Transactions on Environment, Energy and Earth Sciences
ISSN: 2960-1924, eISSN: 2960-2300 | Volume 4
GEMFE 2024

Research on Crop Planting Strategy Problem Based on Optimization Greedy Algorithm

Zhetong Li [1], Peiqi Yuan [2], Sanglin Zhao [3,*], Hao Deng [3], Bingkun Yuan [4]

[1] School of Mathematics and Statistics, Hunan University of Finance and Economics, Changsha, China, 410205

[2] Department of Physics, Capital Normal University, Beijing, China, 100048

[3] School of Engineering Management, Hunan University of Finance and Economics, Changsha, China, 410205

[4] School of Energy Science and Engineering, Central South University, Changsha, China, 410083

* Corresponding Author Email: 202205640108@mails.hufe.edu.cn

Abstract. This article proposes a planting optimization strategy that comprehensively considers crop profits, planting costs, yield per mu, and mutual constraint effects based on crop economic and geographical data from rural areas in North China from 2024 to 2030. By combining greedy algo rithm, K-means clustering analysis, and Pearson correlation analysis, we ensure that crop selection is both efficient and sustainable, while adhering to crop rotation principles to avoid repeated cropping. The results show that the optimized planting plan has a total profit of 86382 538.00 yuan between 2024 and 2030, an increase of 211.22% compared to the original strategy profit. The study not only verified the scientific and feasibility of the model, but also provided specific planting strategy suggestions for agricultural production, aiming to improve agricultural production efficiency and promote the sustainable use of land resources.

Keywords: Greedy algorithm; Multi-objective-planning; K-means; Multiple-linearregression.

1. Introduction

According to the actual situation in rural areas, selecting suitable crops, adapting to local conditions and optimizing planting strategies are conducive to facilitating field management, improving production efficiency and reducing planting risks that may be caused by various uncertain factors. Zhang Can et al. [1] used cluster analysis to get the crop planting plan, while Yang [2] used deep learning algorithm to get the crop plan, However, none of these studies have considered the actual planting of crops and their influencing factors. In this paper, the economic and geographical data of various crops from 2024 to 2030 in a village in North China, China, combined with greedy algorithm of dynamic adjustment factors, K-means cluster analysis and Pearson correlation analysis, were used to determine the types of crops that can be planted, and the profit per unit area was calculated according to their sales price, planting cost and yield per mu. Give priority to planting crops with high profits, consider the principle of rotation, and follow the constraint that crops cannot be repeated, so as to ensure the fertility and sustainability of the land within three years. In this paper, the mathematical model of dynamic goal programming is constructed, and the parameters are set by combining cluster analysis and correlation analysis. Finally, the optimal planting strategy is obtained by using dynamic adjustment greedy algorithm. The profit after optimizing the planting strategy, the profit increased by 211% compared with the original planting plan. This paper optimizes and improves the greedy algorithm, and innovatively proposes the dynamic picking method, which completely considers the actual planting situation of crops, and gives a reasonable planting strategy.

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

2. Methodology

2.1. Literature review

Marta Cardin-Pedrosa Rosa and Carlos José Alvarez-López [3] established a model of agricultural production plan by investigating the planting industry in 53 indigenous people's settlements in Galicia, so as to provide suggestions for the government to formulate agricultural policies and plans. Drawing on the practice of Marta, this paper considers the influencing factors of crop planting.

Mauro Osakil [4] and others, on the basis of studying the unique double-crop production systems in Brazil, From the perspective of risk, the optimization model of agricultural production system under risk conditions-MOTAD model is established. The objective function of the model is to maximize the net operating income, and the constraints are the total planting area, the availability of machinery and cash flow. In this paper, we set the variable constraints and adjust the algorithm parameters.

Basil Manos and Parthenna Chatzinikolaou [5] and others modeled the decision-making process with the help of MCDM model on the premise of sustainable planting process and maximum production interest rate.[6]. The objective function of the model is the maximum gross profit rate, and the constraints are the total planting area, production right, reserved area, market, rotation and irrigation [7].This paper uses a greedy algorithm to simulate the continuous decision-making process.

2.2. Research on the substitutability and complementarity among various crops based on K-means clustering model.

K-means clustering is a widely used clustering algorithm, which divides the data set into k clusters by iteration, so that the similarity between data points within the cluster is high, while the similarity between data points between clusters is low. The core idea of this algorithm is to calculate the Euclidean distance and minimize the sum of the distances from each point to its cluster centroid.

Explore the substitutability and complementarity between crops, and explore the correlation between expected sales volume and price and cost, so as to modify the optimization model. By comparing the yield per mu, planting cost and sales unit price of crops, the similarity between crops is calculated. In this paper, K cluster analysis is used to find crops with similar economic characteristics, which can reveal which crops may be replaceable.

2.3. Regression modeling construction

By using multiple regression model, the relationship equation between expected sales volume, sales price and planting cost is constructed, and the objective programming model is solved by using its linear relationship [8, 9].

The multiple linear regression model is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon \tag{1}$$

The sum is the sales price and planting cost, and it is the expected sales volume. X_1 and X_2 for the sale price, Planting cost, Y for expected sales [10].

3. Experiment

This paper needs to build an objective function to maximize the sum of the total profits from 2024 to 2030, that is, to sum up the annual profits of each crop and finally summarize them.The following are the objective functions and the constraints:

Objective function:

$$Totalprofit = \sum_{t=2024}^{2030} \sum_{i=1}^{n} \sum_{j=1}^{m} profit^{t} + \sum j = \sum_{k=1}^{m} R_{jk} X_{ij} X_{ik} \tag{2}$$

Constraint function:

Because the sales of wheat and corn are increasing year by year, and the non-fixed value increases, the exponential function can be used.

Expected Sales of wheat and corn:

$$S_j^t = S_j^{2023} \cdot \left(1 + \gamma_j^t\right)^{t-2023}, j \in \{\text{wheat， corn}\} \tag{3}$$

Sales of other crops coincided with the growth in wheat and corn.

Expected sales volume for other crops:

$$S_j^t = S_j^{2023} \cdot \left(1 + \varphi_{\mathrm{j}}\right)^{t-2023}, j \in \{\text{wheat， corn}\} \tag{4}$$

Non-stubble continuous planting restraint:

Given the constraint that the crops cannot be planted continuously, the logical variable x is set. If *x=0* means that the crops are not planted in the plot, otherwise. Then the product of logical variables of adjacent plots must be 0.

$$x_{i,1,k,m} \cdot x_{i,1,k+1,m} = 0, \forall i,k; x_{i,1,k,m} \cdot x_{i,2,k,m} = 0, \forall i,k \tag{5}$$

Ordinary greenhouse constraints:

The greenhouse planting area of a single plot must not be too small. And can not exceed the greenhouse area.

$$\sum_{m\in[17,37]} x_{i,j,k,m} \text{剟} z_{i.j}, i \in \{E\}, j = 1; \sum_{m\in[38,41]} x_{i,j,k,m} \quad z_{i.j}, i \in \{E\}, j = 2 \tag{6}$$

Smart greenhouse constraints:

$$\sum_{m\in[17,34]} x_{i,j,k,m},, z_{i,j}, i \in \{F\}, j = 1,2 \tag{7}$$

Planting cost: As the sales volume increases exponentially year by year, and the sales volume is directly related to the planting volume, the planting cost also meets the increase year by year.

$$C_j^t = C_j^{2023} \cdot (1.05)^{t-2023} \tag{8}$$

Area constraints: The planting area of each crop shall not be greater than the total area of the zoning plot.

$$\sum_{j=1}^{m} X_{ij} \le A_i, \forall i \tag{9}$$

Rotation constraints: Bean crops need to be rotated annually, and the variable B is set as the logical variable of bean crop planting. If there is planting, the logical variable is 1, otherwise it is 0, so the sum of planting times in the three years between 2024 and 2026 must be greater than 1

$$\sum_{t=2024}^{2026} B_{ij} \ge 1, \forall i \tag{10}$$

Acre yield: The yield per mu fluctuates randomly with the season, increasing and decreasing.

$$Y_j^t = Y_j^{2023} \cdot \left(1+\delta_j^t\right)^{t-2023}, \delta_j^t \in [-0.1, 0.1] \tag{11}$$

Selling price of crops: According to the survey, corn sales in the region will decrease.

$$P_j^t = P_j^{2023} \cdot \left(1-\alpha_{\mathrm{j}}\right)^{t-2023}, \ \ \mathrm{j} \in crop \tag{12}$$

Variable description:

The annual growth rate is $r_j^t, r_j^t \in \{0.05, 0.1\}$.

(2) The price of edible fungi crops is declining every year, among which α_{j}, $\alpha_{\mathrm{j}} \in [0.01, 0.05]$.

According to the growth law of crops, each crop can not be planted continuously in the same plot (including greenhouse), otherwise it will reduce production;Since the soil containing rhizobacteria of leguminous crops is beneficial to the growth of other crops, it is required to plant leguminous crops at least once every three years from 2023. At the same time, the planting plan should take into account the convenience of farming operations and field management. For example, the planting area of each crop should not be too scattered every season, and the planting area of each crop in a single plot (including greenhouse) should not be too small.

According to experience, the expected sales volume of wheat and corn will increase in the future, with an average annual growth rate between 5% and 10%, and the expected sales volume of other crops will change by about 5% compared with 2023. The yield per mu of crops is often influenced by climate and other factors, and it will change by 10% every year. Due to the influence of market conditions, the planting cost of crops increases by about 5% every year on average. The sales price of grain crops is basically stable; The sales price of vegetable crops tends to increase, with an average annual increase of 5%.

The sales price of edible fungi has decreased steadily, about 1%~5% every year, especially the sales price of Morchella has decreased by 5% every year.

4. Result

4.1. Correlation test results

In this paper, the expected sales volume, sales price and planting cost are analyzed by Spearman correlation test, and the following results are obtained in Table 1:

Table 1. Pearson correlation test results

	Average sales volume	Average price	Total planting cost
Average sales volume	1 (0.000***)	-0.171 (0.286)	-0.121 (0.452)
Average price	-0.171 (0.286)	1 (0.000***)	-0.067 (0.676)
Total planting cost	-0.121 (0.452)	-0.067 (0.676)	1 (0.000***)

Note: * * *, * * and * represent the significance levels of 1%, 5% and 10% respectively.

4.2. Clustering results

According to the K-means clustering model algorithm, the following gene clustering trend is obtained:

From the above cluster analysis results, it can be seen that wheat, corn and soybean can be aggregated into one kind of crops according to the comprehensive analysis of the cluster results of total mu of crops and planting cost, so they are complementary. According to the clustering results of unit price of sales, water spinach and vegetables can be clustered into one category, so they have complementary economic benefits. Drawing from the practices of Lu Zhiyong et al, in solving the model, greedy algorithm sets wheat, corn, soybean, water spinach and vegetables with the same attribute variables [10].

4.3. Regression results

The coefficient test results are shown in theTable 2 and Table 3, and it can be seen that the P value corresponding to each regression coefficient is less than 0.05, so the original hypothesis is rejected and the regression effect is significant. X_1, X_2, Y.

Table 2. Inspection Table of Linear Regression Model

Linear regression model test result n=41						
	t	P	VIF	R²	Adjust r	F
constant	4.173	0.000***	-			
retail price	-1.132	0.265	1.005	0.047	-0.003	F=0.931 P=0.001
Total planting cost	-0.837	0.408	1.005			
Dependent variable: expected sales volume						
Note: * * *, * * and * represent the significance levels of 1%, 5% and 10% respectively.						

Table 3. Results of Multiple Linear Regression

The result of linear regression analysis is n=41			
	Non-standardized coefficient		Standardization coefficient
	B	Standard error	Beta
constant	23691.097	5677.348	-
retail price	-251.92	222.53	-0.18
Total planting cost	-0.143	0.171	-0.133
Dependent variable: expected sales volume			
Note: * * *, * * and * represent the significance levels of 1%, 5% and 10% respectively.			

According to Table 2, the multivariate regression model was successfully constructed by F test. There is a good linear relationship between the total planting cost, the sales unit price and the expected sales volume, so the correlation coefficient is set as the corresponding parameter in the actual greedy algorithm solution.

4.4. Planting strategy results

Based on Pearson correlation analysis and K-means cluster analysis of 3.1 and 3.2, the complementarity and substitution results among crops are obtained, and the potential correlation between crops is found. Then, based on the regression model of 3.3, the relationship equation of each index (between expected sales volume, sales price and planting cost) is found, and the objective function and constraint function in chapter3 are established by combining the above conditions. The simulation data is obtained to solve the optimized planting results. The following results are the results of the crop planting optimization scheme from 2024 to 2030 after comprehensively considering the relevant factors under each index. This result (Table 4) is a simple statistical result of crop planting optimization regardless of regional blocks, which is the sum of the planting area results under all types of plots under the optimized planting scheme:

Table 4. Table of planting area of crop planting optimization scheme from 2024 to 2030 after comprehensive consideration of relevant factors

Table of Planting Area (Table 4)							
crop	year						
	2024	2025	2026	2027	2028	2029	2030
cayenne pepper	29.60	27.90	26.51	25.70	24.77	23.24	22.86
water spinach	42.19	40.08	39.19	36.56	35.89	33.64	31.70
Yellow heart vegetables	42.33	42.33	40.22	37.59	36.38	35.20	33.33
celery	97.56	94.17	88.79	85.38	80.99	77.80	73.25
celery cabbage	37.73	36.06	36.17	33.24	31.10	29.99	28.98
ternip	116.45	112.41	106.79	104.91	98.27	95.23	91.53
garden radish	91.55	87.03	82.12	77.70	75.75	70.83	67.37
Elm yellow mushroom	0.00	0.00	0.00	0.00	0.00	0.00	0.00
lentinus edodes	0.00	0.00	0.00	0.00	0.00	0.00	0.00
pleurotus nebrodensis	0.00	0.00	0.00	0.00	0.00	0.00	0.00
toadstool	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 5. Table of planting area of crop planting optimization scheme from 2024 to 2030 after comprehensive consideration of relevant factors

Table of planting area							
crops	age						
	2024	2025	2026	2027	2028	2029	2030
soybean	24.67	23.65	22.30	21.51	20.27	19.48	18.53
black soya bean	30.97	29.64	28.34	27.40	25.83	24.57	22.58
red bean shrub	28.60	26.84	25.52	24.69	23.32	22.50	21.15
mung bean	24.49	23.38	22.18	21.10	20.65	19.45	18.24
Climb beans	28.90	27.79	26.85	25.21	24.06	23.10	22.11
wheat	43.67	41.78	40.80	37.80	36.66	34.07	33.04
corn	50.03	47.63	44.91	42.71	40.97	38.51	36.76
millet	27.26	26.17	25.46	23.92	22.43	21.51	20.66
kaoliang	39.06	36.47	35.24	33.59	31.54	30.93	28.94
broomcorn millet	36.77	35.34	33.10	31.73	29.76	28.16	26.69
buckwheat	7.72	7.13	7.09	6.70	6.55	6.00	5.90
pumpkin	74.20	71.69	66.43	65.03	60.69	59.04	55.73
sweet potato	26.94	25.73	24.44	23.52	22.72	21.21	20.64
naked oats	25.76	24.93	23.60	22.30	21.28	20.14	19.15
barley	37.33	35.42	33.89	32.24	30.80	28.88	27.33
paddy	11.55	10.82	10.17	9.94	9.21	8.75	8.48
cowpea	27.66	26.86	25.22	24.33	23.74	22.75	21.74
sword bean	35.49	34.40	32.88	31.47	29.07	28.69	27.34
kidney bean	27.90	26.94	25.08	24.20	23.79	22.75	22.26
potato	19.90	19.31	18.61	17.83	17.24	16.77	16.11
tomato	23.00	22.30	21.22	20.82	19.92	18.86	18.22
aubergine	53.89	50.72	48.49	46.50	45.82	43.58	42.22
spinach	22.80	21.76	21.23	20.37	19.62	18.93	18.28
green pepper	28.17	27.17	25.21	24.38	23.39	23.04	21.85
cauliflower	25.28	24.68	24.05	22.42	22.31	21.30	20.84
cabbage	24.52	23.32	21.98	21.57	20.86	20.24	19.26
Leaf lettuce	42.60	41.64	39.90	38.53	36.35	34.95	33.18
Small green vegetables	35.61	34.27	32.98	31.20	29.13	29.05	28.08
cucumber	68.39	64.13	62.26	58.68	55.93	53.30	51.75
lettuce	44.81	41.23	39.29	37.93	37.44	35.62	33.69

The following is the planting scheme of each crop under the optimized scheme-broken line trend chart: see Fig1.

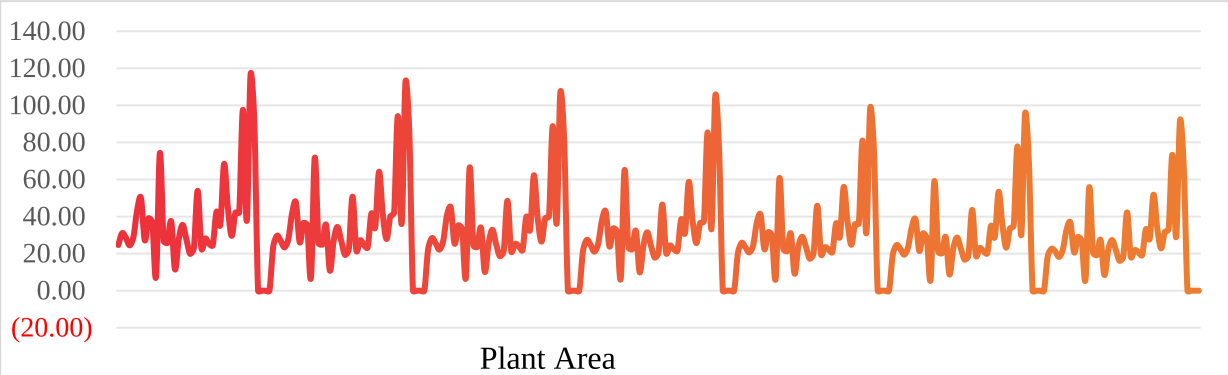

Figure 1. Trend diagram of planting area under crop planting optimization scheme from 2024 to 2030 after comprehensive consideration of related factors

4.5. Total Profit Result

The planting cost and sales profit of crops in 2023 are obtained through mathematical calculation, and the final optimized planting cost and sales profit of crops in 2024-2030 are obtained through the above multi-objective optimization model and greedy algorithm as follows as Table 6:

Table 6. Comparison of total profit before and after optimization

condition	time	Planting cost	profit	Annual average profit
Before optimization	2023	200580	5842558.25	5842558.25
After the optimization	2024-2030	1394146	86382538	12340362.57

According to the table 6, the average annual profit of crops in 2024 before optimization is 5842558.25 yuan, and after optimization by multi-objective optimization model, the average annual profit of crops in 2024-2030 is 12340362.57 yuan, with an increase of 211.22%, which shows that this optimization scheme has obvious advantages.

5. Conclusion

In this paper, the optimization strategy of planting strategy in actual agricultural production is taken as the research object, and on the basis of domestic and foreign research status and related theoretical research, combined with the analysis of the current situation of agricultural production laws in China, taking sales profit as the objective function and crop rotation and land resource occupation as the constraints, a decision-making model of agricultural planting production is constructed. Using the greedy algorithm in the optimization algorithm, the model is solved by python software, and the scientific and feasibility of the model is verified by model checking. This paper optimizes the greedy algorithm and explains the influencing factors of crop planting profit and its actual numerical influence. Put forward the relevant decision basis for agricultural production and life.

References

[1] Zhang Can, Ma Zhiyan, Cheng Yuan, et al. Obstacle Avoidance Path Planning of Robotic Arm with Greedy Algorithm and Improved RRT [J]. Mechanical Design and Research, 2024, 40 (03): 92-96.

[2] Yang Zichao, Wu Heng, Wu Yuewen, et al. Summary of deep learning training task scheduling based on performance modeling [J/OL]. Journal of Software, 1-21 [2024-09-08].

[3] Marta Cardin-Pedrosa, Carlos José Alvarez-López. Model for decision-making in agricultural production planning [J].Computers and Electronics in Agriculture, Volume 82, March 2012: 87-95.

[4] Mauro Osaki, Mário Otavio Batalha. Optimization model of agricultural production system in grain farms under risk, in Sorriso, Brazil [J]. Agricultural Systems, Volume 127, May 2014: 178-188.

[5] Basil Manos, Parthena Chatzinikolaou, Fedra Kiomourtzi. Sustainable Optimization of Agricultural Production [J]. APCBEE Procedia, Volume 5, 2013: 410-415.

[6] SU Zi-zhen, Chen Shaokuan, YANG An-an, et al. Optimization of unpaired train diagram for metro tidal passenger flow [J]. Journal of Railway Science and Engineering, 2024, 21 (06): 2183-2193.

[7] Wang Lili. Roaming path planning method of underground roadway based on improved RRT* algorithm [J]. Modern Electronic Technology, 2024, 47 (12): 62-68.

[8] Huang Wenhui, Wang Jianhui, Zhou Liping, et al. User demand-driven location method for 5G base stations [J/OL]. Computer Research and Development, 1-11 [2024-09-08].

[9] Ren Zongwei, Zhijun Qian, Zheng Wei, et al. Study on joint optimization of vehicle loading and distribution path under mass and volume constraints [J]. Packaging Engineering, 2024, 45 (09): 232-242.

[10] Lu Zhiyong, Liu Huibo, Ma Cuifang. Comparison of sorting methods for typhoon emergency box areas of automated container terminals [J]. Journal of Shanghai Maritime University, 2024, 45 (02): 46-52.