Optimization Of Multi-Objective Crop Planting Strategies Based on Genetic Algorithms

Qitong Fang *

School of Electrical Engineering and Computer,Jilin Jianzhu University, Changchun, China, 130118

* Corresponding Author Email: fangqitong@student.jlju.edu.cn

Abstract. In order to prioritize agriculture and rural development, and to place rural revitalization in an important position, optimizing crop planting strategies in rural areas has become an important research issue in agriculture. In this paper, by integrating and analyzing the given crop data and field information, and by meeting the conditions of crop rotation, avoiding repeated planting, convenient tillage and field management required by the title, a practical mathematical model was established by combining genetic algorithms. Then, a corresponding program was written using MATLAB software. The model was solved, and the optimal planting strategy that can maximize the planting benefits in the future 2024-2030 was obtained, as well as the related data analysis table. This study not only provides a better solution to the problem of scarce farmland in rural areas, but also broadens the practical application scope of genetic algorithms and provides more thinking and possibilities for solving other problems.

Keywords: Genetic Algorithms, Crop Planting Strategy, Multi-Objective Optimization.

1. Introduction

Food security has become a global problem due to the scarcity of land for food cultivation due to the urbanization of rural areas. Therefore, for large agricultural countries, addressing this critical issue is an urgent priority. Similarly, China, which is in the process of rapid urbanization, is facing the problem of scarcity of agricultural land [1]. After scholars improving 3D spatial understanding [2] and detecting objects in varied environments from a bird's-eye view [3], it is important to highlight that in northern China, due to the shrinking of agricultural land, there is an urgent need to make effective progress in land utilization in order to face the problem of food security. Such approaches could be valuable in managing the scattered 1,201 acres of agricultural land in North China.

According to the actual situation of the countryside, making full use of the limited cultivated land resources, adapting measures to local conditions, and developing the organic planting industry are of great practical significance for the sustainable development of the rural economy. Selecting suitable crops and optimizing planting strategies is conducive to facilitating field management, improving production efficiency, and reducing planting risks that may be caused by various uncertain factors. Through the query of a large number of rural land information, the agricultural land information of a village in North China was found. The village is scattered into 34 plots of different sizes, including four types: flat and dry land, terraced land, hillside land and irrigated land. Flat and arid land, terraced fields and hillside land are suitable for planting one season of grain crops per year; Watered land is suitable for planting one crop of rice or two crops of vegetables every year. There are 16 ordinary greenhouses and 4 smart greenhouses in the township, each with a cultivated area of 0.6 acres. Ordinary greenhouses are suitable for planting one season of vegetables and vegetables and one season of edible fungi every year, and smart greenhouses are suitable for planting two seasons of vegetables every year. Different crops can be planted in the same plot each season.

Therefore, this study raises the following research questions (RQs):

RQ1: Given the diversification of land types and the limited availability of agricultural land in mountainous regions, how can this study optimize agricultural yields per season to ensure that if the total yield of a crop exceeds its expected sales volume, the surplus is minimized, assuming stable future sales volumes, planting costs, and selling prices compared to 2023?

RQ2: According to the survey, the expected sales of wheat and corn in China in the future have a tendency to increase [4]. Considering the increasing future demand for staples like wheat and corn in China, coupled with fluctuating sales volumes for other crops, how can agricultural planning be adjusted to mitigate the risks associated with cost volatility and unsold inventory?

RQ3: In most cases, there may be some substitutability and complementarity between various crops [5]. In what ways can the substitutability and complementarity between different crops be leveraged to optimize the correlation between expected sales volumes, sales prices, and planting costs, thereby enhancing overall agricultural profitability?

To address the RQs, this study implements Multi-Objective Optimization and Genetic Algorithms:

(1). Optimization through Genetic Algorithms: Through simulation of 30 populations, cross-inheritance, and fitness-based selection, this study continuously evolves the populations until the total profit stabilized, thereby identifying the most profitable planting strategy.

(2) Linear Programming and Genetic Algorithm Integration: The use of Pareto Optimality [6] in this multivariate optimization enhances the efficiency of the genetic algorithm.

(3) Leveraging Crop Substitutability and Complementarity: The differential evolution method was utilized to establish a linear regression equation, factoring in these interdependencies to refine the model with additional constraints and relevant proportional coefficients.

The multiple linear programming model considers the distribution of crop planting area and the maximum profit, and can be developed into the engineering optimization in the social field, such as network design optimization [7] and logistics distribution management optimization [8].

2. Problem Formulation

Various types of land and their corresponding areas are provided, as well as the planting requirements for various crops. this study provides detailed information on the planting and output of various crops in 2023.this study should first preprocess the given data, convert the areas and names of the 34 plots and 20 greenhouses into visual images, and check and integrate the relevant data of crops. After processing the data, this study can analyze it and establish a reasonable and useful mathematical model to effectively derive the optimal solution. The dataset is available at https://github.com/btxq-ily/annexyctd.

For RQ1, the general model of artificial neural network consists of four basic elements, which are:

(1) This study aims to maintain relevant crop data for the year 2023 and identify the optimal planting scheme that maximizes revenue, particularly when the total output of crops in a season exceeds their expected sales volume, even if some produce cannot be sold through conventional channels. To achieve this objective, this study first formulates a target function focused on maximizing revenue. Subsequently, this study establishes multiple constraints based on two key requirements: (1) Different crops may be planted in each season on the same plot of land. (2) No single crop can be continuously cultivated in the same plot (greenhouse) during any given season. By inputting necessary data into both a linear programming model and a genetic algorithm, this study optimizes the strategy effectively. The total output for each specific crop is calculated by multiplying its yield per acre by its respective planting area, setting this total equal to the expected sales volume for that crop. Once numerical values are established, this study constructs a target function aimed at maximizing revenue alongside six constraints designed to regulate planting conditions due to potential surplus unsold inventory. Furthermore, it is essential that crops planted within each season exhibit minimal dispersion to facilitate efficient farming operations and field management practices. This study designates the dispersal pattern observed in 2023 as a standard reference point for optimizing our planting scheme. Notably, any surplus can be liquidated at 50% of its selling price; thus, altering parameters within our target function while keeping constraints and modeling algorithms constant.

(2) The input signal and the linear signal are the combination of the signals for each input signal.

(3) The function of the nonlinear activation function: making the neuron output signal within a certain range.

In the data given in Annexes 1 and 2, the total planting amount of each crop in 2023 is expressed using multiple linear programming.

For RQ2, Taking the factors into consideration and the planting restrictions of this rural area, this study constructs the objective function of maximum crop returns and 16 constraint conditions. Through multiple linear programming model and genetic algorithm, this study dynamically adjusts the planting strategy and selected the scheme that could maintain high returns under the changing environment.

For RQ3, from complementarity, it can be inferred that certain crops can promote the growth of each other and have a certain correlation in the market (there is a bundling effect, which can be sold together to increase the overall income, the most typical is that the planting of legumes can increase the yield of other crops, and the total income of the two will increase).From the correlation, it can be inferred that with some uncertainties in the market and the actual situation, the sales of crops will change and have correlation. Through the analysis of historical data, this study transforms these complex changing situations into multiple linear regression models, constantly optimize the constraints, and solve the planting scheme under the changing environment.

3. Methodology

Through data outlier detection processing [9], this study can ensure that the data used in this paper is accurate and ensure that the model solves the optimal strategy with high accuracy. The large data prediction model for the user's electricity consumption is implemented in the Clementine software. By applying the genetic algorithm [10] in the optimization algorithm model, data iteration analysis is conducted to analyze the total profit under different scenarios, and additional constraints are added based on different problem types.

3.1. Objective Function

3.1.1 Profit Maximization Objective Function (RQ1)

The equation of the profit maximization is:

$$B = (\sum_{m,n,t,z} Q_{(n,t,z)} * L_{(n,t,z)} * S_{(m,n,t,z)} + x * Q_{(n,t,z)}[L_{(n,t,z)}S_{(m,n,t,z)} - R_{n,t,z}])_{max} \quad (1)$$

where $Q_{(n,t,z)}$ represents the sales price of the z crop in the t quarter of the n year, $L_{(n,t,z)}$ represents the per mu yield of the z crop in the t quarter of the n year, $S_{(m,n,t,z)}$ represents the area of the m plot to plant the z crop in the t quarter of the n year, x represents the sales coefficient exceeding part of the price reduction, and $R_{(n,t,z)}$ represents the expected sales volume of the z crop in the t quarter of the n year.

3.1.2 Profit Maximization Objective Function Considering Cost (RQ2)

Based on RQ1, this study considers the cost problem, so this study adjusts the objective function:

$$D = (\sum_{m,n,t,z} Q_{(n,t,z)} * L_{(n,t,z)} * S_{(m,n,t,z)} - C_{(n,t,z)}S_{(m,n,t,z)}])_{max} \quad (2)$$

where it represents the planting cost of the z crop in the t season of the n year, represents the area of the m plot to plant the z crop in the t season of the n year, represents the sales price of the z crop in the t season of the n year, and the mu yield of the z crop in the t season of the n year$C_{(n,t,z)}S_{(m,n,t,z)}Q_{(n,t,z)}L_{(n,t,z)}$.

3.1.3 Objective Function for RQ3

$$E = \sum_{m,n,t,z}(Q_{(n,t,z)} * L_{(n,t,z)} * S_{(m,n,t,z)} - C_{(n,t,z)} * S_{(m,n,t,z)} + \sum(\partial * S_{(m,n,t,z)} * S_{(m,n,t,z)} + \sum \frac{\beta *}{S_{(m,n,t,z)} * S_{(m,n,t,z)}}, \quad (3)$$

where represents the sales price of the z crop in the t season of the n year, the per mu yield of the z crop in the t season of the n year, represents the area of the m plot of land planted with the z crop in the t season of the n year.

3.2. Constraints

The planting area of each plot shall not exceed the available area:

$$\sum_{z} s_{(m,n,t,z) \le k_i, \forall i,n,t} \tag{4}$$

where k_i represents the total available planting area of Plot i.

All the land in each plot (greenhouse) is planted with legumes at least once in three years:

$$-10 \le \sum_{n=1}^{7} \sum_{t=1}^{2} x\,(m,n,t,z) \le 1, \tag{5}$$

where $x(m,n,t,z)$ represents whether the z crop is planted in the t season of the n year in the m plot, +1 is planted and -1 is not planted.

The actual production of each crop does not exceed the maximum sales volume:

$$L_{(n,t,z)} {}^{*} S_{(m,n,t,z)} \le R_{(n,t,z)}, \forall m,n,t,z \tag{6}$$

where $R_{(n,t,z)}$represents the expected sales volume of z crops in the t season of the n year.

Each crop cannot be planted in the same plot (greenhouse):

$$X_{(m,n,t,z)} + X_{(m,n,t+1,z)} = 0, \forall m,n,t,z. \tag{7}$$

The planting area of the crop in a single plot should not be too small:

$$S_{(m,n,t,z)} \ge P_i \ \forall m,n,t,z. \tag{8}$$

In order to facilitate the cultivation operation and field management, the crops should not be too dispersed in the single season planting area:

Let the total frequency $F(z)$ of the occurrence of the z crop on the planted plot, and the total number of plots of the z crop is G, and assume that the dispersion is derived from the data in 2023.

$$F(z) = \frac{G}{34}. \tag{9}$$

The planting area of each plot shall not exceed the available area:

$$\sum_{z} S_{(m,n,t,z) \le k_i, \forall i,n,t,\cdot} \tag{10}$$

where k_i represents the total available planting area of block i.

All the land in each plot (greenhouses) is planted with legumes at least once in three years:

In order to facilitate the cultivation operation and field management, the crops should not be too scattered in the planting area of a single season.

Suppose that the total frequency of the occurrence of the z crop on the planted plot, and the total number of plots of the z crop is G, and assume that the dispersion can be obtained according to the data of 2023:$F(z)F(z) = \frac{G}{34}$.

4. Result analysis and discussion

4.1. Results for RQ1

The results of seven years from 2024 to 2030 are regarded as a population by genetic algorithm, and a total of 30 populations are simulated. As shown in the Figure 1, the planting decision of each year is an individual, and the planting situation of each plot within the individual is a gene. Fitness function repair infeasible solutions, such as crop replanting.

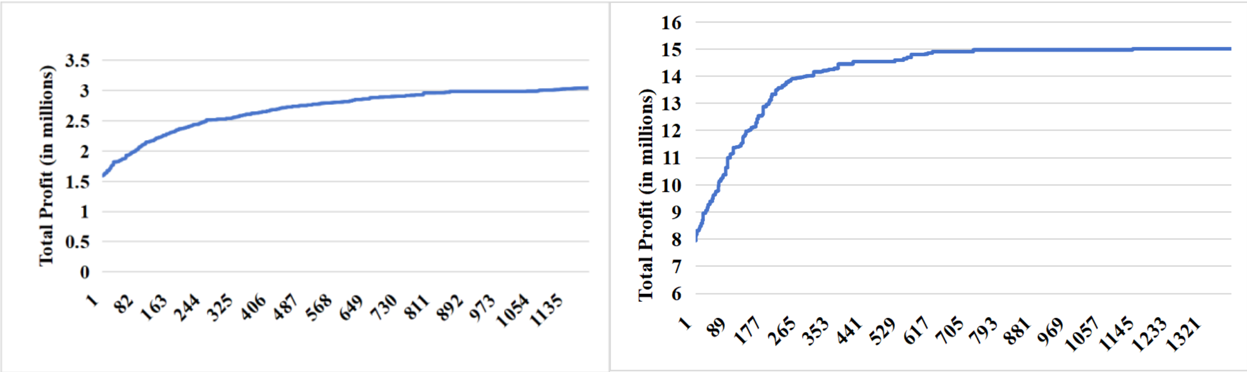

Figure 1. Iteration Curve for RQ1

Genetic algorithm was used to determine the population size of 30, the number of iterations for the first and second cases were 1200 and 1400 respectively, the crossover rate was 0.7, the variation rate was 0.3, the linear programming default single point crossover function was used, and the Gaussian variation was used as the variation function. As shown in the Table 1, the following are examples of single point crossover, convergence curve obtained by iterative process and partial iterative results of genetic algorithm.

Table 1: Results of Genetic Iteration Part

Number of iterations	Case 1 Convergence Curve	Case 2 Convergence Curve
1	1590313.265	7944578.833
2	1590313.265	7946651.83
3	1599752.075	8125092.638
4	1600748.707	8176010.702
5	1603629.214	8187230.499

4.2. Results for RQ2

Based on RQ1, RQ2 adds multiple stochastic constraints, so it is necessary to find the convergence domain of average profit and minimum profit iteratively to determine the optimal planting scheme under the condition of random variable parameters. Since the emergence of multivariate solution increases the computational load, Pareto Optimality is adopted to optimize the multivariate problem. The following is the resulting Figure 2.

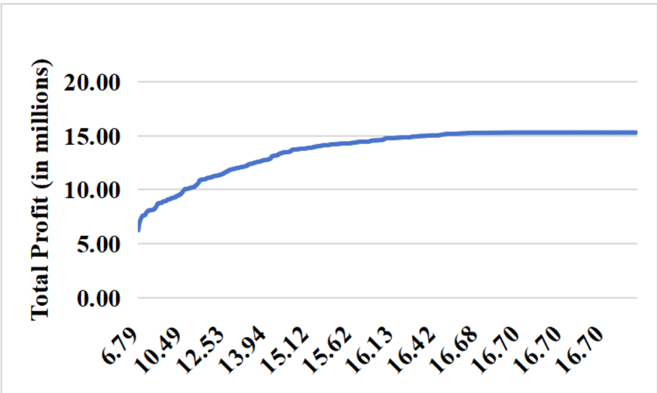

Figure 2. Iterative curve for RQ2

4.3. Results for RQ3

Using 1stOpt, the linear regression equation is obtained through the differential evolution method.

$$z=a+b/(1+((x-c)/d^2)+e/(1+((y-f)/g)^2) \tag{11}$$

Linear regression equation was obtained for correlation analysis (as indicated in Table 2).

Table 2. Correlation analysis results

	Sales volume	Selling price	Cost
Sales volume	1	-0.448	-0.513
Selling price	-0.448	1	0.19
Cost	-0.514	0.19	1

This is the correlation table for problem 3, which is solved under multiple constraints.

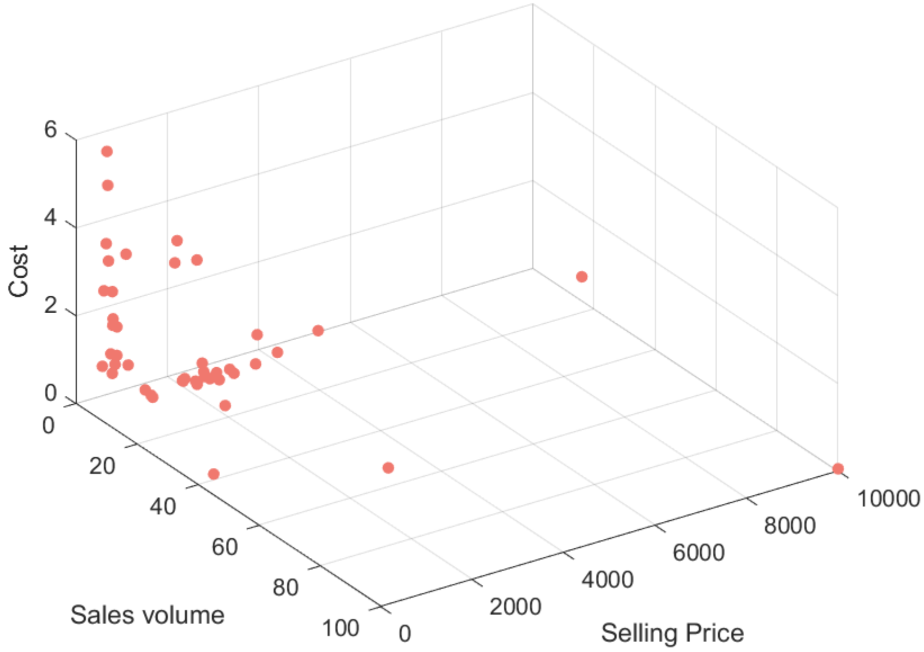

Figure 3. Graphs of analysis data on planting costs, expected sales volume, and selling price.

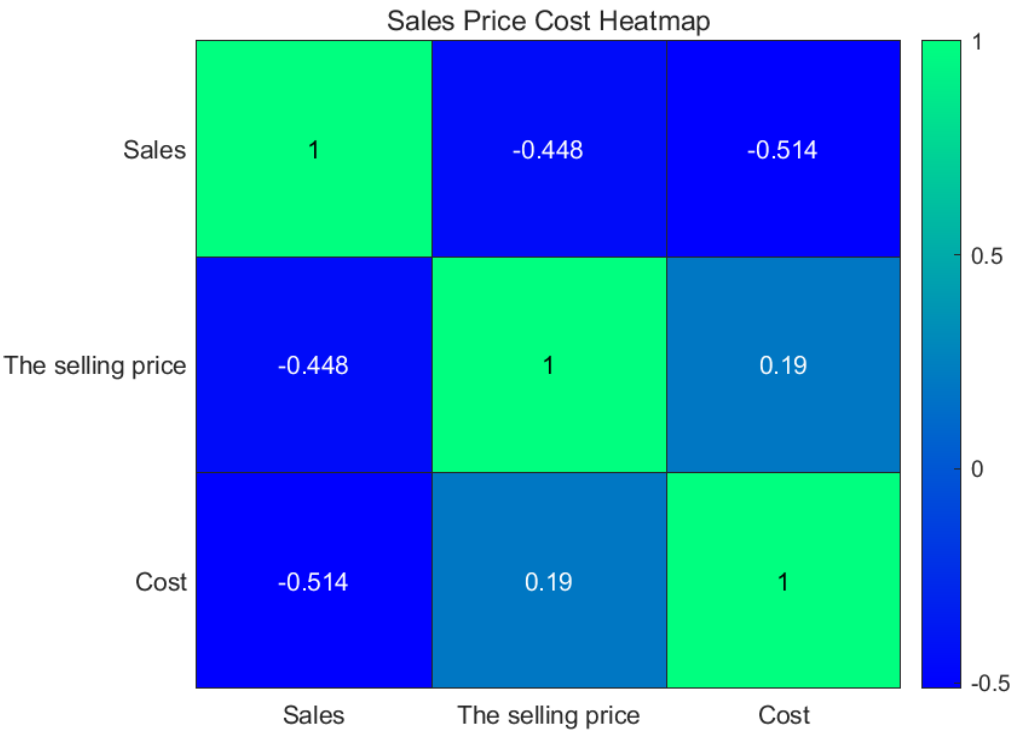

Figure 4. The heatmap of the correlation analysis results.

Table 3. Sales price, planting cost, and expected sales volume are some of the results.

Crop number	1	2	3	4	5	6
The price of Season 1	0	0	0	0	0	0
Sales of Season 1	0	0	0	0	0	0
The cost of Season 1	0	0	0	0	0	0
First Season Yield	0	0	0	0	0	0
The price of Season 2	3.25	7.5	8.25	7	6.75	3.5
Second Season Sales	42201.11868	28432.12954	10270.91655	10983.02762	11188.16934	49638.46136
Cost of Season 2	420	420	367.5	367.5	367.5	472.5
Second Season Yield	428	535	428	374.5	444.05	856

Through the correlation table, this study can obtain the data on the correlation between sales volume (as indicated in Table 2), cost, and price. The specific values of sales, costs, and prices presented in a three-dimensional graph (as indicated in Figure 3) can be intuitively accessed through a scatter plot. Through the heat map (as indicated in Figure 4), this study can intuitively analyze the different degrees of correlation between the three. The results of the partial solution, including the sales price, cultivation cost, and expected sales volume, are presented in Table 3.

5. Conclusions

This study addresses the challenge of limited arable land in certain regions by optimizing planting strategies to enhance agricultural yields. The methodology employs optimization models, including genetic algorithms, multi-linear programming, and linear regression, to tackle specific challenges through the definition of objective functions and constraints. The resulting optimized planting strategy improves the efficiency of arable land utilization, contributes to food security, stabilizes social order, and opens new avenues for enhancing land use practices. Furthermore, it broadens the applicability of genetic algorithms and encourages various sectors to leverage this technology for innovative advancements. This study integrates genetic algorithms with agriculture, providing a conceptual framework and theoretical foundation for future combinations. It facilitates a deeper integration of optimization algorithms with practical applications while offering solutions to diverse societal issues.

References

[1] Liu, Y., & Zhou, Y. (2021). Reflections on China's food security and land use policy under rapid urbanization. Land use policy, 109, 105699.

[2] Yao, J., Li, C., Sun, K., Cai, Y., Li, H., Ouyang, W., & Li, H. (2023, October). Ndc-scene: Boost monocular 3d semantic scene completion in normalized device coordinates space. In 2023 IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 9421-9431). IEEE Computer Society.

[3] Yao, J., Lai, Y., Kou, H., Wu, T., & Liu, R. (2024). QE-BEV: Query evolution for bird's eye view object detection in varied contexts. In ACM Multimedia 2024.

[4] Rattray, J., & Brokaw, S. C. (2012). The implications of the increasing global demand for corn. UW-L Journal of Undergraduate Research, 15, 1-10.

[5] Censor, Y. (1977). Pareto optimality in multiobjective problems. Applied Mathematics and Optimization, 4(1), 41-59.

[6] León, J. A., & Tumpson, D. B. (1975). Competition between two species for two complementary or substitutable resources. Journal of theoretical Biology, 50(1), 185-201.

[7] Neri, M., Guelpa, E., & Verda, V. (2024). Two-stage stochastic programming for the design optimization of district cooling networks under demand and cost uncertainty. Applied Thermal Engineering, 236, 121594.

[8] Zhang, L., Fu, M., Fei, T., Lim, M. K., & Tseng, M. L. (2024). A cold chain logistics distribution optimization model: Beijing-Tianjin-Hebei region low-carbon site selection. Industrial Management & Data Systems.

[9] Meng, Z., Wang, Y., Zheng, S., Wang, X., Liu, D., Zhang, J., & Shao, Y. (2024). Abnormal monitoring data detection based on matrix manipulation and the cuckoo search algorithm. Mathematics, 12(9), 1345.

[10] Bülbül, M. A. (2024). Optimization of artificial neural network structure and hyperparameters in hybrid model by genetic algorithm: iOS–android application for breast cancer diagnosis/prediction. The Journal of Supercomputing, 80(4), 4533-4553.