

Research on the Crop Planting Strategy Based on The Linear Optimization Model

Rongyu Qiao¹, Min Cai², Yifan Ding³, Yijun Yan^{2, *}

¹ School of Art, North China University of Water Resources and Electric Power, Zhengzhou, China, 450046

² School of Water Conservancy, North China University of Water Resources and Electric Power, Zhengzhou, China, 450046

³ School of Foreign Studies, North China University of Water Resources and Electric Power, Zhengzhou, China, 450046

* Corresponding Author Email: yyjmpp@163.com

Abstract. Amidst the backdrop of global population growth and climate change, optimizing crop planting strategies to enhance profitability and reduce costs has become a critical approach to addressing food security issues. This paper, based on agricultural planting statistical data from 2023, conducts preliminary data visualization, constructs multi-constraint models step by step, and employs linear programming, robust optimization algorithms, and other numerical solution methods to predict and analyze the optimization of crop planting strategies. Specifically, this study is based on the linear planning model, around the crop area for the decision variable, establish the target function, determine the constraints, optimization strategy after finally get two set scenarios under the next six years of the maximum profit of 41.6709 million yuan and 40.7512 million yuan respectively, unsalable all types of total profit is greater than the price of the total profit. Further, combined with the use of robust optimization model, solve the optimal planting scheme under extreme conditions, and give full consideration to the sales and price fluctuations, and the influence of seasonal change on the output of unit area, aims to achieve risk minimization and benefit maximization, optimization strategy for the final solution after seven years maximum total profit of 41.3328 million yuan.

Keywords: Data Profiling, Data Preprocessing, Linear Programming, Robust Optimization Model.

1. Introduction

In recent years, with the continuous growth of the global population and the intensification of climate change, food security has become a focal point of concern for governments and research institutions worldwide [1]. How to maximize crop yields on limited land resources to ensure the stability and sustainability of food supply has become a critical issue that needs to be addressed in agricultural production [2]. Against this backdrop, optimizing crop planting strategies to increase per-unit-area revenue and reduce production costs has become an effective way to improve agricultural production efficiency and economic benefits.

This paper, under the constraints of matching plot types with suitable crops, crop rotation planting, requirements for legume crop planting, other considerations in planting plans, the dispersion and size of plot areas, and other conditions, establishes an appropriate mathematical model. To simplify this study, it is assumed that the future expected sales volume, planting costs, yield per acre, and sales prices of various crops remain relatively stable compared to the current year, and crops planted each season are sold in that season. If the total output of a certain crop per season exceeds the corresponding expected sales volume, the excess cannot be sold normally. To more closely approximate the actual situation, this study sets two scenarios: in the first scenario, the excess is unsold, causing waste; in the second scenario, the excess is sold at 50% of the 2023 sales price. Optimal planting plans for crops in the countryside over the next seven years are provided for both scenarios. On this basis, this study considers the uncertainty of expected sales volume, yield per acre, planting costs, and sales prices of various crops, as well as potential planting risks under multiple constraints. It provides the optimal planting plan for crops in the countryside over the next six years.

2. Data Sources and Visualization

The data for this study is sourced from <https://www.mcm.edu.cn/>. Considering the diversity and complexity of the data, this study profiles the data at the beginning of model establishment. It uses auxiliary software such as Excel and Python for visual output of existing data to initially understand the overview of the model data and to better understand and use appropriate models to solve problems.

2.1. Analysis of Average Unit Revenue for Different Plots

By using Excel bar charts for data profiling analysis [3] on the "relevant data for 2023" in the table, the sales price range in the statistical data table is first processed into an average, and then the average unit revenue is calculated using the formula:

$$\text{per mu yield} * \text{selling price} - \text{planting cost} = \text{average revenue} \quad (1)$$

After calculating the average unit revenue, statistical analysis is performed with the corresponding plot types. From Figure 1 Average unit gain across plots, it can be concluded that the average return of ordinary greenhouses is relatively high. In contrast, the average unit revenue for terraced fields and hillside land is lower. This suggests that when optimizing the crop planting strategy in this study, profit maximization can be achieved by prioritizing the allocation of plots with high average unit revenue.

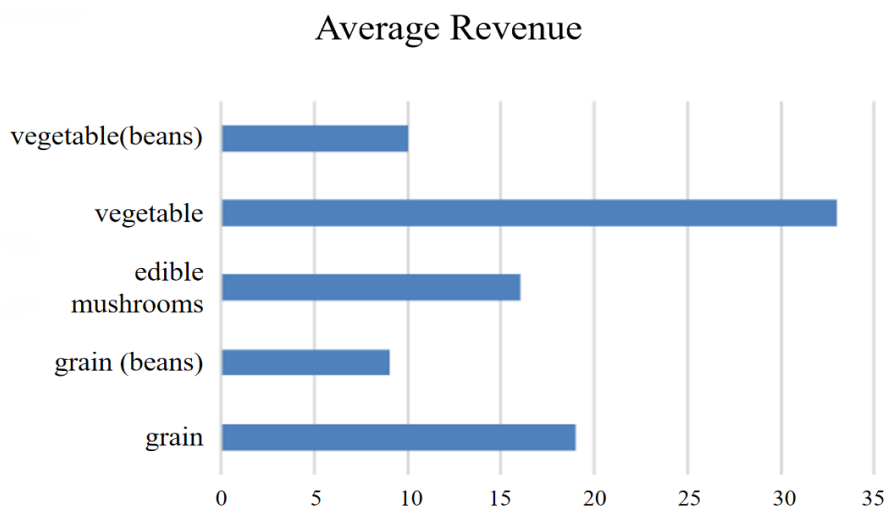


Figure 1. Average unit gain across plots

2.2. Analysis of Average Unit Revenue for Different Crops

After a comparative analysis of the average unit income of different plots, a statistical visualization of the average unit income of different crops can also find very useful information. Figure 2 Average unit income for different crops shows that edible greengrocery had the highest average unit income[4] and the lowest average units for cereals and legumes. This reflects that farmers do not earn high profits from grain cultivation. To improve the enthusiasm of farmers and ensure adequate national grain reserves, the government can appropriately provide preferential subsidies and policy support.

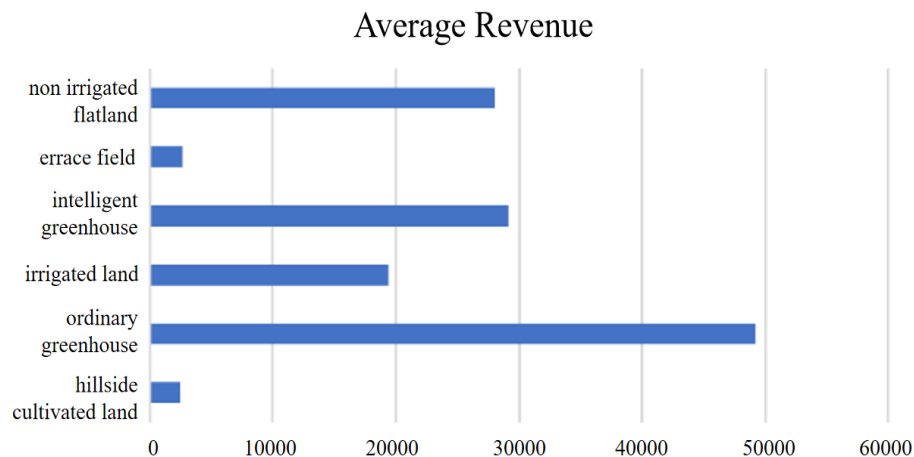


Figure 2. Average unit income for different crops

2.3. Crop Planting Category Richness

Excel's built-in visualization tools were utilized to analyze the collected statistical data on "Crop Planting Situation in 2023", resulting in Figure 3 Analysis of the richness and proportion of crop planting categories. The analysis of the statistical data indicates that grain crops account for the majority of crop types. This suggests that farmers have a high expectation for grain production and return rates in 2023. Grain, as the foundation and essence of the nation, is an important component of national security and people's well-being. It can be anticipated that the rational use of crop resources, optimization of planting land allocation, and overall planning of grain proportion are of utmost importance.

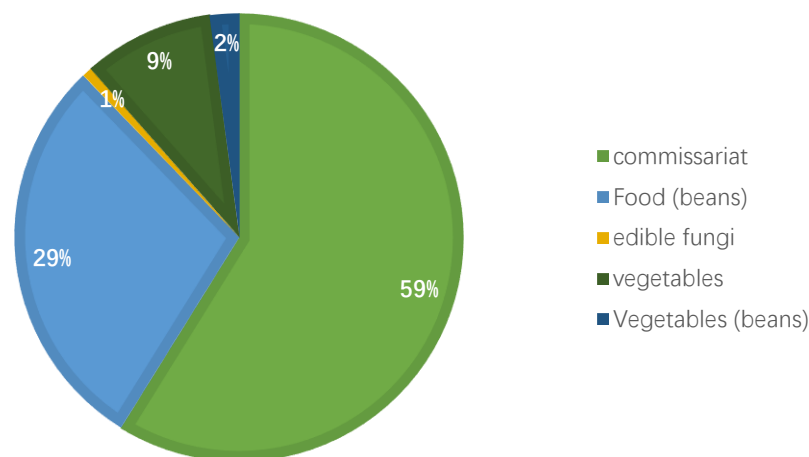


Figure 3. Analysis of the richness and proportion of crop planting categories

3. Establishment and Solution of the Crop Planting Strategy Optimization Model

For the establishment and solution of the rural planting strategy optimization model, data preprocessing based on the 2023 data is carried out, including data cleaning, data aggregation, sales price calculation, and generating a 0-1 matching table based on the data provided. Then, the processed data is analyzed using a linear programming model, with the planting area of crops on various plots as decision variables, to establish the objective function and determine constraints, including planting restrictions for different types of plots, crop rotation restrictions, legume crop planting requirements, and restrictions on crop yield and sales volume. The optimal planting plan for crops over the next six years is solved for the two possible scenarios.

3.1. Data Integration and Processing

Since the selection of planting plans is based on the 2023 data, preprocessing of the 2023 data is a crucial step. High-quality data not only directly affects the accuracy of the analysis results but also lays a good foundation for subsequent modeling.

For the collected data in this study, data cleaning is first performed to check for missing values and analyze whether there are outliers in combination with the actual situation of plot areas, planting areas, and crop cultivation. It is found that there are no missing or outlier values. Secondly, the cleaned data is aggregated by merging crop names and plot types to generate a merged primary key, integrating plot information and crop cultivation information into one table. Then, the total planting area and sales volume of each crop on each plot are calculated by grouping and summing according to crop type, plot type, and crop name. Subsequently, the sales price is calculated for the processed data. Since the sales price of each crop in different seasons and on different plots is a range and not an exact number, to obtain the sales price of each crop, this study defines a sales price calculation function. The sales price is obtained by generating a random number, and the calculation formula is as follows:

$$\text{sales_price} = \text{min_price} + \text{random_value} * (\text{max_price} - \text{min_price}) \quad (2)$$

Where min_price is the minimum value of the price range, max_price is the maximum value of the price range, and random_value is a number randomly generated between 0 and 1. Finally, based on the existing data provided on crop planting information, a 0-1 matching table for crops with plot types and seasons is generated.

3.2. Establishment of the Mathematical Model

On the basis of data preprocessing, this study also needs to clarify the objective function, which is a profit maximization model. The constraints include crop-plot matching, flat dry land, terraced fields, hillside land, irrigated land, ordinary greenhouses, smart greenhouses, crop rotation, etc.

3.2.1. Establishment of the Objective Function

According to the profit formula, which equals total sales revenue minus total planting costs, or equivalently, the total revenue from selling all crops minus the total planting costs minus the revenue from unsold parts, the following function can be written:

$$C_{ijk} = c_{ijk} * P_{ijk} \quad (3)$$

$$\max W_1 = \sum_{i,j} (P_{ijk} B_{ijk} X_{ijk} - C_{ijk}) - \sum_j D_{jk} B_{ijk} \quad (4)$$

$$\max W_2 = \sum_{i,j} (P_{ijk} B_{ijk} X_{ijk} - C_{ijk}) - \sum_j \frac{B_{ijk}}{2} D_{jk} \quad (5)$$

3.2.2. Constraints

Let J represent the set of crop types, I the set of plot types, k the planting season, $x_{i,j}$, the decision variable of planting crop type j on plot type i , $x_{i,j} = 1$ otherwise $x_{i,j} = 0$.

a. Crop-plot matching constraints

Based on the existing data on crops and cultivated land, it is known that various crops can only be planted in specific plots and seasons. Therefore, this study takes it as a fundamental premise for crop planting. To simplify the assumptions and facilitate the solution of the model, the study merges plots with seasons, transforming the three-dimensional relationship of crop-plots seasons into a two-dimensional relationship of crop-plots and seasons. A two-dimensional 0-1 matrix is obtained, showing which plots and seasons are suitable for each crop type. Using i to represent the plot and season, j to represent the crop, a_{ij} to represent the crop j cannot be grown on the plot and season of i , the constraint can be expressed as:

$$a_{ij} = \begin{cases} 1 & j \text{ cannot be planted in plot } i \text{ season} \\ 0 & j \text{ cannot be planted in plot } i \text{ season} \end{cases} \quad (6)$$

$$x_{i,j} \leq a_{ij} \quad (7)$$

It means that crops can only be grown in specific areas.

b. Plains, terraced fields, and sloping land constraints:

Maximize the profit objective function by using the build:

$$E(W_1^1) = E(T) \times E(P_{ijk}X_{ijk}) - E(C_{ijk}) \quad (8)$$

Set confidence level: Maintain yield [10] at 90% confidence level, can define constraints:

$$P_{ijk}X_{ijk}(W_1^1 \geq W_{1 \text{ threshold}}^1) \geq 0.9 \quad (9)$$

Finally, the optimization target is set to maximize the expected return: Maximize $E(W_1^1)$.

3.3. Mathematical Model Solving and Result Presentation

The first scenario of the research corresponds to the crop planting allocation results for the year 2024, as partially shown in Table 1 below. The first scenario, which corresponds to the crop planting allocation results for 2024 (Partial), yields a maximum profit of 41, 674, 373.09 yuan for the crops over six years.

Table 1. corresponds to the crop planting allocation results in 2024 for the first scenario (mu)
(Partial)

Seasonal	Land parcel name	Soybean	Black soya bean	Red beans	Mung bean	Vine grows beans	Wheat
Season 1	A1	0	0	0	0	0	0
Season 1	A2	0	13.75	0	0	13.75	13.75
Season 1	A3	0	0	0	0	0	10.5
Season 1	A4	0	0	14.4	14.4	0	14.4

The second scenario of the research corresponds to the partial example of the crop planting allocation results for the year 2024. The second scenario in Table 2 corresponds to the crop planting allocation results for 2024 (mu)(Partial), and the maximum profit of crops in the first scenario obtained by solving is 40756499.12 yuan within six years:

Table 2. corresponds to the crop planting allocation results in 2024 for the second scenario (mu)
(Partial)

Seasonal	Land parcel name	Soybean	Black soya bean	Red beans	Mung bean	Vine grows beans	Wheat
Season 1	A1	0	0	0	18	18	0
Season 1	A2	0	0	0	11	0	0
Season 1	A3	0	0	7	0	7	7
Season 1	A4	18	0	18	18	0	0

4. Modeling the Uncertainty of Economic Parameters in Crop Planting

Considering the use of linear programming models[5] for crop optimization strategies, which assume the linearity of the objective function and constraints, it is difficult to verify the correlation between data and the impact of future market uncertainties (such as sales volume, price, planting costs, and climate change) on planting strategies under the premise of diversified data. This simplifies the real-world problem and proposes further research, establishing a robust optimization model [6].

This model is used to solve for the optimal planting plan under the most adverse conditions, thereby maximizing the consideration of fluctuations in sales volume and price, the impact of climate change on yield per acre, and changes in planting costs [7], resulting in a planting plan that minimizes risk while maximizing profit.

4.1. Establishment of the Mathematical Model

4.1.1. Modeling of Uncertain Parameters

For each type of crop, define the range of uncertainty for sales volume, yield per acre, planting cost, and price: The sales volume growth rate for wheat and corn is 5%~10%, while for other crops, the sales volume changes by $\pm 5\%$. The yield per acre varies by $\pm 10\%$. The average annual increase in planting cost is 5%. The average price increase for vegetable crops is 5%. The price decline for edible fungi ranges from 1% to 5%, The drop in morels was 5%. At this point, the average return is calculated by -12%.

Using mathematical language, the uncertainties can be expressed as follows:

$$t_1 \in [0.05, 0.1] \quad (10)$$

$$t_2 \in [-0.05, 0.05] \quad (11)$$

$$t_3 \in [-0.1, 0.1] \quad (12)$$

$$t_4 = 0.05 \quad (13)$$

$$t_5 = 0.05 \quad (14)$$

$$t_6 \in [-0.01, -0.05] \quad (15)$$

$$t_7 \in t_6 = -0.05 \quad (16)$$

4.1.2. Select the robustness measure of uncertain variables and solve the objective function

a. Minimizes the worst-case losses

First, define the loss function:

$$L = \max[C_{ijk} - W_1^1] \quad (17)$$

The optimization target is then assumed: Minimize L [9].

b. Maximize the gains at a certain level of confidence

Maximize the profit objective function by using the build:

$$E(W_1^1) = E(T) \times E(P_{ijk}X_{ijk}) - E(C_{ijk}) \quad (18)$$

Set confidence level: Maintain yield [10] at 90% confidence level, can define constraints:

$$P_{ijk}X_{ijk}(W_1^1 \geq W_{1 \text{ threshold}}^1) \geq 0.9 \quad (19)$$

Finally, the optimization goal is set to maximize the expected return: Maximize $E(W_1^1)$.

4.2. Solving Mathematical Models and Displaying Results

The above process is also implemented using Python code, and the results of the planting strategy in 2025 are shown as follows in Table 3 Forecast results of crop planting distribution in 2025 (Partial). According to the optimization of the loss function, the average return of the optimization scheme is -4%, and the maximum total profit in seven years is 41,332,800 yuan:

Table 3. Forecast Results for Crop Planting Allocation in 2025 (Partial)

Seasonal	Land parcel name	Soybean	Black soya bean	Red beans	Mung bean	Vine grows beans	Wheat
Season 1	A1	0	0	0	0	0	0
Season 1	A2	0	0	16.5	0	16.5	22
Season 1	A3	7.875	0	7.875	0	0	0
Season 1	A4	0	0	0	0	0	23.4
Season 1	A5	0	0	0	0	13.6	0

5. Conclusion

This study reveals that an optimized reallocation of planting strategies in rural areas, supported by a research framework and applied to the field of agricultural economic optimization, can significantly enhance the efficiency and resilience of farming practices. By employing mathematical models such as linear programming and robust optimization, the study demonstrates the feasibility of arranging crop cultivation methods under multiple constraints. The integration of traditional models with realistic issues allows for a more refined comparison of optimization strategies, leading to more reasonable and reliable outcomes. The results indicate that the optimized plan exhibits a higher risk resistance and robustness, with an average revenue decrease of only 4% under conditions of significant market price fluctuations and seasonal change, compared to a 12% decrease with traditional planting schemes. In conclusion, these modeling outcomes robustly illustrate the effectiveness of the proposed optimization control strategies in the rational allocation of arable land resources and the maximization of sales revenue.

References

- [1] Guo D. Research on High-Yield Cultivation Techniques of Crops and Application of Agricultural Technology Extension [J]. Seed Science and Technology, 2024, 42 (17): 146 - 148.
- [2] Yu F-H, Zhang Y-Z, Wang L-Q. Analysis of the Optimization Path of Cultivated Land Protection Policy in the Face of Climate Change [J/OL]. China Land Economy, 1 - 16 [2024-10-23].
- [3] Zhou X-J, Tian B-W, Zheng H-F, et al. Research on Sewage Monitoring Sampling Data Based on Data Analysis [J]. Leather Making and Environmental Protection Science and Technology, 2022, 3 (23): 150 - 152.
- [4] Wang R-L. Analysis of the Attributes and Benefits of Land around Bus-Only Lanes Based on Big Data [J]. Transportation and Transportation, 2021, 37 (01): 11 - 15.
- [5] Liu S, Ren X-G, Wang S-X, et al. Linear Analysis of Lightweight Cryptographic Algorithm ACE and SPIX Based on MILP [J/OL]. Acta Electronica Sinica, 1 - 10 [2024 - 10 - 21].
- [6] Yu D-W, Yang M, Zhai H-F, et al. A Review of the Application of Robust Optimization in Power System Scheduling Decisions [J]. Automation of Electric Power Systems, 2016, 40 (07): 134 - 143+148.
- [7] Lv W. Analysis of Organic Agriculture Cultivation Techniques and Measures [J]. Contemporary Agricultural Machinery, 2023, (10): 56+58.
- [8] Fu J, Yang J, Wang Y, et al. SMRFnet: Saliency multi-scale residual fusion network for grayscale and pseudo color medical image fusion [J]. Biomedical Signal Processing and Control, 2025, 100 (PB): 107050 - 107050.
- [9] Zhou J-L, Zhang Y-G, Xiao Y, et al. Multi-Objective Path Optimization Model and Algorithm for Multimodal Transport under Uncertain Time [J/OL]. Systems Engineering and Information Science for Transportation, 1 - 17 [2024 - 10 - 23].
- [10] Wang N-N, Han S. Prediction Algorithm for the Reliability Confidence Interval of Wireless Communication Links Based on GA-Elman [J/OL]. Journal of Jilin University (Engineering and Technology Edition), 1 - 6 [2024 - 10 - 23].