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

Optimisation of Crop Planting Strategies Based on Dynamic Planning Models

Jiaqi Wu [1, *, †], Zhe Cao [2, †], Weidi Zhai [3, †]

[1] Reading Academy, Nanjing University of Information Science and Technology, Nanjing, China, 210044

[2] School of Computer Science and Cyberspace Security, Nanjing University of Information Science and Technology, Nanjing, China, 210044

[3] School of Atmospheric Science, Nanjing University of Information Science and Technology, Nanjing, China, 210044

* Corresponding Author Email: gy808448@student.reading.ac.uk

† These authors contributed equally.

Abstract. In today's society, agricultural production faces many challenges, including population growth, resource scarcity, climate change and changes in market demand. In order to meet the growing demand for food, farmers need to optimise their cropping strategies to improve the efficiency of land and resource use for sustainable development. In this context, the optimisation of crop cultivation has gradually attracted extensive attention from academics and agricultural producers. Through the use of modern data analysis techniques, researchers can better understand the economic benefits of different crops and their interrelationships, and thus provide farmers with practical planting advice. In order to study to get the optimal strategy of crop planting in a certain countryside, this paper proposes a dynamic planning model to determine the optimal planting scheme of crops. In this paper, firstly, we take the profit as a clustering index, and use the K-means++ clustering algorithm to classify crops into three classes of high profit, medium profit and low profit. It is found that mushroom, vegetable and grain crops almost exactly correspond to these three profit classes, which provides effective theoretical support for the development of optimal planting strategies. Subsequently, for many problems in real life, this paper establishes a dynamic planning model, systematically solves a variety of practical constraints, and proposes the optimal planting scheme. Finally, this paper finds the optimal planting programme from 2024 to 2030 and the maximum profit to provide guidance for practical agricultural production activities.

Keywords: K-means++ clustering model; Dynamic planning optimal; Planting scheme for crops.

1. Introduction

The sustainable development of the rural economy is one of the most important issues of concern to society today. In this process, farmland resources play a crucial role, not only guaranteeing food security in the countryside, but also promoting economic development and ecological balance, which is the cornerstone for achieving sustainable rural development. However, in order to ensure the maximum profitability of the use of farmland resources, the crop cultivation programmes that currently exist still face major problems. For example, a study by Huazhong Agricultural University [1] used a linear programming model to explore the maximisation of profitability of agricultural supply, but only part of the influencing factors were considered and the planning model was complex, making it very difficult to solve. In addition, the study of Jilin Agricultural University [2] in the use of dynamic programming to maximise the profit of crop cultivation, only discusses the theoretical basis, not the actual situation in-depth elaboration. Therefore, it is necessary to establish a more reasonable crop cultivation optimisation scheme by combining modern data analysis techniques and dynamic planning models to address the actual situation of a specific region.

For this reason, this paper selects the planting data of a region in 2023 to optimise the existing planting strategy, with a view to proposing a more practical and effective improvement plan. By analysing the actual data, this paper hopes to provide scientific decision support for the sustainable development of rural economy. (Data from: http://cumcm.cnki.net.)

2. Analysis of optimal planting schemes based on dynamic programming

2.1. Standard unit of revenue treatment

In this paper, in order to summarise and analyse the mu yield, planting cost and corresponding selling unit price, profit/mu is used as the standard unit for judging the return of this type of crop, with the following formula:

$$G_i = S_i \times O_i - P_i \tag{1}$$

G_i represents the value of profit per acre for crops, S_i indicates the average selling price, O_i denotes yielding per acre for this crop, P_i represents the cost per acre of that crop.

In order to ensure the reasonableness of the selling prices, the paper averages the interval prices of the crops. This treatment processing is based on the overall level of the data, which ensures the universality and reasonableness of the standard price unit. At the same time, this paper draws a three-dimensional bar chart for visual display as shown in Figure 1:

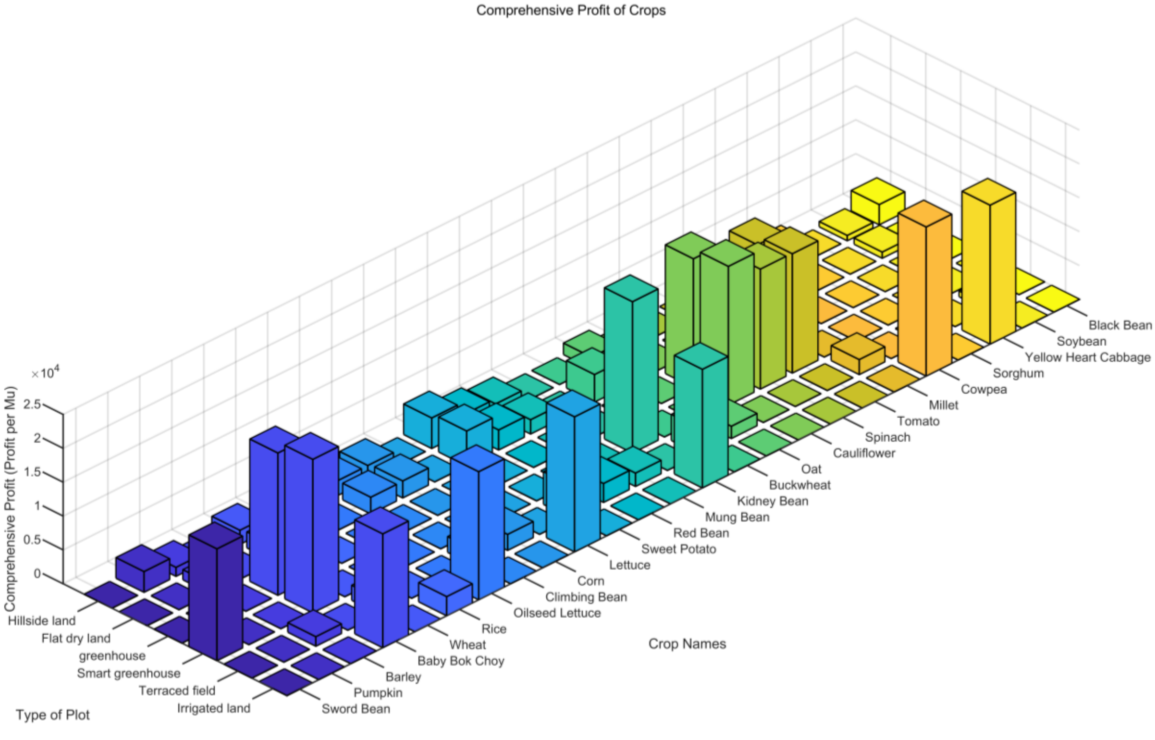

Figure 1. Three-dimensional histogram of integrated crop profits

As can be seen from Fig. 1, the profit value of the same crop in different plot types is different, so it is possible to find out the corresponding best suitable land for the same crop in order to determine the optimal planting programme subsequently.

2.2. Crop identity label codes

This paper identifies some features, such as: planting season, arable land type, crop type, etc. are key indicators affecting the final decision, and the indicators are qualitative data, this paper is corresponding to the label code for the crops, to facilitate the implementation of the final decision function and the selection of optimal planting plan for the crops. Take maize as an example: the label coding of the corresponding features is obtained as follows:

$$C(Crop\ number，Planting\ season，Land\ type，Crop\ type) = C_{i\,k}^{j\,g} \tag{2}$$

i represents crop number; j represents the growing season. When $j=1$ represent single-season planting, $j=2$ represents second season planting; g are a list representing land types; k for crop type.

Coding crop data not only helps to determine the optimisation objective and the corresponding functional relationship, but also significantly simplifies the representation of qualitative features, thus laying a solid foundation for subsequent model building. The coding process transforms complex, unstructured qualitative data into structured numerical data, which can be effectively processed and analysed by machine learning models.

2.3. Subclassification of crops for profit characteristics using K-means++ algorithm

Based on each type of crops to select the sales profit in 2023 as an indicator of its subclass division of the data for preliminary screening, this paper uses K-means + + clustering method [3] to carry out a preliminary cluster analysis of crops.

2.3.1. Determine the best value as the initial clustering centre.

Since the cluster partitioning results of the K-means clustering algorithm are significantly affected by the initial number of cluster classes, the optimal number of cluster classes needs to be pre-determined by other methods to pre-determine the optimal number of cluster classes. However, the clustering results from the number of cluster classes determined by different methods may not be consistent. For this reason, in order to ensure the reliability of clustering centres, the elbow rule is used in this paper to find the optimal number of clusters.

The elbow rule usually employs the sum of squared errors to determine the initial number of cluster classes in K-means++ clustering method, as shown in Equation (3):

$$SSE = \sum_{i=1}^{n} \sum_{j}^{k} p^{(i,j)} = \| g^{(i)} - b^{(j)} \|^{2} \tag{3}$$

$g^{(i)}$ denotes the value of crop returns in the cluster category, $b^{(j)}$ is the mean value of all crop returns in the cluster class. The sum of squared errors (SSE) is the distance of all crop returns from the cluster centre, the smaller the SSE value, the better the clustering effect. As the number of clusters gradually approaches the optimal number of clusters, the SSE value will be significantly reduced at this point, we can determine the ‘elbow’ corresponding to the value of k.

The optimal k value of 3 is determined by the above elbow rule, so this paper divides different crops into high profit crops, medium profit crops and low profit crops according to the profit. According to this division method, the crops are divided accordingly.

2.3.2. Distance between individual samples of the method based on Euclidean distance.

The Euclidean distance is used to calculate the shortest distance of each sample from the currently existing clustering centre i.e. the distance from the nearest clustering centre, the larger the value, the greater the probability of being selected as a clustering centre. The formula is shown in (4):

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^{d} (x_{ik} - c_{jk})^{2}} \tag{4}$$

2.3.3. The sample is assigned to k clustering centres selected by the optimisation iteration.

$$\mu_j = \frac{\sum_{i=1}^{m} 1\{c^{(i)} = j\} x^{(i)}}{\sum_{i=1}^{m} 1\{c^{(i)} = j\}} \tag{5}$$

μ_i is the position of the centre point after j iterations, $c^{(i)}$ is the distance to the cluster centroid.

Recalculate the position of the centroid of each cluster, which is the mean of all the sample points in the class, i.e. the centre of mass of all the sample points in the class.

Steps 2 and 3 are repeated until some termination condition is met, e.g. no (or minimum number of) objects are reassigned to different clusters and no (or minimum number of) cluster centres are changed again.

Clustering results as shown in Table 1:

Table 1. Analysis Summary of Test Results

Group Size	Data Source (Average Comparative Variability)			F	P
	low profit (n=22)	Medium profit (n=15)	High Profit (n=4)	NAN	NAN
Profit	13280.439	105887.5	284500	206.9	0.00***

Note: *** indicates p<0.001, ** indicates p<0.01, * indicates p<0.05; no mark indicates not significant.

Through the external indicator F-value is 206.983, found that its P-value results for 0.000, performance and its significant, that is, this paper that there is a significant difference between the means of different groups of clustering that the clustering effect is good [4], this paper can be followed up by using the clustering results for the corresponding analysis. The clustering results are shown in Figure 2:

Figure 2. Heat map of crop clustering results

This paper found that high profit crops are usually mushroom crops, medium profit crops are mostly vegetables, and low profit crops are mainly grain crops. Specifically, high profit mushroom crops are suitable for planting in greenhouse growing land, medium profit vegetable crops are suitable for watered land, and low profit grain crops are suitable for planting on the other three types of land.

Enhancing the stability of the dynamic programming objective algorithm: limiting the study to the more profitable crops helps to reduce the impact of data noise or uncertainty and improves the stability and reliability of the optimisation algorithm. This ensures that the optimisation results have a higher practical application value.

3. Design of optimal planting scheme based on dynamic programming

Dynamic programming provides an efficient technique for solving complex optimisation problems by preserving intermediate results and the ability to solve in stages, it is able to deal with a wide range of problems with constraints and ensures that a globally optimal solution is found. This approach improves the comprehensibility and manageability of the problem through structured and visualised state transfer tables, making it a powerful tool for solving a wide range of practical problems.

Since crops are grown over a season or a year dynamic programming can effectively deal with these time-dependent constraints [5] and find the best planting strategy that satisfies all the constraints through an optimisation algorithm. Specifically, dynamic programming can be considered to solve the following three problems: the recropping problem [6], the legume planting constraints problem [7], and the stagnant marketing problem [8].

3.1. Define the profit state representation

Based on the above four subproblems a state representation of profit can be derived as

$$dp[i][j][g][y] \tag{6}$$

i represents crop number, j represents the growing season, g indicates the type of plot planted, y indicates year.

3.2. Determine the initial state of the recurrence relation

Firstly the dp state matrix is initialised to an all 0 matrix then the initial state is divided into the following three types for assignment:

(1) There is a mismatch between the adaptive relationships of crops and the corresponding cropping land types:

If the crop cannot be planted on the corresponding type of land then the data value of the corresponding matrix position is updated to -1. After this round of screening the matrix can be divided into 0/-1 matrix.

(2) There is a mismatch in the adaptive relationship between the crop and the corresponding season:

Similarly, in this round of screening, if the crop cannot be planted during the corresponding quarter, the matrix features are further compressed by updating the matrix nodes with a data value of 0 at the corresponding matrix position to -1 on the basis of the matrix obtained in (1).

(3) Initialisation of the profit state corresponding to planting crops in 2023:

Finally, in this round of updating, the points that are 0 in the updated matrix are filled in with the profit value of the corresponding 2023 crop, and if the crop has a profit value of 0 in this state in 2023, i.e., no ploughing is carried out in this condition, the value of the matrix point remains unchanged at the value 0.

3.3. Define the state transfer equation

(1) Solving the transfer equation for the heavy cropping problem

When the crop planted in the second quarter is selected, it is determined whether the profit of planting the crop in the first quarter is zero (a profit of zero means that the crop is not planted [9], which is expressed as shown in Equation (6):

$$dp[i][j-1][g][y] == 0 \tag{7}$$

i represents crop number, j represents the growing season, g indicates the type of plot planted, y indicates year.

If the above conditions are met then the current state can be temporarily updated to:

$$dp[i][j][g][y] = G_{C_{i\,k}^{jg}} \tag{8}$$

$G_{C_{i\,k}^{jg}}$ denotes the profitability of growing a ith crop on a g land type in the j quarter of the y year.

(2) Crop rotation requirements for legumes

The selection of whether a legume crop is required in a particular year is based on a combination of whether the land has been planted with legumes in the previous three years, i.e. whether the condition is met:

$$\sum_{k'=1}^{3}\sum_{k\epsilon i_g} dp[i_k][j][g][y_{k'}] > 0 \tag{9}$$

i_k represent the category number of the g plot where pulses can be grown, $y_{k'}$ indicated k' of three consecutive years.

If the above conditions are met and $G_{C_{i_k k}^{jg}}$ is larger than $dp[i_k][j][g][y_{k'}]$ then the current state can be temporarily updated to:

$$dp[i_k][j][g][y_{k'}] = G_{C_{i_k k}^{jg}} \tag{10}$$

(3) Lagging sales in excess of expected sales

For the expected sales volume stagnation problem, this paper adopts the method of selling all stagnant goods with 50% price reduction , for the case that the total production is greater than the expected expected sales volume: $O_y(i) \times G_i > O_i \times G_i$, it is necessary to update the state transfer equation in the dynamic programming as [10]:

$$dp[i][j][g][y] = \max\left((O_y(i) - O_i) \times 0.5 + \min(O_y(i) \times G_i, O_i \times G_i), G_{C_{i\,k}^{jg}}\right) \tag{11}$$

Otherwise update the state transfer equation to:

$$dp[i][j][g][y] = \max\left(\min(O_y(i) \times G_i, O_i \times G_i), dp[i][j][g][y]\right) \tag{12}$$

3.4. Getting the optimal planting scheme

Based on the establishment of the above dynamic programming equations, the optimal cropping programme for the crops in the village for the year 2024 is derived, which are demonstrated in Table 2.

Table 2. Crop Acreage Results Presentation

Crop	Land Type	Planting area	Crop	Land Type	Planting area
Black bean	Terraced field	55.2	Sweet potato	Arid land	19.6
Red bean	Terraced field	70.4	Oat	Terraced field	27.3
Mung bean	Arid land	66.0	Rice	Irrigated land	50.4
Runner bean	Hillside land	31.6	Cowpea	Irrigated land	13.6
Wheat	Terraced field	158.7	Sword bean	greenhouse	7.8
Corn	Arid land	158.3	Kidney bean	greenhouse	1.9
Millet	Terraced field	180.4	Potato	Terraced field	10
Sorghum	Terraced field	44.3	Tomato	greenhouse	2.9

4. Conclusion

In this paper, it firstly uses profit as the clustering index, and use K-means+ clustering algorithm to divide crops into three categories of high profit, medium profit and low profit, and find that fungus, vegetable and grain crops correspond to these three categories almost perfectly, which provides excellent explanatory power for the establishment of the optimal planting programme and planting strategy. After that, this paper adopts the dynamic programming model with profit orientation and takes into account the problems of heavy cropping, crop rotation of legumes, marketing of agricultural products, and interplanting and intercropping to give the optimal planting strategy for the next ten years in a village in North China. The algorithmic strategy in this paper reduces a large number of constraints in the initialisation and state transfer process of the dynamic programming model, simplifies the model solving steps, and at the same time significantly improves the efficiency of the model solving and ensures that the optimal solution can be found.

References

[1] Wen Hui. Research on profit maximisation strategy of fresh agricultural products supply chain under the domination of different subjects [D]. Huazhong Agricultural University, 2020.

[2] Sun LQ. Application of dynamic planning model in determining the optimal planting programme for crops [J]. Jilin Agriculture, 2015, (08): 84.

[3] LIU Jiahui, ZHANG Ping, CAO Jinyin, et al. Cluster analysis and fine management of disease cost based on K-means algorithm [J]. Health Economics Research, 2024, 41 (08): 37-40+44

[4] LI Tianyi, LIU Qinming. Construction of value assessment index system for paintings based on BERT-LDA and K-means clustering [J]. Software Engineering, 2024, 27 (01): 68-73.

[5] Ma Maosheng. Influence of factors of agricultural cultivation in the structure of agricultural cultivation [J]. Hebei Agricultural Machinery, 2023, (07): 142-144.

[6] JIN Lin, GUO Wenqi, ZHANG Peitong, et al. Effects of heavy cropping on growth and nutrient uptake of shallow-growing Chinese yam [J/OL]. Zhejiang Agricultural Science, 1-8 [2024-09-15].

[7] PU Xiao, ZHI Yuxin, LI Hongying, et al. Evaluation of the effects of different crop rotation patterns on nitrogen and phosphorus utilisation and economic benefits of rice in Chaohu Lake Basin [J/OL]. Journal of Anhui Agricultural University, 1-10 [2024-09-15].

[8] WANG Juncheng, HU Hongxia. Study on the Problem of Agricultural Products' Lagging Sales in Key Supporting Counties for Rural Revitalisation [J]. Zhongzhou Journal, 2023, (09): 46-54.

[9] Li Shuo. Optimisation of crop cultivation structure in the Weigan River Basin under water resource constraints [D]. University of Chinese Academy of Sciences (Chinese Academy of Sciences, Institute of Space and Astronautical Information Innovation), 2022.

[10] Fan Linlin. Multi-objective optimisation model and its application in crop germplasm resources statistics [D]. Changchun University of Technology, 2022.