# Research on Crop Planting Strategies in Rural Areas of North China Based on Bi-Level Optimization

Zeyang Zhang [1,*], Yichen Li [1], Yunshan Zhang [2]

[1] School of Statistics, University of International Business and Economics, Beijing, 100029

[2] UIBE Business School, University of International Business and Economics, Beijing, 100029

* Corresponding Author Email: 202357039@uibe.edu.cn

**Abstract.** To address the increasing demand for food and maximize both yield and economic returns in agricultural production, it is crucial to develop scientifically informed planting strategies for farmers. This study constructs a mathematical model based on Bi-Level Optimization theory to determine the optional crop planting combinations and predict future yields under the specific conditions of northern China. To enhance the model's comprehensiveness, a state transition function was incorporated, enabling a more dynamic representation of planting scenarios. Additionally, the integration of Monte Carlo simulation further equipped the model with the capability to identify optical solutions across diverse conditions. After model calculations, a detailed planting strategy and an expected income table were given. Overall, this research provides valuable insights into the design of planting strategies that simultaneously maximize yield and economic benefits, laying a solid foundation for future studies in sustainable agricultural planning.

**Keywords:** Bi-level optimization model, Genetic algorithm, Monte Carlo simulation, Rural sustainable development.

## 1. Introduction

Food is a fundamental aspect for human survival. And its supply significantly impacts societal stability and well-being. Since 1980, China's population has maintained growing, with a natural growth rate 11.87% (sources: National Bureau of statistics of China, statistical communique, 1980, https://data.stats.gov.cn/easyquery.htm?cn=C01&zb=A0301&sj=1980, accessed in Dec 1, 2024), necessitating an increasing demand for agricultural production. A stable and diverse food supply soon became the priority for the authorities. In the past few decades, the widespread use of chemical fertilizers enhanced crop yields. However, side-effects led to land deterioration not only by changing the chemical composition of the soil but also eliminating microbes that benefit the nutrition intake for crops, which contributes to declines in both quality and quantity of crops during harvest seasons. Existing practices for planting are increasingly inadequate to meet the growing needs for food while preserving land for sustainable use. Additionally, crop planting strategy involves the intersection of agronomy and statistics, which raises the difficulty of research and leads to almost no research in this area.

To sum up, most of the methods to improve crop yield through reasonable planning of crop planting area and cycle are based on empirical theories, rather than specific quantitative analyses. To conduct more in-depth research in this area and contribute to subsequent studies, this paper has carried out a preliminary exploration of this issue. Therefore, the paper investigates the problem and addresses finding an optimal growing solution of varied crops by using limited scale and different types of lands. At the meantime, farmers' income will be maximized. Since northern China is the nation's critical agricultural base, with about 21.06 million hectares of arable land (sources: National Bureau of statistics of China, statistical communique, annual by province, sown areas of major farm crops, 2022 https://data.stats.gov.cn/adv.htm?m=advquery&cn=E0103, accessed in Dec 1, 2024), the paper selected this area as a case to develop and test the proposed model.

In the selection of crop planting strategy, it is essential not only to consider the economic benefits of the crops but also to optimize the planting quantities of the same crop after the initial optimization. Otherwise, the planting of the same crop may become too dispersed, which is not conducive to efficient harvesting. The amount of planting obtained through profit maximization affects the dispersion coefficient of the crop, and the process of reducing the dispersion of the crop also affects the planting amount. In such problems, where two solution sets mutually influence each other, a bi-level optimization approach is widely adopted [1].

Bi-Level Optimization (BLO) is a hierarchical mathematical program where the feasible region of one optimization problem is constrained by the solution set of another optimization problem. The outer optimization task is commonly referred to as the Upper-Level (UL) problem, while the inner optimization task is commonly referred to as the Lower-Level (LL) problem. BLOs involve two kinds of variables, corresponding to the UL and LL problems respectively. In this problem, a BLO is constructed by setting profit maximization as the UL problem and crop dispersion minimization as the LL problem [2]. Furthermore, to enhance the comprehensiveness of the study and align with the principles of sustainable cultivation, the crop planting strategy planning is extended from one-year horizon to ten- year horizon. In this paper, a state transition equation is established to complete the iterating process across years and the aforementioned BLO fusion is utilized to construct an improved dynamic optimization mathematical model [3-8]. Finally, to address the inherent complexity, dynamism and uncertainty of agricultural parameters, the enhanced model also incorporates with Monte Carlo simulation, a computational algorithm that uses random sampling method for computing the results [8-11], to find the optimal solutions in all scenarios.

In conclusion, the research into this problem offers significant practical implications by providing a well-built framework for optimizing planting strategies, which assists farmers to better design their plans while reaching the goal of maximizing revenue and minimizing the planting dispersive rate at the same time. Due to the lack of attention to this domain in earlier research, this paper serves as a basis for future studies. From a methodological perspective, this paper employs relatively fundamental and widely used mathematical models, making it possible to conduct follow-up research to enhance and optimize the model herein.

## 2. Methodology

### 2.1. Description of the Dataset

The research is grounded in statistical data collected in 2023 from a representative village in northern China. The data includes 2 primary variables: (1) land types, flat and dry land, terraced fields, hillsides, irrigated land, ordinary greenhouses and smart greenhouses. (2) crop species, comprising 41 species in total, categorized into 6 groups according to the biological relationship. At the planting side, the paper considered 6 types of lands and 41 types of crops to represent the main cultivated land types and main crops of agriculture in North China. The six types of land are flat and dry land, terraced fields, hillsides, irrigated land, ordinary greenhouses and smart greenhouses; And the 4l crops cover grain, vegetables, edible fungi and beans. The detailed data of Chinese cultivated land use in 2023 is available at https://www.stats.gov.cn/sj/ndsj/2023/indexch.htm. And the unit adopted in the context to measure the size of lands is Mu, a Chinese unit of area, approximately 0.0667 hectares, description available at https://www.wikidata.org/wiki/Q836651.

### 2.2. Conditions and Assumptions

To meet the objective conditions of crop growth, ensure the reasonable planting of various crops and make the crop planting strategy more realistic, the paper put forward the following objective conditions:

1) Due to the climate in North China, crops can only be planted once on flat dry land, terraces and hillsides a year. Rice can be planted twice a year on irrigated land. Vegetables can be planted twice a year on irrigated land or smart greenhouses, and once a year in Ordinary greenhouses. Edible fungi can only be planted once a year in ordinary greenhouses.

2) Crops are prohibited from being continuously planted on the same land, except for vegetables grown in irrigated land, ordinary greenhouses and smart greenhouses. In other cases, the same type of crops cannot be planted on the same land in the following year to maintain land fertility.

3) Cabbage, white radish and carrot cannot be planted in the smart greenhouse.

Then, to optimize the crop planting strategy, the paper also set up additional planting constraints:

1) To ensure fertility, all soil must be planted with beans every three years.

2) Crops should not be excessively dispersed, and the planting area for each crop on a single plot of land should not be too small.

Next, on the sales side, the paper intends to find out the relationship between selling price and sales volume as well as cost and yield. Obviously, the expected sales volume cannot be fixed for each of the next seven years, so we must make our optimal solution as consistent as possible with market fluctuations. By referring to the data of China Statistical Yearbook for a total of 10 years from 2014 to 2023, this paper puts the average sales volume fluctuation of all crops at 5%.

At the same time, to make the results of this paper more general, the following assumptions are designed to avoid the interference of extreme cases:

Rational planting: Farmers can plant rationally and predictably.

Maximize profits: The goal is to maximize farmers' profits.

Fixed planting area: The land available for planting each year is constant and will not change due to uncertainties.

Limited risk impact: Potential planting risks will not lead to large-scale crop loss or even collapse.

Offset: The output is always roughly the same as the market demands: that is, almost al1 the crop output can be consumed.

Land diversity is stable: Different types of land possess varying levels of fertility, which affect the per-Mu yield of the same crop, and these differences remain stable from 2024 to 2030.

Crop-related factors are stable: The expected sales volume, yield, planting cost and selling price of crops will not fluctuate sharply from 2024 to 2030, such as the planting cost soaring due to fertilizer shortage.

In fact, the purpose of this paper is to solve the cycle of "What crops can be planted in this field this year" --"what crops should be planted here this year" -- "How many crops should be planted here this year"--"What crops should be planted next to this field" --"What crops should be planted here next year.

## 2.3. Building Mathematical Models

1) Build a Bi-level multi-objective optimization model

As mentioned above, the goal of this paper is to maximize farmers' income. At the same time, to ensure that the planting areas of the same crop are close to each other and avoid over-fitting of the algorithm leading to too scattered crop planting, this paper introduced the crop dispersion coefficient and solved the minimum value of the coefficient at the same time to constrain the solution of income.

However, the paper found that if the two sets of functions were directly solved by nonlinear programming, it would lead to nonlinear non-convexity problems, and the optimal solution could not be obtained. The paper then found that the minimization of crop planting coefficient and the maximization of total revenue had hierarchical, independent and conflicting characteristics, which met the conditions for establishing a bi-level multi-objective optimization model. After analysis, it is found that the effect of maximizing the total revenue is better than that of minimizing the species planting dispersion coefficient.

Therefore, the maximum total revenue is set as the upper function, and the minimum crop planting dispersion coefficient is the lower function, and restriction conditions are added to make the two functions have solutions under the condition of maximum profit.

The function of bi-level multi-objectives model as follows:

Upper function:

$$\max\sum_{i=1}^{34}\sum_{j=1}^{41}[p_j \cdot Y_j \cdot \min(\sum_T A_{ij}, Q_j^{2023}) \cdot x_{ij}^t - C_{ij} \cdot \sum_T A_{ij} \cdot x_{ij}^t] \tag{1}$$

Where $P_j$ is the unit price of $j$-th crop, $Y_j$ denotes yield per unit area, $C_{ij}$ refers to the unit cost of planting $j$-th crop on $i$-th land, $\sum_T A_{ij}$ is the areas of $j$-th crop planted on six types of land. $x_{ij}^t$ is a constraint that defines feasibility of planting $j$-th crop on $i$-th land.

This objective function selects the smaller value of actual planting and expected yield to calculate the profit, ensuring that the portion of planting and production that exceeds the expected portion is not accounted for.

Lower function:

$$\varepsilon = \min\sum_{y=2024}^{2030}\sum_{i=1}^{54}\sum_{j=1}^{41}(Z_i - A_{ij}^y)x_{ij} \tag{2}$$

Where $\varepsilon$ is the minimum of crop dispersion index, $A_{ij}^y$ denotes the areas of $j$-th crop planted on $i$-th land in $k$-th year, $T$ refers to types of the land.

2) Establish the dynamic optimization equation

Subsequently, due to the limitations of conditions, considering that the optimization decisions of each year are more likely to give the optimization results of the previous year, this paper adopts a dynamic model through the state transition equation, and iterates year by year from the data of the first year.

Dynamic programming by state transition equation is a common strategy to solve problems with overlapping substructures and optimal substructures. The core idea of dynamic programming is to decompose a large programming problem into a series of sub-problems, and by storing the solutions of these sub-problems, the problems in each stage are solved step by step, and finally the global optimal solution is obtained. This method solves the excessive distortion of the two-year gap in the planning after discrete simulation of planting cost, unit price of sales, yield per mu and expected sales volume.

Following the objective mentioned above, we can determine a state transition equation:

$$Q_j^y = Q_j^{y-1} \cdot (1+\theta_j) \tag{3}$$

Where $Q_j^y$ denotes the sales volume of $y$-th year. $\theta_j$ is a coefficient. Generally, $\theta_j \in (-0.05, 0.05)$, when the crop is wheat or corn, $\theta_j \in (0.05, 0.1)$.

3) Fusion of bi-level optimization model and dynamic programming

In this way, we can form a new upper function through replacing the yield of following years with the state transition equation:

$$\text{Max}\sum_{y=2024}^{2030}\sum_{j=1}^{41}\sum_{t=1}^{2}[P_j^y \cdot \min(Q_j^y, x_{ij}^{ty} \cdot \sum_T A_{ij}^{ty} \cdot Y_{ij}^y) - C_{ij}^y \cdot \sum_T A_{ij}^{ty}] \tag{4}$$

4) Monte Carlo simulation

In this paper, considering the complexity, dynamics and uncertainty of parameters, that is, multiple constraints are in the form of intervals, in order to find the optimal solution in all scenarios, Monte Carlo simulation is used to solve. Monte Carlo simulation is a calculation method based on "random numbers", which is often used to solve complex problems involving random variables. The core idea of the method is to simulate the future scenario through a large number of random samples, and calculate the value of the objective function according to the results of each simulation, evaluate its distribution, and finally get the optimal solution.

## 3. Results

Through the optimization of the mathematical model, the final income is ￥857256.4, representing an increase of 52.15% compared to the overall income in 2023. Furthermore, both crop yields and land utilization are lower than those in 2023. This indicates that excess production has been largely optimized, land resources have been efficiently allocated, and profits have been increased. The projected planting area of some crops in 2024-2030 obtained by Monte Carlo simulation is shown in Table 1.

**Table 1.** Comparison of power load forecasting of 403 line

| Land type | Soybean | Black Soybean | Red Bean | Mung Bean | Runner Bean | Wheat | Corn | Millet |
|---|---|---|---|---|---|---|---|---|
| A1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 72.9 |
| A2 | 0 | 0 | 0 | 0 | 24.73 | 5.695 | 10.28 | 0 |
| A3 | 0 | 0 | 0 | 0 | 0 | 0 | 25.18 | 0 |
| A4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 71.58 |
| A5 | 0 | 0 | 0 | 0 | 0 | 0 | 36.85 | 30.66 |

Through analysis and literature review, it was found that a nitrogen-fixing endophytic bacterium isolated from *Dendrobium devonianum* significantly increased the fresh weight of test plants. Field experiments demonstrated that the fresh weights of maize and rapeseed treated with its fermentation broth were significantly higher than those of the water control group, with an increase of 59.96%. This indicates that nitrogen-fixing bacteria play a role in enhancing crop yield. Therefore, the paper deducted one of the reasons for increase in yield is probably microbe effects. It is given that leguminous crops host nitrogen-fixing bacteria, such as rhizobia, in their root nodules, planting legumes on any given land can promote the yield of other crops cultivated on the same land. According to the reviewed data, the yield of all crops can increase by approximately 25%. Moreover, some organic compounds naturally released by crops, such as Gibberellin and Ethylene, have been shown to play significant roles in regulating plant growth. These phytohormones can enhance various physiological processes, including cell elongation, seed germination, and fruit ripening, thereby promoting overall crop growth and development under certain environmental conditions.

Obviously, we achieved our goal of exceeding the original profit by using only part of the land. The remaining land can be used to grow other crops for more profit, ensuring that farmers have the ability to cope with uncertainties, such as large-scale weather disasters and locust infestations, in actual planting and production.

In addition, the comparison between the predicted data and the actual data shows that the model has strong forecasting performance, fast forecasting speed and high integration degree, which can provide accurate and valuable reference for crop planting strategy. With the help of this model, we can plan crops in a more scientific and practical way, to maximize the benefits of crop planting

## 4. Conclusion

This paper employs a bi-level optimization model based on linear programming for crop planting strategy, incorporating state transition equations and Monte Carlo simulation to handle uncertainties and offer practical solutions. Despite focusing on 2023 data and omitting international trade impacts, the research is significant for China's agriculture, aiming to enhance production efficiency, farmers' income, and food security, thereby contributing to rural development, common prosperity, and global challenges like sustainable development and climate change.

## References

[1] Qi J, Wang S, Psaraftis H. Bi-level optimization model applications in managing air emissions from ships: A review [J]. Communications in Transportation Research, 2021, 1: 100020.

[2] Bard J F. Practical bilevel optimization: algorithms and applications [M]. Springer Science & Business Media, 2013.

[3] Kistler T, Franz M. Continuous program optimization: Design and evaluation [J]. IEEE Transactions on Computers, 2001, 50 (6): 549-566.

[4] Kistler T, Franz M. Continuous program optimization: A case study [J]. ACM Transactions on Programming Languages and Systems (TOPLAS), 2003, 25 (4): 500-548.

[5] Bertsekas D P. Dynamic programming: deterministic and stochastic models [M]. Prentice-Hall, Inc., 1987.

[6] Zou F, Yen G G, Tang L, et al. A reinforcement learning approach for dynamic multi-objective optimization [J]. Information Sciences, 2021, 546: 815-834.

[7] Branke J, Schmeck H. Designing evolutionary algorithms for dynamic optimization problems [J]. Advances in evolutionary computing: theory and applications, 2003: 239-262.

[8] Chudasama B. Fuzzy inference systems for mineral prospectivity modeling-optimized using Monte Carlo simulations [J]. MethodsX, 2022, 9: 101629.

[9] Mitchell F J. Monte Carlo Simulation [M]. Nova Science Publishers, Incorporated, 2017.

[10] James F. Monte Carlo theory and practice [J]. Reports on progress in Physics, 1980, 43 (9): 1145.

[11] Mardani Najafabadi M, Taki M. Robust data envelopment analysis with Monte Carlo simulation model for optimization the energy consumption in agriculture [J]. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2024, 46 (1): 9436-9450.