.

To facilitate the subsequent model solving, in this study, the crop types such as soybeans, black beans,..., white shiitake mushrooms, and morel mushrooms are numbered from 1 to 41 respectively, and the plots A1, A2, A3,..., F3, F4 are numbered from 1 to 50 respectively.

The related variables of crops include expected sales volume, planting cost, acre yield, and selling price. Since the dataset does not provide specific data on expected sales volume, this study adopts an ideal processing method, that is, production and sales balance. It is assumed that the total output of each crop in 2023 is equal to the expected sales volume of each crop in the future, thereby simplifying the model and providing a basic reference point for subsequent analysis and prediction. The expected sales volume of each crop is shown in Figure 2 Each crop corresponds to the total output.

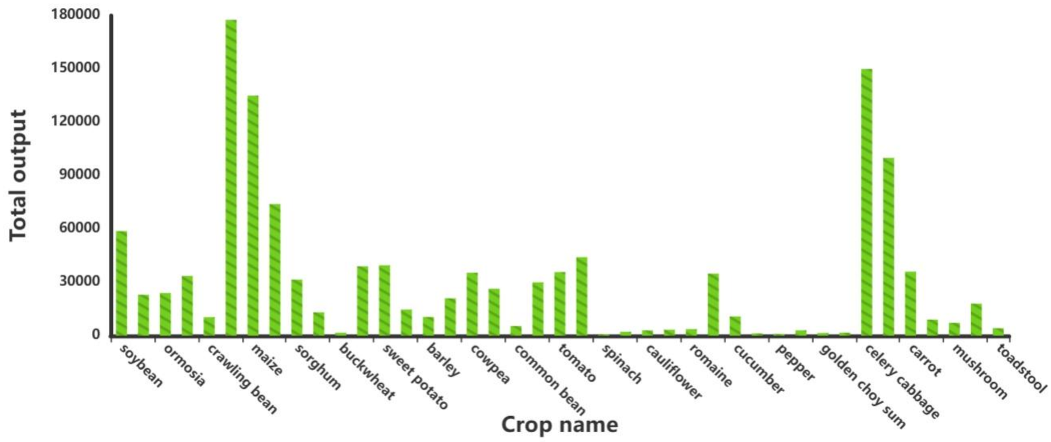

Figure 2 Each Crop Corresponds to the Total Output

This study adopts the mean approach to handle the interval data of sales unit prices, taking advantage of its characteristic of representing the central tendency to make a preliminary estimation of the data. Given that the expected sales volume, acre yield, planting cost, and sales price of crops in the dataset may all be influenced by multiple factors, which could lead to discrepancies between actual results and predictions, this study considers introducing a fluctuation factor to address the inherent uncertainties and risks in agricultural production.

2.2 Linear Programming Model

2.2.1 Model analysis

In this paper, we construct an optimal planting scheme model which seeks to maximize planting profit under various constraints. Specifically, if the total production of a crop each season exceeds the expected sales volume, the excess will be either unsalable or sold at a 50 percent discount. In order to distinguish these two sales scenarios, this study introduced an adjustment coefficient K into the objective function, so as to adjust the parameters in the objective function according to different scenarios, so as to obtain the optimal planting plan when maximizing revenue.

The constraint conditions are shown in Figure 3 Crop planting constraints, which are the constraint of season number, the constraint of total area of block number, the constraint of planting of legume crop, the constraint of repeated cropping, the constraint of crop type of plot, the constraint of special crop type and the constraint of planting area.

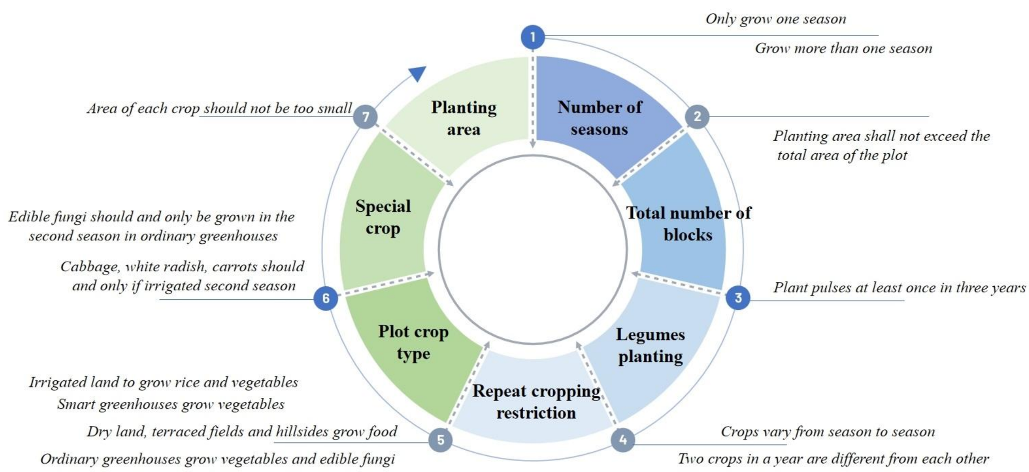

Figure 3 Crop Planting Constraints

2.2.2 Model building process

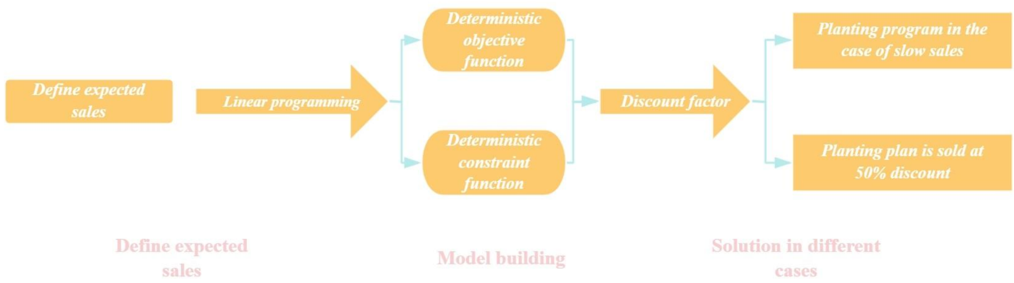

Figure 4 Linear Programming Model Building

As shown in Figure 4 Linear programming model building, the expected sales volume was first defined as the total production of each crop in 2023. A linear programming model was adopted to determine the crop planting scheme whose objective function was to maximize profit, and the constraint function was determined. The objective function contains discount coefficient K, and this study can solve the optimal planting scheme in two different situations by setting the value of K to 0 and 0.5 [8].

2.2.3 Objective function

Based on the above analysis, a linear programming model of the maximum profit from the sale of cultivated crops is established to guide the crop planting plan in future years.

Objective function:

Where $i \in \{1,2,\dots,54\}$ indicates the plot number, $j \in \{1,2,\dots,41\}$ denotes the crop number, $t \in \{2024,2025,\dots,2030\}$ represents the planting year, $s \in \{1,2\}$ represents the quarter number, $q_{i,j,t,s}$ represents the actual yield of the j crop planted in the s season of the t year in the i plot, $p_{i,j,t,s}$ represents the selling price of the j crop planted in the s season of the t year in the i plot, and d_j represents the sales volume of the j crop. $e_{i,j,t,s}$ represents the cost of planting the j crop in the s season of the t year for plot i.

$$\text{Max} \sum_{t=2024}^{2030} \sum_{i=1}^{54} \sum_{s=1}^{2} \sum_{j=1}^{41} \left[\min\left(q_{i,j,t,s}, d_j\right) \cdot p_{i,j,t,s} + \max\left(0, q_{i,j,t,s} - d_j\right) \cdot K \cdot p_{i,j,t,s} - e_{i,j,t,s}\right] \quad (1)$$

Yield of each crop:

The yield of each crop is the area planted($x_{i,j,t,s}$) multiplied by the yield per acre($y_{i,j,t,s}$).

$$q_{i,j,t,s} = x_{i,j,t,s} \cdot y_{i,j,t,s}, \forall i, \forall j, \forall t, \forall s \quad (2)$$

Planting cost of each crop:

The cost of growing each crop is the area planted($x_{i,j,t,s}$) it multiplied by the cost of growing($c_{i,j,t,s}$).

$$e_{i,j,t,s} = x_{i,j,t,s} \cdot c_{i,j,t,s}, \forall i, \forall j, \forall t, \forall s \quad (3)$$

2.2.4 Constraint function

(1) Season constraints

According to the different planting characteristics of different plot types, the number of seasons is restricted:

The land types are flat dry land, hill land and terrace, that is, $i \in \{1,2,\dots,26\}$, the plot can only grow one season of crops, that is, s = 1:

$$x_{i,j,t,1} = 0, \forall i \in \{1,2,\dots,26\}, \forall t, \forall j \quad (4)$$

When the local type is irrigated land, ordinary greenhouses, and smart greenhouses, that is, $i \in \{27,28,\dots,54\}$, the plot can be planted with one or two crops, i.e. s = 1 or s = 2:

$$x_{i,j,t,1} \geq 0, x_{i,j,t,2} \geq 0, \forall i \in \{27,28,\dots,54\}, \forall t, \forall j \quad (5)$$

(2) Constraints on the total land area

The total area of crops grown in different plots cannot exceed the plantable area of the plots A_i:

$$\sum_{j=1}^{41} x_{i,j,t,s} \leq A_i, \forall i, \forall t, \forall s \quad (6)$$

(3) Planting constraints of legumes

Every plot of land needs to be fertilized at least once every three years with legumes, which are numbered j_{legumes}:

$$\sum_{t=t_0}^{t_0+2} \sum_{s=1}^{2} x_{i,j_{\text{legumes}},t,s} \geq 1, \forall i, \forall t_0 \in \{2024,2025,2026,2027,2028\} \quad (7)$$

(4) Repeated crop restriction

Crops cannot be grown consecutively in the same planting season on the same plot or in two adjacent planting seasons:

The land types are flat dry land, terraced land and hilly land, that is, $i \in \{1,2,...,26\}$,The crops planted this year and next year cannot be the same:

$$x_{i,j,t,1} \cdot x_{i,j,t+1,1} = 0, \forall i \in \{1,2,...,26\}, \forall t, \forall j \quad (8)$$

Because in mixed integer linear programming, constraints on product form are usually not allowed because the model becomes nonlinear. Therefore, this study needs to transform the constraints of this product form into a linear form. In this study, two bivariate variables $w_{i,j,t}$ and $w^{'}_{i,j,t}$ are introduced by large M method [9].

$$x_{i,j,t,1} \le M \cdot w_{i,j,t}, \forall i \in \{1,2,...,26\}, \forall t, \forall j \quad (9)$$

$$x_{i,j,t+1,1} \le M \cdot w^{'}_{i,j,t}, \forall i \in \{1,2,...,26\}, \forall t, \forall j \quad (10)$$

$$w_{i,j,t} + w^{'}_{i,j,t} \le 1, \forall i \in \{1,2,...,26\}, \forall t, \forall j \quad (11)$$

Which $w_{i,j,t}$ and $w^{'}_{i,j,t}$ is a binary variable, constant M is large enough, this constraint means that if $x_{i,j,t,1} > 0$, then $w_{i,j,t} = 1$, thus $w^{'}_{i,j,t} = 0$, then $x_{i,j,t,1}$ must be zero, and vice versa.

The block type is irrigated land, i.e. $i \in \{27,28,...,34\}$, rice cannot be planted for two consecutive years:

$$x_{i,16,t,1} \cdot x_{i,16,t+1,1} = 0, \forall i \in \{27,28,...,34\}, \forall t, \forall j \quad (12)$$

The same can be obtained by using large M method:

$$x_{i,j,t,1} \le M \cdot w_{i,j,t}, \forall i \in \{27,28,...,34\}, \forall t, \forall j \quad (13)$$

$$x_{i,j,t+1,1} \le M \cdot w^{'}_{i,j,t}, \forall i \in \{27,28,...,34\}, \forall t, \forall j \quad (14)$$

$$w_{i,j,t} + w^{'}_{i,j,t} \le 1, \forall i \in \{27,28,...,34\}, \forall t, \forall j \quad (15)$$

When the block type is smart greenhouse, that is, $i \in \{51,52,53,54\}$, the crops planted in the second quarter of last year and the first quarter of this year cannot be the same, and the crops planted in the first quarter of this year and the second quarter of this year cannot be the same:

$$x_{i,j,t-1,2} \cdot x_{i,j,t,1} = 0, \forall i \in \{51,52,53,54\}, \forall t, \forall j \quad (16)$$

$$x_{i,j,t,1} \cdot x_{i,j,t,2} = 0, \forall i \in \{51,52,53,54\}, \forall t, \forall j \quad (17)$$

The same can be obtained by using large M method:

$$x_{i,j,t,2} \le M \cdot w_{i,j,t}, \forall i \in \{51,52,53,54\}, \forall t, \forall j \quad (18)$$

$$x_{i,j,t+1,1} \le M \cdot w^{'}_{i,j,t}, \forall i \in \{51,52,53,54\}, \forall t, \forall j \quad (19)$$

$$w_{i,j,t} + w^{'}_{i,j,t} \le 1, \forall i \in \{51,52,53,54\}, \forall t, \forall j \quad (20)$$

$$x_{i,j,t,1} \le M \cdot w_{i,j,t}, \forall i \in \{51,52,53,54\}, \forall t, \forall j \quad (21)$$

$$x_{i,j,t,2} \le M \cdot w^{'}_{i,j,t}, \forall i \in \{51,52,53,54\}, \forall t, \forall j \quad (22)$$

$$w_{i,j,t} + w^{'}_{i,j,t} \le 1, \forall i \in \{51,52,53,54\}, \forall t, \forall j \quad (23)$$

(5) Constraints of crop types in plots

Different plot types are suitable for different types of crops:

On flat dry land ($i \in \{1,2,...,6\}$), terraces ($i \in \{7,8,...,20\}$), hill land ($i \in \{21,22,...,26\}$), can only grow food crops ($j \in \{1,2,...,15\}$) :

$$x_{i,\text{nonfood},t,s} = 0, \forall i \in \{1,2,...,26\}, \forall t, \forall s \quad (24)$$

For irrigated land ($i \in \{27,28,...,34\}$), can only grow rice ($j = 16$) or vegetables ($j \in \{17,18,...,37\}$) :

$$\sum_{j \in \{1,2,...,15\} \cup \{38,39,40,41\}} x_{i,j,t,s} = 0,\ \forall i \in \{27,28,...,34\}, \forall t, \forall s \quad (25)$$

For ordinary greenhouses ($i \in \{35,36,...,50\}$), only vegetable crops can be planted in the first season ($j \in \{17,18,...,37\}$) :

$$\sum_{j \in \{1,2,...,16\} \cup \{38,39,40,41\}} x_{i,j,t,1} = 0, \forall i \in \{51,52,53,54\}, \forall t, \forall s \quad (26)$$

For smart greenhouses ($i \in \{51,52,53,54\}$), only vegetable crops can be grown ($j \in \{17,18,...,37\}$) :

$$\sum_{j \in \{1,2,...,16\} \cup \{38,39,40,41\}} x_{i,j,t,s} = 0, \forall i \in \{51,52,53,54\}, \forall t, \forall s \quad (27)$$

(6) Special crop constraints

Celery cabbage($j = 35$), mooli ($j = 36$) and carrot ($j = 37$) can only be irrigated land ($i \in \{27,28,...,34\}$) (s=2) and in the second season of irrigated land can grow any of these three crops:

$$x_{i,j,t,2} = 0, \forall i \notin \{27,28,...,34\}, \forall j \in \{35,36,37\}, \forall t \quad (28)$$

$$x_{i,j,t,1} = 0, \forall i \in \{27,28,...,34\}, \forall j \in \{35,36,37\}, \forall t \quad (29)$$

$$x_{i,j,t,2} \geq 0, \forall i \in \{27,28,...,34\}, \forall j \in \{35,36,37\}, \forall t \quad (30)$$

$$\sum_{j \in \{17,18,...,34\} \cup \{35,36,37\}} x_{i,j,t,2} = 0 , \forall i \in \{27,28,...,34\}, \forall t \quad (31)$$

For ordinary greenhouses ($i \in \{35,36,...,50\}$), in the second season (s=2) can only grow edible fungi ($j \in \{38,39,40,41\}$):

$$x_{i,j,t,2} = 0, \forall i \in \{35,36,...,50\}, \forall j \notin \{38,39,40,41\}, \forall t \quad (32)$$

$$x_{i,j,t,2} \geq 0, \forall i \in \{35,36,...,50\}, \forall j \in \{38,39,40,41\}, \forall t \quad (33)$$

Edible fungi can only be grown in the second season of ordinary greenhouses:

$$x_{i,j,t,2} = 0, \forall i \notin \{35,36,...,50\}, \forall j \in \{38,39,40,41\}, \forall t \quad (34)$$

(7) Planting area constraints

The planting area of each crop should not be less than $30\%A_i$:

$$x_{i,j,t,s} \geq 0.3A_i, \forall i, \forall j, \forall t, \forall s \quad (35)$$

2.3 Introduction Of Wave Factor

2.3.1 Model analysis

In order to comprehensively consider the uncertainty of the expected sales volume, per mu yield, planting cost and selling price of various crops, as well as the potential planting risk, this study will introduce a fluctuation factor into the model. These fluctuation factors are intended to reflect the inevitable uncertainties in actual agricultural production. Specifically, this study will introduce fluctuation factors to the expected sales volume, per mu yield, planting cost and selling price of various crops to reflect the uncertainty of these variables in actual production. This not only improves the realism of the model, but also provides a more reliable basis for decision making. The set range of the Wave factor is as follows: Figure 5 Wave factor.

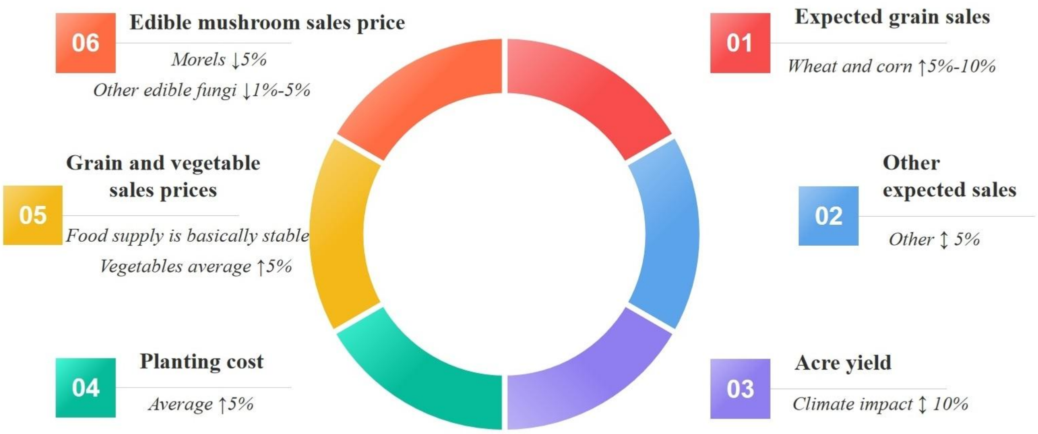

Figure 5 Wave Factor

2.3.2 Model building process

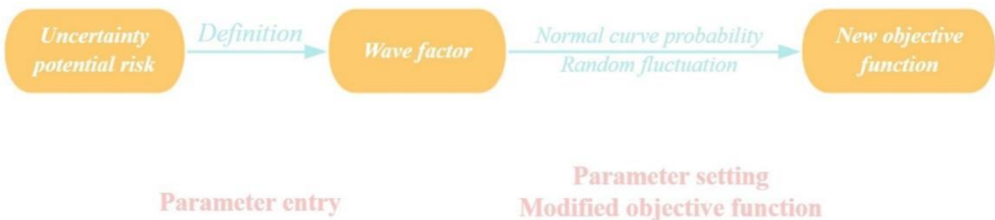

Figure 6 The Introduction Process of Wave Factor

As shown in The introduction process of wave factor in Figure 6, due to the need to comprehensively consider the expected sales volume, acre yield, uncertainty of planting cost and selling price of various crops as well as potential planting risks, the volatility factor is first defined. The constructed expression is used as the coefficient of expected sales volume, per mu yield, planting cost and selling price, and is solved by modifying the input value of the objective function through the random fluctuation of normal curve probability, and then the optimal planting scheme.

2.3.3 Definition of wave factor

(1) Uncertainty of expected sales volume

Assume that the expected sales volume fluctuation coefficient is $\theta_{j,t}^{\text{sales}}$:

$$d_{j,t} = d_{j,t-1} \times \left(1 + \theta_{j,t}^{\text{sales}}\right) \tag{36}$$

Among them, the expected sales volume fluctuation coefficient $\theta_{j,t}^{\text{sales}}$ of corn and wheat is 5% ~ 10%, and the expected sales volume fluctuation coefficient $\theta_{j,t}^{\text{sales}}$ of other crops is 5% on average.

(2) Uncertainty of acre yield

Assume that the yield fluctuation coefficient per mu is $\theta_{j,t}^{\text{yield}}$ and introduce the risk factor $\phi_{j,t}^{\text{risk}}$:

$$y_{j,t} = y_{j,t-1} \times (1 + \theta_{j,t}^{\text{yield}}) \times \left(1 - \phi_{j,t}^{\text{risk}}\right) \tag{37}$$

The area yield fluctuation coefficient of theta $\theta_{j,t}^{\text{yield}}$ of 10% ~ 10%, the risk factor $\phi_{j,t}^{\text{risk}}$ of 5% ~ 5%.

(3) Uncertainty of planting cost

Suppose the growth coefficient of planting cost is $\theta_{j,t}^{\text{cost}}$:

$$c_{j,t} = c_{j,t,s} \times (1 + \theta_{j,t}^{\text{cost}}) \tag{38}$$

The growth coefficient of planting cost $\theta_{j,t}^{\text{cost}}$ is 5% on average.

(4) Uncertainty of sales price

Suppose the selling price volatility coefficient is $\theta_{j,t}^{\text{price}}$:

$$P_{j,t} = P_{j,t} \times (1 + \theta_{j,t}^{\text{price}}) \tag{39}$$

Including grain sales price fluctuation coefficient theta $\theta_{j,t}^{\text{price}}$ for an average of 0%, grain sales price fluctuation coefficient theta $\theta_{j,t}^{\text{price}}$ for an average of 5%, morchella sales price fluctuation coefficient theta $\theta_{j,t}^{\text{price}}$ for the average -5%, The fluctuation coefficient of other edible fungi sales price $\theta_{j,t}^{\text{price}}$ is -5% ~ -1%.

2.3.4 Normal curve probability random fluctuation

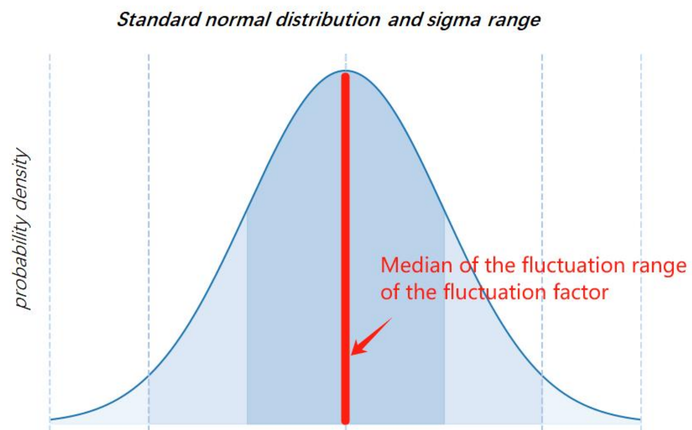

Figure 7 Normal Distribution Probability Distribution Curve

The central limit theorem states that for any sequence of independent uniformly distributed random variables with a distribution form, when the sample size is large enough, the distribution of the sample mean will tend to be normal. This theory provides the basis for crop fluctuation factor probability to be biased to normal distribution. In this study, it is more scientific to replace the normal curve probability random fluctuation factor with the main function[10]. As shown in Figure 7 normal curve probability random fluctuation, the probability of obtaining different sizes of the fluctuation factors within the range is not the same. For example, the fluctuation coefficient of mu yield $\theta_{j,t}^{\text{sales}}$ is -10% ~ 10%, then the probability of obtaining the random factor $\theta_{j,t}^{\text{sales}}$ is greater the closer it is to 0, and decreases to both sides.

2.2.5 Solution method for model

The expected sales volume, acre yield, planting cost and selling price of fluctuation factors were introduced into the objective function of Linear Programming Model for solving.

$$\text{Max} \sum_{t=2024}^{2030} \sum_{i=1}^{54} \sum_{s=1}^{2} \sum_{j=1}^{41} \left[\min\left(q_{i,j,t,s}, d_{j,t}\right) \cdot p_{i,j,t,s} - x_{i,j,t,s} \cdot c_{i,j,t,s}\right] \tag{40}$$

2.4 Introduction Of Spearman Correlation Coefficient

2.4.1 Model analysis

In real life, there may be substitutability and complementarity among various crops, and there is also correlation between expected sales and selling price and planting cost. In order to explore the relationship between them, this study needs to use the correlation analysis model to explore the possible relationship between them. Therefore, this study defines the substitution and complementation coefficient, the correlation coefficient between sales volume and price, and the correlation coefficient between planting cost and price, and introduces them into the linear programming model based on fluctuation factor by establishing constraint conditions.

2.4.2 Model building process

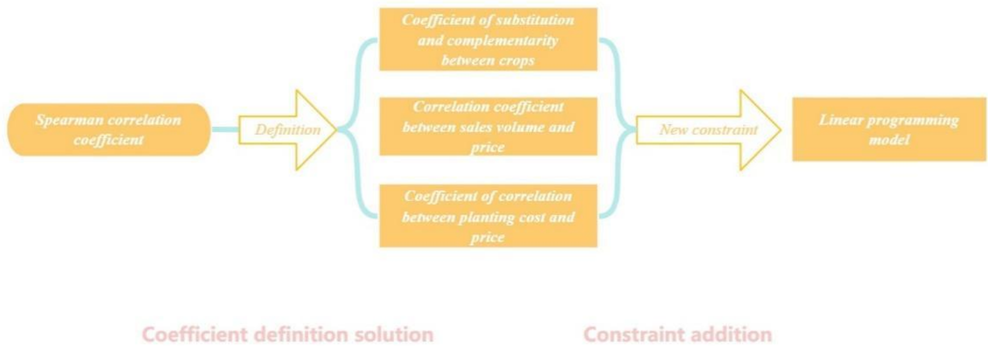

Figure 8 Spearman Correlation Coefficient Introduction Process

As shown in Figure 8 Spearman Correlation Coefficient introduction process, firstly, this study defined the substitution and complementarity coefficient between crops by Spearman. The correlation coefficient between sales volume and price and the correlation coefficient between planting cost and price. Then, the substitutability and complementarity constraints between crops, the correlation constraints between sales volume and price, and the correlation constraints between planting cost and price are added to the linear programming model based on fluctuation factors to solve.

2.4.3 Substitution and complementarity between crops

In this study, the following definitions are given for substitution and complementarity:

Substitution relationship: If two crops are substitutable, an increase in acreage for one crop may reduce the need for the other.

Complementary relationship: If two crops are complementary, an increase in acreage planted with one crop may increase demand for the other.

Based on the linear programming model based on fluctuation factors, the annual planting area variables of each crop are divided into grain, vegetable and edible fungi, and then the correlation analysis is carried out. Since the annual planting area variables of each crop did not conform to the normal distribution, Spearman correlation analysis model was used for analysis, and the results were shown as follows figure 9-11[11]:

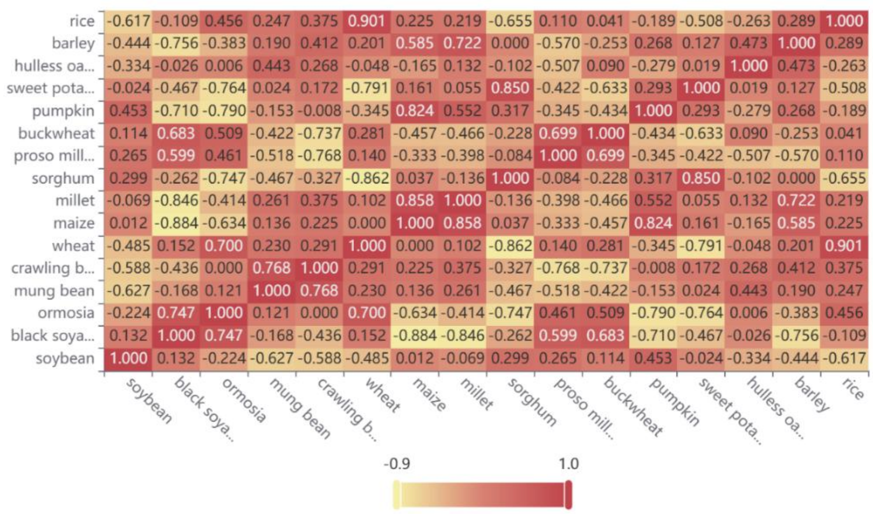

Figure 9 Thermal Matrix of Annual Planting Area Correlation of Food Crops

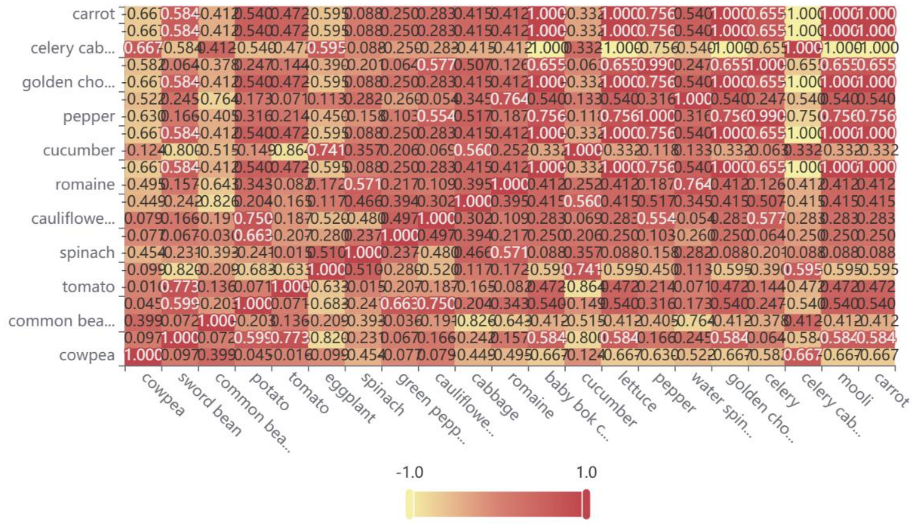

Figure 10 Thermal Matrix of Annual Planting Area Correlation of Vegetable Crops

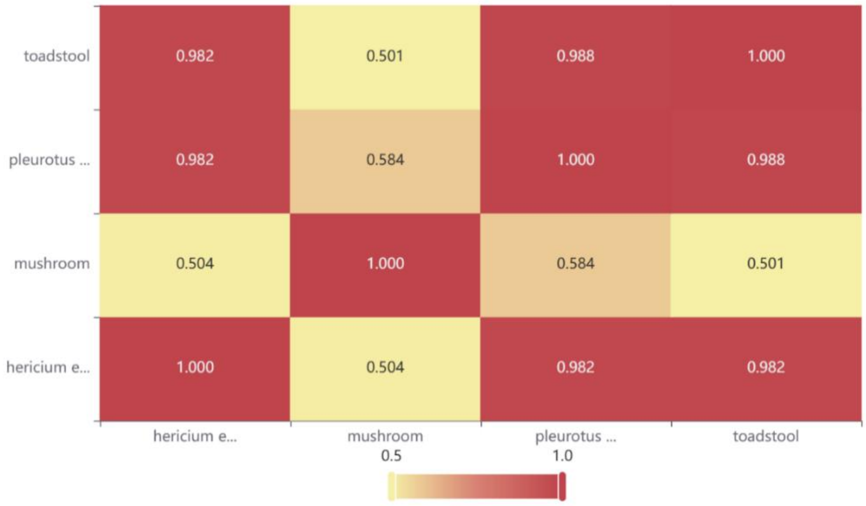

Figure 11 Thermal Matrix of Annual Planting Area of Edible Fungi

From the above three correlation thermodynamic matrices, it can be seen that when crops are negatively correlated with each other, this study believes that there is a substitution relationship between them. On the contrary, there is a complementary relationship. Therefore, the correlation coefficient λ_{y_1,y_2} between crops was defined as the substitution and complementarity coefficient, representing the substitution and complementarity relationship between y_1 and y_2. λ_{y_1,y_2} is negative, and the greater the absolute value of the coefficient, the stronger the substitution. Vice versa.
Therefore, this study optimized the expected sales volume of a certain crop in the current year as follows:

$$d_{y_2,t} = d_{y_2,t-1} + \sum_{j=1}^{41} \lambda_{j,y_2} \cdot x_{i,j,t,s} \tag{41}$$

This study uses the above formula as the first constraint added.

2.4.4 The correlation between expected sales volume, selling price and planting cost

Correlation between sales volume and price: There may be a positive or negative correlation between sales volume and price for some crops. For example, a rise in sales leads to a rise in prices, and vice versa.

$$p_{j,t} = p_{j,t-1} \cdot \left(1 + \gamma_{j,d,p} \cdot (d_{j,t} - d_{j,t-1})\right) \tag{42}$$

where $\gamma_{j,d,p}$ is the correlation coefficient representing the expected sales volume of j crop and the selling price.
Use the above formula as the new second constraint.
Correlation between planting cost and price: There may be a positive or negative correlation between planting cost and price of some crops. Rising planting costs lead to rising prices, and vice versa.

$$p_{j,t} = p_{j,t-1} \cdot \left(1 + \gamma_{j,c,p} \cdot (c_{j,t} - c_{j,t-1})\right) \tag{43}$$

Where $\gamma_{j,c,p}$ is the correlation coefficient representing the expected sales volume and selling price of j crop.
Use the above formula as the new third constraint.
In this study, Spearman correlation analysis model was used to analyze the expected sales volume and selling price of crops in this year and the planting cost and selling price, and the correlation coefficients were defined as $\gamma_{j,d,p}$ and $\gamma_{j,c,p}$, respectively. Figure 12 below shows the thermal matrix visualization of the correlation coefficients of $\gamma_{j,d,p}$ and $\gamma_{j,c,p}$ of tomatoes.

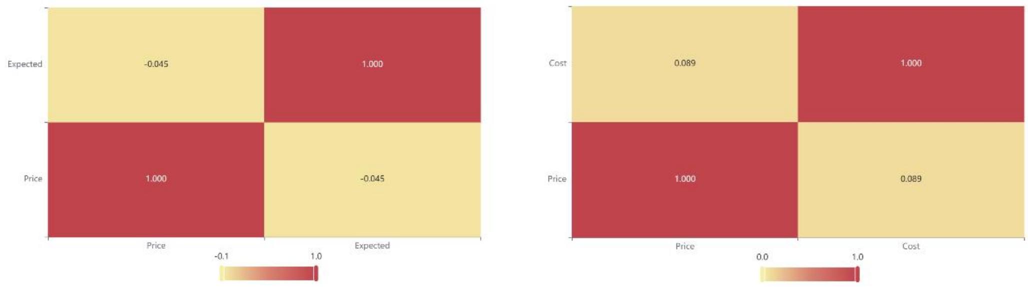

Figure 12 Tomato $\gamma_{j,d,p}$ And $\gamma_{j,c,p}$ Coefficients

Due to space constraints, $\gamma_{j,d,p}$ and $\gamma_{j,c,p}$ coefficients of other crops are shown as follows in Table 1 $\gamma_{j,d,p}$ and $\gamma_{j,c,p}$ coefficients of various crops.

Table 1 $\gamma_{j,d,p}$ And $\gamma_{j,c,p}$ Coefficients of Various Crops

crop	γ_(j,d,p)	γ_(j,c,p)	crop	γ_(j,d,p)	γ_(j,c,p)	crop	γ_(j,d,p)	γ_(j,c,p)
soybean	0.2	0.5	barley	0.12	0.121	cucumber	0.4	0.47
black soya bean	0.52	0.7	rice	0.14	0.214	lettuce	-0.5	0.687
ormosia	-0.5	0.456	cowpea	-0.42	0.75	pepper	-0.643	0.78
mung bean	-0.12	0.72	sword bean	-0.32	0.69	water spinach	-0.964	0.69
crawling bean	-0.1	0.46	common bean	-0.857	0.745	golden choy sum	0.5	0.77
wheat	-0.2	0.345	potato	-0.92	0.698	celery	0.143	0.685
maize	0.54	0.14	tomato	-0.046	0.089	celery cabbage	0.786	0.75
millet	0.11	0.93	eggplant	-1	0.87	mooli	0.036	0.97
sorghum	0.19	0.75	spinach	-0.464	0.61	carrot	-0.929	0.69
proso millet	-0.17	0.27	green pepper	-0.643	0.763	hericium erinaceus	0.786	-0.46
buckwheat	-0.4	-0.17	cauliflower	-0.714	0.473	mushroom	-0.286	-0.69
pumpkin	0.22	-0.12	cabbage	-0.107	0.479	pleurotus nebrodensis	0.75	-0.76
sweet potato	0.14	0.24	romaine	0.286	0.48	toadstool	-0.5	-0.7
hulless oat	-0.14	0.11	baby bok choy	0.857	0.147			

2.4.5 Solution method for model

Objective function:

$$\text{Max} \sum_{t=2024}^{2030} \sum_{i=1}^{54} \sum_{s=1}^{2} \sum_{j=1}^{41} \left[\min\left(q_{i,j,t,s}, d_j\right) \cdot p_{i,j,t,s} + \max\left(0, q_{i,j,t,s} - d_j\right) \cdot K \cdot p_{i,j,t,s} - e_{i,j,t,s}\right] \tag{44}$$

Constraints:
In comprehensive consideration of the expected sales of various crops, the uncertainty of planting cost and selling price, as well as the potential planting risks and the following constraints: Season number constraint, total area constraint, legume planting constraint, repeat crop constraint, crop type constraint, special crop type constraint, planting area constraint, substitution and complementarity constraint between newly introduced crops, correlation constraint between sales volume and price and correlation constraint between planting cost and price. The newly introduced constraint mathematical expressions are shown as follows:

$$d_{y_2,t} = d_{y_2,t-1} + \sum_{j=1}^{41} \lambda_{j,y_2} \cdot x_{i,j,t,s} \tag{45}$$

$$p_{j,t} = p_{j,t-1} \cdot \left(1 + \gamma_{j,d,p} \cdot (d_{j,t} - d_{j,t-1})\right) \tag{46}$$

$$p_{j,t} = p_{j,t-1} \cdot \left(1 + \gamma_{j,c,p} \cdot (c_{j,t} - c_{j,t-1})\right) \tag{47}$$

3 CONCLUSION

To sum up, the Spearman-Normal stochastic linear programming model is used in this study to obtain the optimal strategy for crop planting and marketing. It can be seen that under the circumstances of considering multiple dynamic factors and constraints, The significance of effectively optimizing the planting scheme of crops is to provide scientific decision support for agricultural production, so as to improve the economic benefits and sustainability of crops. However, due to the high computational complexity of the model and the risk of local optimal solutions, future research directions can consider introducing more intelligent optimization algorithms, such as simulated annealing, particle swarm optimization, etc., to further improve the global search ability and adaptability of the model.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Lee C C, Zeng M, Luo K. How does climate change affect food security? Evidence from China. Environmental Impact Assessment Review, 2024, 104: 107324.

[2] Guo X X, Li K L, Liu Y Z, et al. Toward the economic-environmental sustainability of smallholder farming systems through judicious management strategies and optimized planting structures. Renewable and Sustainable Energy Reviews, 2022, 165: 112619.

[3] Yang L, Wang L, Chu J, et al. Improving soil quality and wheat yield through diversified crop rotations in the North China Plain. Soil and Tillage Research, 2024, 244: 106231.

[4] Filippi C, Mansini R, Stevanato E. Mixed integer linear programming models for optimal crop selection. Computers & Operations Research, 2017, 81: 26-39.

[5] Maa H. Optimization of Crop Planting Strategies Based on Linear Programming and Monte Carlo Simulation. 2024.

[6] Annetts J E, Audsley E. Multiple objective linear programming for environmental farm planning. Journal of the operational research society, 2002, 53(9): 933-943.

[7] Alotaibi A, Nadeem F. A review of applications of linear programming to optimize agricultural solutions. International Journal of Information Engineering and Electronic Business, 2021, 15(2): 11.

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

[9] Soleimani-Damaneh M. Modified big-M method to recognize the infeasibility of linear programming models. Knowledge-Based Systems, 2008, 21(5): 377-382.

[10] Fradi A, Feunteun Y, Samir C, et al. Bayesian regression and classification using Gaussian process priors indexed by probability density functions. Information Sciences, 2021, 548: 56-68.

[11] Ali Abd Al-Hameed K. Spearman's correlation coefficient in statistical analysis. International Journal of Nonlinear Analysis and Applications, 2022, 13(1): 3249-3255.