The Study of Crop Planting Schemes Based on Hybrid Nonlinear Programming

Jialu Xu[1] Nier Wu[1]

1 College of Science, Minzu University of China, Beijing 100081;

Abstract: This study focuses on the specific conditions of rural areas and explores the impact of rational land use and the promotion of organic farming on the sustainable development of rural economies. Using crop yield data from 2023 and relevant plot information, the study applies hybrid nonlinear programming models, stochastic programming models, genetic particle swarm algorithms, and other methods to analyze the effects of land use, crop selection, and planting techniques on rural economies. The results indicate that rational crop selection and optimized planting methods can significantly simplify management, increase yields, and reduce risks. The findings highlight the crucial role of efficient land use in rural economic development and suggest the need for relevant interventions to promote sustainable development in rural areas.

Keywords: Hybrid nonlinear programming model, stochastic programming model, genetic particle swarm algorithm, planting schemes

DOI:10.69979/3041-0843.25.01.014

1. Introduction

1.1. Problem Background

This study examines a rural village in North China's mountainous region to assess the impact of locally adapted organic farming on sustainable rural economic development. With 1,201 acres of arable land and 20 greenhouses, the village supports diverse crop cultivation. An optimized planting scheme is proposed, incorporating crop rotation with legumes every three years to enhance soil quality while simplifying management and reducing fragmentation. The findings indicate that these strategies boost yields, mitigate risks, and contribute to rural economic growth.

1.2. Problem Requirements

The question provides information on the current crops planted on the arable land in the countryside and the planting situation in 2023.The goal is to establish a mathematical model to answer the following questions:

Question 1: Assuming that the future sales volume, planting costs, yield per acre, and sales prices of various crops remain the same as in 2023, with crops harvested in each season being sold within that season, if the total yield exceeds the expected sales volume, the surplus cannot be sold. Under the two scenarios where the surplus either leads to waste or is sold at 50% of the 2023 sales price, determine the optimal planting schemes for the village from 2024 to 2030.

Question 2: Taking into account the expected sales volume, yield per acre, planting costs, and the fluctuation of sales prices, as well as potential planting risks, develop the best planting plan for the village from 2024 to 2030.

1.3. Problem Analysis

1.3.1 Analysis of Question 1

This study aims to maximize crop planting benefits while considering land constraints and planting requirements. If total yield exceeds expected sales, two scenarios are analyzed: surplus waste or sale at 50% of the 2023 price. A hybrid nonlinear programming model is developed, using planting area and binary variables (crop selection) as decision variables to optimize benefits. Given the problem's complexity, a genetic algorithm is employed for solution optimization.

1.3.2 Analysis of Question 2

Building on Question 1, Question 2 incorporates uncertainties in crop sales volumes, planting costs, yield per acre, and sales prices, assuming they follow a uniform distribution. A stochastic programming model is formulated, introducing random variables and modifying the objective function to maximize profits while maintaining the same constraints. The Sample Average Approximation (SAA) method converts the stochastic model into a deterministic one, which is then solved using a hybrid genetic particle swarm algorithm.

2. Research Assumptions

1.It is assumed that the 5% figure in Question 2 is used as a basis for calculations.

2.It is assumed that the random variables involved in Question 2 follow a uniform distribution.

3.Based on the 2023 crop planting data, it is assumed that the number of plots for each crop does not exceed three. This implies that the planting of each crop should not be too dispersed across different plots, and the area planted with each crop on a single plot (including greenhouses) should not be too small.

4.It is assumed that crops of the same type have substitutability and complementarity with one another.

5.It is assumed that the expected sales volume of each crop in each season in 2023 will be the same as the yield in 2023.

3. Data Processing and Explanation

3.1. Notation Explanation

Table 1: Symbol Explanation

Symbols	Explanation
i	Types of planting plots
j	Types of crops
k	Planting seasons
t	Planting year
X_{ijkt}	The planting area of crop j on plot i in the k-th season of the year t
Y_{ijkt}	Binary decision variable indicating whether crop j is planted on plot i in the k-th season of the year t (1 for planted, 0 for not planted)
M_{ijkt}	The yield per acre of crop j on plot i in the k-th season of the year t
N_{ijkt}	The expected sales volume of crop j on plot i in the k-th season of the year t
C_{ijkt}	The cost of planting crop j on plot i in the k-th season of the year t
S_i	The maximum planting area of each plot
ξ_{jkt}	Sales price per unit

3.2. Data Preprocessing

3.2.1. Handling Missing Values

To handle missing values in the dataset, the MATLAB find function is used to locate the missing values. Once identified, the missing values are filled with appropriate estimates or defaults.

3.2.2. Mean Value Treatment for Price Data

In the question, the price range for various crops is given.Directly using the price range in the optimization model could introduce additional complexity, such as the need for uncertainty handling or more advanced methods.

To simplify the model, we use the mean value of the price range as a representative price. This approach reduces the impact of price fluctuations on the results of the optimization model.

Table 2: Sales Prices of Some Crops

Crop Names	Sales price (Yuan per jin)	Crop Name	Sales price (Yuan per jin)
Sword bean	6.24	Wheat	3.83
Cabbage	6.86	Elm mushroom	55.96

Pumpkin	1.38	Rice	7.97
Potato	4.09	Oilseed lettuce	4.72
Chinese cabbage	2.78	Climbing bean	7.29

3.2.3. Processing of Expected Sales Volume

The expected sales volume for each crop in 2023 is set as the product of the planting area and yield per acre for 2023.

4. Model Establishment and Results

4.1. Establishment and Solution of the Model for Problem 1

To address the two scenarios of excess sales volume leading to unsold goods and discounted sales, a mixed nonlinear integer programming model was established. The model considers the regions, years, seasons, and crop types, setting up the objective function and constraints accordingly. The genetic algorithm was then used to solve for the optimal planting scheme.

4.1.1. Decision Variable Setup

1.Symbol Assumptions:

1.Let i represent the types of planting plots, such as flat dry land A1 being 1, flat dry land A2 being 2, and so on.

2.Let j represent the types of crops.

3.Let k represent the planting seasons, where k=1 indicates a single season, k=2 indicates the first season, and k=3 indicates the second season.

4.Let t represent the planting year, where t=0corresponds to the year 2023, t=1 corresponds to the year 2024, and so on.

2.Decision Variables:

X_{iikt} represents the planting area (in acres) of crop j on plot i in year t, season k.

Y_{iikt} is a binary decision variable indicating whether crop j is planted on plot i in year t, season k. A value of 1 indicates that the crop is planted, while 0 indicates that it is not planted.

3. Relevant Parameters:

M_{iikt} represents the yield per acre (in jin per acre) of crop j on plot i in year t, season k.

N_{iikt} represents the pre-sale volume (in tons) of crop j on plot i in year t, season k.

C_{iikt} represents the cost (in yuan per acre) of growing crop j on plot i in year t, season k.

4.1.2. Establishment of the Objective Function

The objective is to maximize the total profit of the planting scheme for each year, based on the two given scenarios. The objective function is established as follows:

1. Unsold Waste:

When $\sum_i M_{ijkt} \times X_{ijkt} \times Y_{ijkt} \le N_{jkt}$, the yield is the sales volume.

when $\sum_i M_{ijkt} \times X_{ijkt} \times Y_{ijkt} > N_{jkt}$ only the expected sales volume can be sold. The total profit is:

$$MaxZ_1 = \begin{cases} \sum_j \sum_k \left[\left(\sum_i X_{ijkt} \times Y_{ijkt} \times M_{ijkt} \right) \times \xi_{jkt} - \sum_i X_{ijkt} \times Y_{ijkt} \times C_{ijkt} \right], & \sum_i M_{ijkt} \times X_{ijkt} \times Y_{ijkt} \le N_{jkt} \\ \sum_j \sum_k \left(N_{jkt} \times \xi_{jkt} - \sum_i X_{ijkt} \times Y_{ijkt} \times C_{ijkt} \right), & \sum_i M_{ijkt} \times X_{ijkt} \times Y_{ijkt} > N_{jkt} \end{cases}$$

(1)

2. Discounted Sales:

When $\sum_i M_{ijkt} \times X_{ijkt} \times Y_{ijkt} > N_{jkt}$时, the excess portion is sold at 50% of the original price. The total profit is:

$$MaxZ_2 = \begin{cases} \sum_j \sum_k \left[\left(\sum_i X_{ijkt} \times Y_{ijkt} \times M_{ijkt} \right) \times \xi_{jkt} - \sum_i X_{ijkt} \times Y_{ijkt} \times C_{ijkt} \right], & \sum_i M_{ijkt} \times X_{ijkt} \times Y_{ijkt} \le N_{jkt} \\ \sum_j \sum_k \left[N_{jkt} \times \xi_{jkt} + \left(\sum_i X_{ijkt} \times Y_{ijkt} \times M_{ijkt} - N_{jkt} \right) \times \xi_{jkt} \times \frac{1}{2} - \sum_i X_{ijkt} \times Y_{ijkt} \times C_{ijkt} \right], & \sum_i M_{ijkt} \times X_{ijkt} \times Y_{ijkt} > N_{jkt} \end{cases}$$

(2)

4.1.3. Establishment of Constraints

1.Plot Area Constraint:

In the three different planting seasons, for each plot i and each crop j, the total planting area for each year should not exceed the available area of the plot:

$$\sum_J X_{ij1t} \le S_i \quad (3)$$

$$\sum_j X_{ij2t} \le S_i \quad (4)$$

$$\sum_j X_{ij3t} \le S_i \quad (5)$$

2.Planting Plot Rule Constraint:

Flat dry land, terraces, and hillside plots can only be used to plant crops for one season each year:

$$\sum_{i=1}^{26} \sum_{k=2,3} \sum_t Y_{ijkt} = 0 \qquad (6)$$

For greenhouses, the constraint allows planting crops for up to two seasons each year:

$$\sum_{i=35}^{54} \sum_{k=1} \sum_t Y_{ijkt} = 0 \quad (7)$$

For flat dry land, terraces, and hillside plots, only one season of grain crops (excluding rice) can be planted each year:

$$\sum_{i=27}^{54} \sum_{j=1}^{15} \sum_k \sum_t Y_{ijkt} = 0 \quad (8)$$

$$\sum_{i=1}^{26} \sum_{j=16}^{41} \sum_k \sum_t Y_{ijkt} = 0 \quad (9)$$

For irrigated land, only one season of rice can be planted each year:

$$\sum_{i=35}^{54} \sum_{j=16} \sum_k \sum_t Y_{ijkt} = 0 \quad (10)$$

$$\sum_{i=27}^{34} \sum_{j=16} \sum_{k=2,3} \sum_t Y_{ijkt} = 0 \quad (11)$$

Vegetable Planting Constraints for Irrigated Land, Regular Greenhouses, and Smart Greenhouses:

If two seasons of vegetables are planted on an irrigated plot, the first season can include multiple types of vegetables (excluding Chinese cabbage, white radish, and red radish). In the second season, only one of the three crops (Chinese cabbage, white radish, or red radish) can be planted.

Additionally, due to seasonal requirements, Chinese cabbage, white radish, and red radish can only be planted in the second season on irrigated land.

Regular greenhouses can plant crops in two seasons each year. The first season can include multiple vegetables (excluding Chinese cabbage, white radish, and red radish). Smart greenhouses can plant crops in two seasons each year, and in both seasons, they can grow multiple vegetables (excluding Chinese cabbage, white radish, and red radish).

$$\sum_{i=27}^{34}\sum_{j=17}^{34}\sum_{k=1,3}\sum_{t} Y_{ijkt}=0 \quad (12)$$

$$\sum_{i=35}^{50}\sum_{j=17}^{34}\sum_{k=3}\sum_{t} Y_{ijkt}=0 \quad (13)$$

$$\sum_{i=27}^{34}\sum_{j=35}^{37}\sum_{k=1,2}\sum_{t} Y_{ijkt}=0 \quad (14)$$

$$\sum_{i=35}^{54}\sum_{j=35}^{37}\sum_{k}\sum_{t} Y_{ijkt}=0 \quad (15)$$

Planting Constraints for Edible Mushrooms in Regular Greenhouses:

Since edible mushrooms thrive in environments with lower and suitable temperatures and humidity, they can only be planted in regular greenhouses during the autumn and winter seasons.

$$\sum_{i=27}^{34}\sum_{j=38}^{41}\sum_{k}\sum_{t} Y_{ijkt}=0 \quad (16)$$

$$\sum_{i=35}^{50}\sum_{j=38}^{41}\sum_{k=2}\sum_{t} Y_{ijkt}=0 \quad (17)$$

3. Legume Crop Planting Frequency Constraint:

Each plot of land must be planted with legume crops at least once every three years.

$$\sum_{j=1}^{5}\sum_{j=17}^{19}\sum_{k}(X_{ijk(t-2)}+X_{ijk(t-1)}+X_{ijkt})\geq S_i,\quad \forall i,t\geq 2 \qquad (18)$$

$$\sum_{j=1}^{5}\sum_{j=17}^{19}\sum_{k} Y_{ijkt}\geq 1,\quad \forall i,t\geq 2 \qquad (19)$$

4. Non-Consecutive Planting Constraint:

Each crop cannot be planted consecutively on the same plot.

$$\forall i,j,t: \qquad Y_{ij1t}+Y_{ij1(t+1)}\leq 1 \quad (20)$$

$$Y_{ij2t}+Y_{ij3t}\leq 1 \quad (21)$$

$$Y_{ij3t}+Y_{ij2(t+1)}\leq 1 \quad (22)$$

$$Y_{ij3t}+Y_{ij1(t+1)}\leq 1 \quad (23)$$

$$Y_{ij1t}+Y_{ij2(t+1)}\leq 1 \quad (24)$$

5. Planting Plan Constraints for Cultivation and Field Management:

Different crops can be intercropped on the same plot each season for easier farming operations and field management. However, the planting of each crop per season should not be too scattered, and the area of each crop planted on a single plot (including greenhouses) should not be too small.

$$\forall i,k,t \qquad \sum_{j} Y_{ijkt}\leq 3 \quad (25)$$

6. Non-Negativity and Binary Constraints:

X must be non-negative, and YYY can only take values of 0 or 1.

$$\forall t\geq 2,\forall j \qquad \sum_{i}\sum_{k} Y_{ijkt}\geq 1 \qquad (26)$$

7. Maximum Number of Plots for Each Crop:

$$\forall jkt \qquad \sum_{i} Y_{ijkt}\leq 7 \quad (27)$$

4.1.4. Model Solution

Due to the complexity of the problem, which involves multiple plots, various crops, and complex constraints with a nonlinear objective function, a heuristic algorithm—genetic algorithm—is used to solve the model.

1. Overview of Genetic Algorithm

The genetic algorithm is a global search method that simulates biological evolution by operating on a population of feasible solutions, each carrying genetic information. Through selection, crossover, and mutation, it follows the principle of "survival of the fittest" to optimize solutions iteratively. The process begins with an initial population, where individuals are evaluated for fitness. If the optimization criteria are met, the best solution is selected; otherwise, selection, crossover, and mutation continue until convergence is achieved.

4.1.5 Solution Results

The obtained solution results show that the total profit for the first scenario is 43,550,000, while the total profit for the second scenario is 46,788,000.

4.2. Establishment and Solution of the Model for Problem 2

4.2.1. Analysis of Uncertain Factors

In Problem 2, the optimal planting scheme needs to be developed under an uncertain environment. The main uncertainties include:

1.Expected Crop Sales Volume: The sales volume of wheat and corn increases by 5%-10% annually, while the expected sales volume of other crops fluctuates by 5% each year.

2.Crop Yield per Acre: Due to factors such as climate, the yield per acre fluctuates by 10%.

3.Planting Cost: The planting cost increases by an average of 5% each year.

4.Sales Price:Prices of grain crops remain relatively stable.Prices of vegetable crops increase by 5% annually.Prices of edible mushrooms decrease each year by 1%-5%, with the price of morel mushrooms decreasing by 5% annually.

4.2.2. Establishment of the Stochastic Programming Model

Due to the uncertainties in expected crop sales volume, yield per acre, planting costs, and sales prices, it is challenging to solve the problem using traditional deterministic programming methods. Therefore, this study adopts a stochastic programming model to handle the random variables in the objective function. The model aims to determine the optimal planting scheme for each plot from 2024 to 2030, maximizing overall profits.

1. Probability Distribution Function of Random Variables

The prerequisite for establishing the stochastic programming model is to know the probability distribution function of the random variables. In this study, it is assumed that all random variables follow a uniform distribution, as detailed in the following table:

Table 2: Probability Distribution Functions of Random Variables

Random Variable	Assumed Distribution	Random Variable	Assumed Distribution
Wheat and corn expected sales volume change rate a	$U(5\%,10\%)$	Expected Sales Volume Change Rate for Other Crops b	$U(-5\%,5\%)$
Crop Yield per Acre Change Rate d	$U(-10\%,10\%)$	Edible Mushroom Price Change Rate f	$U(-5\%,-1\%)$

2. Decision Variable Setup

X_{iikt} a represents the planting area (in mu) of crop j on plot i in year t, season k.

Y_{iikt} represents the binary decision variable indicating whether crop j is planted on plot i in year t, season k. A value of 1 indicates planting, while 0 indicates no planting.

3. Relevant Parameters

M_{iikt} represents the yield per acre (jin/mu) of crop j on plot i in year t, season k.

N_{iikt} represents the presale volume (tons) of crop j on plot i in year t, season k.

C_{iikt} represents the cost (yuan/mu) of planting crop j on plot i in year t, season k.

XI_{ikt} represents the sales price (yuan/jin) of crop j in year t, season k.

e represents the planting cost, which increases by an average of 5% each year.

g represents the price of vegetable crops, which increases by 5% each year.

h represents the price of morel mushrooms, which decreases by 5% each year.

4. Establishment of the Objective Function

The objective is to maximize total profit, defined as the total planting revenue minus the total planting cost. Revenue is determined by the crop yield and sales price, while profit is determined by the planting area and the seasonal crop sales price. It can be expressed as follows:

$$\max E[f(X_{ijkt}, Y_{ijkt,} a, b, d, f)] \qquad (28)$$

In the equation, $X_{ijkt,} Y_{ijkt}$ represents the decision variable to be solved, while a,b,d,f are the random variables in this problem. The specific expression for $f(X_{ijkt}, Y_{ijkt,} a, b, d, f)$ is given as follows:

When $\sum M_{iikt} \times X_{iikt} \times Y_{iikt} \le N_{ikt}$ is:

$$\sum_{k}\left\{\begin{array}{l} \sum_{i}\sum_{j=1}^{16} X_{ijkt}\times Y_{ijkt}\times M_{ijk(t-1)}(1+d)\times XI_{jk(t-1)} + \\ \sum_{i}\sum_{j=17}^{37} X_{ijkt}\times Y_{ijkt}\times M_{ijk(t-1)}(1+d)\times XI_{jk(t-1)}\times(1+e) + \\ \sum_{i}\sum_{j=38}^{40} X_{ijkt}\times Y_{ijkt}\times M_{ijk(t-1)}(1+d)\times XI_{jk(t-1)}\times(1+f) + \\ \sum_{i}\sum_{j=41} X_{ijkt}\times Y_{ijkt}\times M_{ijk(t-1)}(1+d)\times XI_{jk(t-1)}\times(1+g) + \end{array}\right\} - \sum_{i}\sum_{j=1}^{16} X_{ijkt}\times Y_{ijkt}\times M_{ijk(t-1)}(1+d)\times C_{jk(t-1)}\times(1+h)$$

(29)

When $\sum M_{iikt} \times X_{iikt} \times Y_{iikt} > N_{ikt}$ is:

$$\sum_{k}\left\{\begin{array}{l} \sum_{j=1}^{5} N_{jk(t-1)}\times XI_{jk(t-1)} + \sum_{j=6}^{7} N_{jk(t-1)}(1+a)\times XI_{jk(t-1)} + \\ \sum_{j=8}^{16} N_{jk(t-1)}(1+b)\times XI_{jk(t-1)} + \sum_{17}^{37} N_{jk(t-1)}(1+b)\times XI_{jk(t-1)}\times(1+e) + \\ \sum_{38}^{40} N_{jk(t-1)}(1+b)\times XI_{jk(t-1)}\times(1+f) + \sum_{j=41} N_{jk(t-1)}(1+b)\times XI_{jk(t-1)}\times(1+g) \end{array}\right\} - \sum_{i}\sum_{j}\sum_{k} X_{ijkt} Y_{ijkt} C_{jik(t-1)}(1+h)$$

(30)

5. Constraints

The second problem builds upon the constraints established in the first problem. Therefore, the constraints in this study are the same as those in the first problem, as detailed in the previous sections.

4.2.3. Model Solution

First, the Sample Average Approximation (SAA) method is used to convert the stochastic programming model into a deterministic model. Then, a genetic algorithm is designed to solve the model.

Sample Average Approximation (SAA) Method

The SAA method utilizes the probability distribution of random variables to draw samples and compute their average values. This approach approximates the chance constraints and objective function in the stochastic optimization model. It helps to more accurately simulate and solve optimization problems involving uncertainty.

The main idea of the SAA method is to perform N_k random samples of the variable ω, resulting in a set of independently and identically distributed samples $\omega_1, \omega_2 \ldots \omega_i \in \Omega, i = 1, 2 \ldots N_{k.}$ When $k \to \infty$, $N_k \to \infty$, the sample average approximation function can be constructed as follows:

$$\hat{f}_N(x) = \frac{1}{N} \sum f(x, \omega_i) \approx E[f(x, \omega)] \quad (31)$$

1.The specific implementation steps are as follows:

Step 1: Generation of Random Variables and Samples.

Generate samples for the sales growth rate of wheat and corn.For other crops, generate samples for their respective growth rates.Generate samples to simulate the fluctuations in crop yield across different years.Generate samples for changes in planting costs to estimate future planting costs.

Step 2: Constructing Sample Scenarios.

Create a scenario for each set of generated samples.For example, one scenario might include a high growth rate for wheat, low yield fluctuations, and moderate cost reductions.

Step 3: Redefining the Objective Function and ConstraintsEach scenario uses its respective random sample values to recalculate the objective function and constraints.

Specific values such as sales growth rates, yield fluctuations per acre, etc., are substituted into the model to compute expected revenue, costs, and other economic indicators.

Step 4: Solving Each Deterministic Model.

Optimization techniques are applied to solve the deterministic problem for each scenario.

Step 5: Calculating Statistical Averages

The solutions from all optimization scenarios are averaged to obtain the expected objective function value.This value represents the average optimal profit.

2. Further Optimization Using Genetic Particle Swarm Algorithm

This study employs the Genetic Particle Swarm Optimization (GPSO) algorithm, integrating the Genetic Algorithm (GA) into Particle Swarm Optimization (PSO) to avoid local optima. PSO's global search identifies optimal solutions, while genetic operations (selection, crossover, mutation) refine them.

The process begins with parameter initialization and population creation. Particles are evaluated, sorted by fitness, and updated using the PSO formula. New particles are generated through mutation, and the best-performing half is retained. The remaining particles undergo crossover to enhance diversity. Elite selection preserves top solutions, and the process repeats until termination criteria are met, yielding the optimal solution.

3. Solution Results

The obtained total profit is 58,078,141 yuan.

References

[1] Zhang Chunqin. Fuzzy stochastic optimization model and method for crop planting structure [J]. Anhui Agricultural Science, 2019, 47(18): 243-246+250.

[2] Zhang Fan, Guo Ping, Ren Chongfeng. Application of fractional two-stage stochastic optimization model in crop planting structure optimization [J]. China Rural Water and Hydropower, 2016, (09): 111-114.

[3] Zhang Yan. Expectation value model and sample average approximation method for solving stochastic bilevel programming problems [D]. Liaoning University, 2019.

[4] He Yunlong. Smoothing and sample average approximation methods for solving stochastic symmetric cone complementarity problems [D]. Liaoning University, 2018.

*Co-first Author: Jialu Xu, Nier Wu

*Corresponding Author: Jialu Xu E-mail:xujialu@zymzdx.wecom.work