Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 13, No. 2, 2024

Research on Agricultural Products Planting Strategies Based on Integer Programming and Stochastic Programming

Yingshan Liu[1, *], Kaiwen Hu[1], Yuhang Lei[2]

[1]School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China
[2]School of Automation, Nanjing University of Science and Technology, Nanjing, China
[*]Corresponding author: 922104780605@njust.edu.cn

Abstract: In this paper, by applying the algorithms of integer programming and stochastic programming, we have studied in depth the crop planting strategy of a rural area in North China. First, we constructed an optimization model based on integer planning and stochastic planning agricultural planting strategies by exhaustively analyzing the detailed data of cultivated land and crops in the area. Further, we introduced stochastic variables to simulate the market uncertainty and the risk in the planting process, so as to propose an optimal planting scheme based on robust optimization theory. In addition, Spearman's correlation coefficient and ordinary least squares (OLS) regression analysis were used in this paper to further optimize the planting strategy by taking into account the substitutability and complementarity among crops. Through the continuous iteration and refinement of these methods, this paper significantly improves the accuracy and practicability of the planting decision model, and provides a valuable reference strategy for scientific planting of crops.

Keywords: Integer Planning, Stochastic Planning, Market Uncertainty, Substitutability and Complementarity. Integer Programming, Stochastic Programming, Robust Optimization, Spearman's Correlation Coefficient, OLS Regression.

1. Introduction

With the continuous advancement of agricultural modernization, the scientific and reasonable development of crop planting scheme has become the key to improve the efficiency of agricultural production and help rural revitalization [1]. In this paper, by constructing an optimization model based on integer programming and stochastic programming, we constructed a solution for the optimal planting strategy of agricultural products [2,3,4]. Further, we deeply analyze the uncertainty and planting risk of the agricultural products market, including the fluctuation of sales volume, planting cost, mu yield and sales price. Combined with the idea of robust optimization, we proposed a more robust planting decision model [5,6]. Finally, by introducing Spearman's correlation coefficient and ordinary least squares (OLS) regression analysis, we considered the correlation and substitutability between crops to more comprehensively assess the planting benefits of different crop combinations, thus refining the planting scheme optimization model [7,8].

2. Crop Planting Strategies Based on Integer Programming and Stochastic Programming

Agricultural planting planning is a complex decision-making process involving a variety of factors such as plots, crops, and seasons. In order to help rural areas to develop a reasonable agricultural planting plan and a reasonable planting strategy for crops, we establish an integer planning model to analyze and bring more benefits to farmers in rural areas. The research objective of this paper is to formulate the optimal planting strategy by comprehensively analyzing factors such as crop yield, selling price, and planting cost.

We collected the size of each plot type and the type of crops that can be planted with the cultivated arable land in a rural area in North China, as well as the crop cultivation in 2023 and the mu yield, planting cost, and selling unit price of each crop in 2023. The acreage yield of various crops in flat dry land, terraced land, and hillside land is low, but the acreage yield of crops such as pumpkin, sweet potato, and maize is high. The mu yield of all kinds of crops in watered land, ordinary greenhouses and intelligent greenhouses are higher, among which the mu yield of cucumber and hollow vegetable is the highest. Therefore, the above crops can be considered as much as possible while planting.

Our goal is to maximize total profit for all years from 2024 to 2030, with profit equal to sales revenue minus costs. In order to simplify the model analysis, we assume that the expected future sales volume, planting costs, mu yield and sales price of various crops remain stable compared to 2023 and the seasonal crop is sold in the same season, the sales price of agricultural products obeys the normal distribution, the mu yield fluctuates within a reasonable range, and the yields of different crops are the same when they are planted in different plots. By reviewing the literature and information on the relationship between expected sales volume and mu yield, the expected sales volume in North China is about 90% of the yield.

In order to more closely match the reality, this paper assumes that the mu yield fluctuates up and down 5%, the use of decision analysis as well as stochastic planning, to assist in solving. We assume that the sales price is uniformly distributed, the acres of production and sales price is divided into optimistic, neutral, pessimistic three cases, optimistic case acres of production increased by 2.5%, neutral unchanged, pessimistic case acres of production decreased by 2.5%, the sales price of the optimistic case to take the upper quartile, the neutral case to take the average, pessimistic case to take the lower quartile

When total production exceeds the corresponding expected sales volume, the excess is stagnant and can be wasted. If the acreage exceeds the expected sales volume, only the expected sales volume is actually sold, and if the total production does not exceed the expected sales volume, the acreage is actually sold. Profit on sales can be expressed mathematically as:

$$\max Z = \sum_{i=1}^{n} \sum_{j=2024}^{2030} \left[\begin{array}{c} \sum_{k=1}^{2} \sum_{l=1}^{m} Y_{jl} a_{ijkl} (N_{jl} - C_{jl}) (1 - B_{jl}) + \\ \sum_{k=1}^{2} \sum_{l=1}^{m} M_{jl} a_{ijkl} (N_{jl} - C_{jl}) B_{jl} \end{array} \right] \quad (1)$$

where B_{jl} is a logistic variable, $B_{jl} = 0$ represents total production exceeding expected sales, $B_{jl} = 1$ represents total production not exceeding expected sales, Z is sales profit, i denotes the plot, j denotes the year, k denotes the quarter, l denotes the crop number, Y denotes the expected sales, whose value is 0.9 times the acreage, M is the acreage, n denotes the total number of plots, m denotes the total number of crops, alpha denotes the number of acres of the crop planted, and N denotes the sales price. C denotes cost.

Considering that the cultivation of the same crop should be concentrated as much as possible, we consider using the number of acres planted to reflect the degree of concentration of crop cultivation with. The larger the number of acres planted to a particular crop on the same plot, the more concentrated the crop is, i.e., a plot of land is planted to as many of the same crop as possible. This can be expressed in terms of the extreme difference in the number of acres planted with a crop, which is expressed by the formula:

$$\max(\max a_{ijkl} - \min a_{ijkl}) \quad (1)$$

which indicates the maximum difference between the maximum number of acres planted and the minimum number of acres planted in different planting plots for the 1st crop planted in the current season of the year. In this way, the degree of concentration of crop cultivation is determined.

In order to make the optimal crop planting scheme closer to the real situation in this countryside, we propose a rich set of constraints as follows:

No heavy cropping constraints

According to is drop research, each crop can not be planted in the same plot in a continuous heavy crop, which is also called continuous cropping, refers to the continuous planting of the same crop on a field. That is, the same crop cannot be planted in two consecutive years. In order to facilitate the distinction between the ith field j year k season with or without crop l planting, this paper with the help of logic variables, with 0 and 1 to quickly distinguish between the presence of crop planting, set A_{ijkl} as a logic variable, if $A_{ijkl} = 0$ then that is not planted in the ith field l crop, if $A_{ijkl} = 1$ then that is planted in the ith field l crop.

$$A_{ijkl} = \begin{cases} 0 \text{ unplanted} \\ 1 \text{ planted} \end{cases} \quad (2)$$

The inability to grow the same crop two years in a row can be expressed mathematically as:

$$A_{ijkl} + A_{i(j+1)kl} \leqslant 1 \quad (3)$$

In addition, the crops planted in 2023 are known, and crops already planted in 2023 cannot be planted in 2024 in the same plots, and the data need to be filtered and processed to make them fit the constraints.

Plant legume crop at least once in three years constraint

The roots of legume crops contain rhizobacteria, they can act as nitrogen fixers. Soil containing rhizobacteria of legume crops is favorable for other crops and all the land in each plot is planted with legume crops at least once in three years from 2023. It is shown that at least once in every three years from 2024 to 2030 all the land in each plot will be planted with legume crop and the mathematical formula is expressed as:

$$\sum_{i=y}^{y+2} \sum_{k=1}^{2} A_{ijkl} \geqslant 1, y = 2024 \sim 2028 \quad (4)$$

Plot size constraints

The size of the plot is limited, with 1,201 acres of open cropland scattered into 34 plots of different sizes, and another 16 ordinary greenhouses and 4 smart greenhouses, each of which has a cropland area of 0.6 acres. The sum of the total area of crops cultivated in the respective plots should not exceed the area of the respective plots, i.e.

$$\sum_{l=1}^{m} a_{ijkl} \leqslant S_i \quad (5)$$

where S is the plot area.

Crops can be planted constraints

After research, the types of crops that can be grown on flat dry land, terraced land, hillside land and watered land are different, and the crops that can be grown in ordinary greenhouses and smart greenhouses are different in different seasons, and some crops have special requirements for the return season. Flat and dry land, terraced land and hillside land are suitable for planting grain crops other than rice in a single season every year. Watered land can grow vegetables in two seasons; in the first season, a variety of vegetables can be grown except cabbage, white radish and carrot, and in the second season, only one of cabbage, white radish and carrot can be grown. Cabbage, white radish and carrot can only be grown in the second season on watered land. Ordinary greenhouses grow two crops per year. In the first season, you can grow a variety of vegetables except cabbage, white radish and carrot, and in the second season, you can only grow edible mushrooms, which can only be grown in ordinary greenhouses in the fall and winter. Smart greenhouses can grow two seasons of vegetables every year, except cabbage, white radish and red radish. Let D_{ikj} be a logistic variable, if $D_{ikl} = 0$, then it means that the lth crop cannot be grown in the i-th plot, if $D_{ikl} = 1$, then it means that the lth crop can be grown in the i-th plot. Process the data to facilitate calculations.

$$D_{ikl}=\begin{cases}0 \text{ unplantable}\\ 1 \text{ can grow}\end{cases} \tag{6}$$

Crop species as the x-axis, planting plots as the y-axis, the establishment of a coordinate system, each crop can be planted plots as shown in Figure 1, in which the intersection of the two axes have a blue labeling represents can be planted.

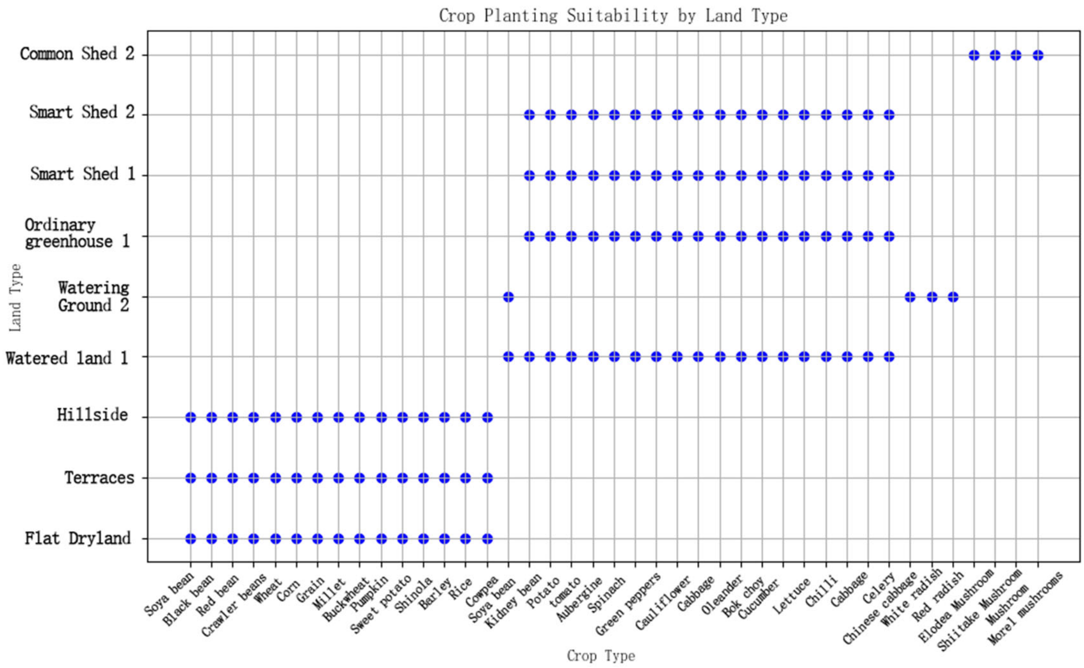

Figure 1. Plots and seasons in which each crop can be grown

We aggregate the constraints and sum to obtain the final optimization model as:

$$\min \frac{\min a_{ijkl}-\max a_{ijkl}}{Z=\sum_{i=1}^{n}\sum_{j=2024}^{2030}\left[\begin{array}{c}\sum_{k=1}^{2}\sum_{l=1}^{m}Y_{jl}a_{ijkl}(N_{jl}-C_{jl})(1-B_{jl})+\\ \sum_{k=1}^{2}\sum_{l=1}^{m}M_{jl}a_{ijkl}(N_{jl}-C_{jl})B_{jl}\end{array}\right]}$$

$$s.t.\begin{cases}A_{ijkl}+A_{i(j+1)kl}\leqslant 1\\ \sum_{j}^{j+2}\sum_{k=1}^{2}A_{ijkl}\geqslant 1, j=2024\sim 2028\\ \sum_{l=1}^{m}a_{ijkl}\leqslant S_i\\ A_{ijkl}\in\{0,1\}\ B_{jl}\in\{0,1\}\\ D_{ikl}=1,\quad D_{il}\in\{0,1\}\end{cases} \tag{7}$$

In the process of solving the model for optimizing the planting scheme for agricultural products, it is affected by the conditions that the same crop cannot be planted in two consecutive years and that there should be beans planted in three years. The crops that can be planted in each year are determined by the crops planted in the previous year, so it can be further simplified in the solution process, making the model solution process more convenient, without affecting the accuracy of the results. In the objective function of the model, this paper considers the dynamic changes of mu yield and unit sales price, and uses decision analysis and stochastic planning to incorporate uncertainty into the optimization model, and finds the solution that performs optimally in the whole by analyzing it under optimistic, neutral, and pessimistic future scenarios. After the above analysis and optimization, this paper uses integer programming to solve the planting scenarios from 2024 to 2030 year by year.

In the process of establishing the objective function, due to the uncertainty of the sales price and, this paper divides the final result into optimistic, neutral and pessimistic three cases, taking the upper quartile for the optimistic case of the sales price, taking the average for the neutral case, and taking the lower quartile for the pessimistic case. The final three results are obtained, and then the profit values of the three scenarios are added together as the expectation of profit under different market conditions, and the planting plan with the largest expectation is the best planting plan. In our calculations, the yields of each crop in the optimistic, pessimistic, and neutral scenarios are different, but basically converge. In the pessimistic scenario, the yields of the various crops are summarized in Figure 2 (some crops are not shown because their yields are too small).

The fact that some of the crops were not planted or were planted in very small quantities means that such crops may not be suitable for local cultivation or the cost of cultivation is too high for large-scale cultivation. Further analysis of the data reveals that some plots are not planted with crops, and after in-depth discussion, this paper concludes that due to the limited expected sales volume of the crops, there may be a situation where no money is made from planting the crops. Therefore, some plots are abandoned. Finally, the total profit of various crops from 2024 to 2030 is optimistic 41577113 yuan, neutral 38288978 yuan, pessimistic 35495018 yuan.

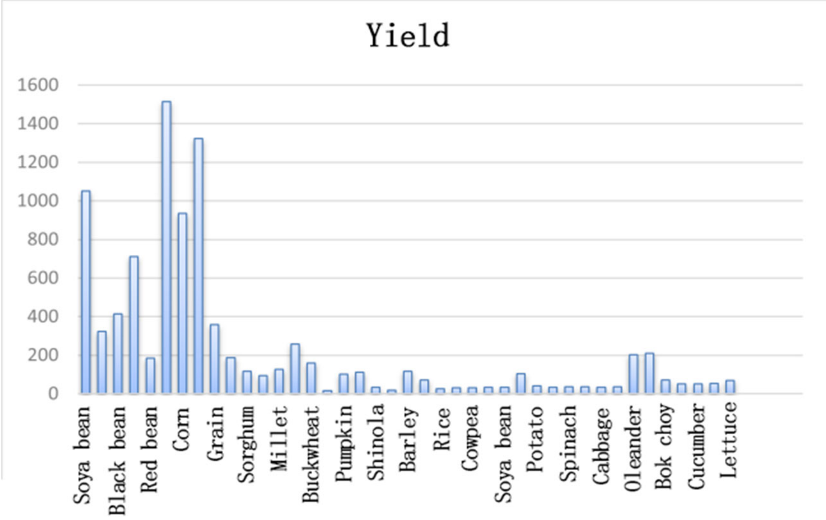

Figure 2. Yield of each crop under the pessimistic scenario

3. Crop planting Strategies That Consider Uncertainty and Planting Risk

In reality, expected crop sales, acreage, growing costs, and selling prices are not likely to remain stable, but rather vary from season to season, year to year, market to market, and other factors. This chapter explores how to adjust the planting strategy of agricultural products to maximize returns in the face of uncertainty and potential planting risks.

Uncertainty in expected crop sales, acreage, growing costs, and selling prices is reflected in the following: expected future sales of wheat and corn are expected to increase, on average, between 5 and 10 percent per year; expected future sales of other crops are expected to change by about ±5 percent per year relative to 2023; acreage is expected to change by ±10 percent per year; and growing costs are expected to increase on average by about 5 percent per year. planting costs will increase by an average of about 5% per year; the sales price of grain crops will be essentially stable; the sales price of vegetable crops will increase by an average of about 5% per year; the sales price of edible mushrooms could decrease by 1% to 5% per year; and the sales price of morel mushrooms could decrease by 5% per year.

Due to the expected sales volume, mu yield, planting costs, sales prices from a stable amount of fluctuations in the range of quantities, the need to describe the variables, analyze its probability distribution, by consulting the relevant information, mu yield, planting costs, sales prices of incremental and decremental to obey the uniform distribution. We transform stochastic planning into deterministic planning by stochastic simulation of random variables to find the expected value of the maximum profit, the expected value of the minimum profit, and the expected value of the moderate profit. In this paper, we use stochastic planning to determine the mu yield, planting cost, and selling price by adding a random variable. Considering the diverse changes in the market environment, the market situation is divided into optimistic, neutral, pessimistic three kinds of optimistic situation to take the upper limit of the range of changes, pessimistic situation to take the lower limit of the range of changes, the neutral situation to take the middle of the range of changes, and now the profit value of the three kinds of situations will be added up as the expectation of profit in different market situations, and the planting program with the largest expected value will be the best planting program.

In addition, more than part of the stagnant direct waste is not in line with real life, and with sustainable development and the concept of green, so this paper uses more than part of the sales price in 2023 by 50% of the sales price price reduction to sell this situation to solve. Random variables will be added to the mu yield, planting costs, sales price changes in the range of fluctuations into a certain amount, can be obtained from the following formula indicates that the 2024 ~ 2030 crops mu yield, planting costs, sales price.

$$
\begin{gathered}
Y_{(j+1)l} = (1+R_Y)Y_{jl} \begin{cases} R_Y \in 5\%\sim 10\% \ l=6,7 \\ R_Y \in -5\%\sim 5\% \end{cases} \\
M_{(j+1)l} = (1+R_M) \ R_M \in -10\%\sim 10\% \\
C_{(j+1)l} = (1+R_C) \ R_C = 5\% \\
N_{(j+1)l} = (1+R_N) \begin{cases} R_N = 0\% \ l \in 1\sim 16 \\ R_N = 0\%\sim 5\% \ l \in 17\sim 37 \\ R_N = -1\%\sim -5\% \ l \in 38\sim 40 \\ R_N = -5\% \ l = 41 \end{cases}
\end{gathered} \tag{8}
$$

where $R_Y R_M R_C R_N$ are the expected sales volume, mu yield, planting costs, sales price of random variables, obeying a uniform distribution. We also use integer programming to solve the model, cf. above. In the solution process, robust optimization can be used to simplify the solution process.

The goal of robust optimization is to find a decision that is as optimal as possible for each possible scenario, i.e., for all uncertainty sets. The performance of this decision in the face of the worst-case scenario is also referred to as "worst-case performance". Under the optimal cropping strategy, Figure 3 shows the yield of each crop under the pessimistic scenario.

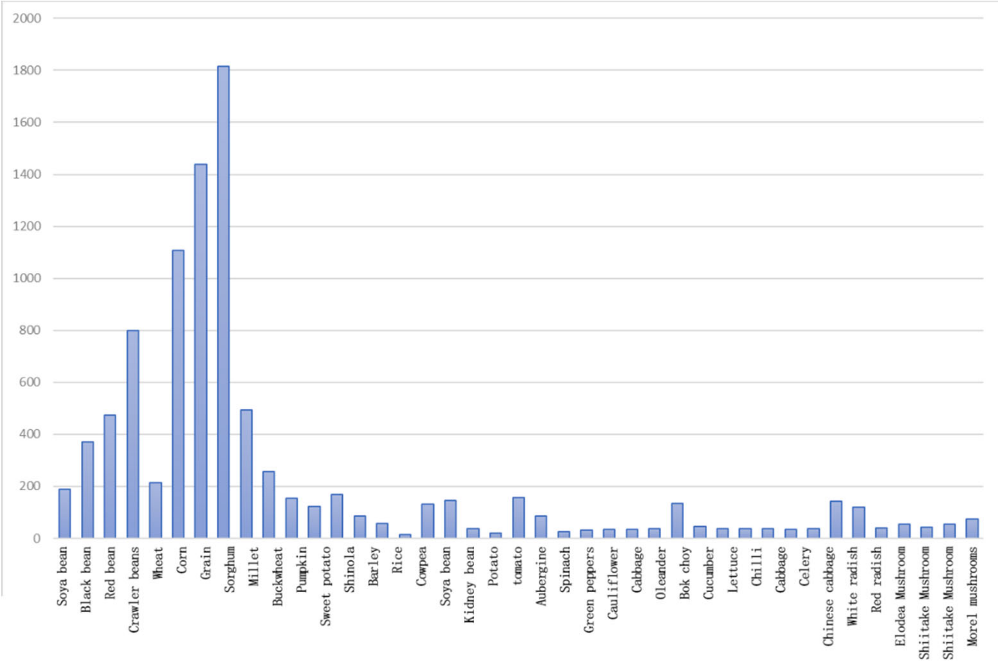

Figure 3. Yield of each crop under uncertainty and planting risk

The crops grown on a large scale under the optimal planting scheme obtained in this chapter remain essentially unchanged, and there is a tendency for the expected future sales of wheat and corn to increase, with a substantial increase in the amount planted. The sales price of vegetable crops increases by about 5% per year on average, and there is a small increase in the amount of planting, but vegetables can be planted on smaller plot sizes, so there is a substantial increase in the relative amount of planting. The sales price of edible mushrooms can decrease by 1% to 5% per year, and the sales price of morel mushrooms decreases by 5% per year. Although the sales price decreases year by year and the cost increases year by year, which makes the available income decrease, the initial selling price of edible mushroom crops is high, and it can only be planted in ordinary greenhouses in the second season, so its planting quantity remains relatively stable.

4. Optimal Planting Scheme Considering Crop Relatedness and Substitutability

In agricultural production, one crop can be grown in place of another, a situation known as crop substitutability. Substitutability of agricultural products is determined by factors such as type of cropland, acreage, cost of cultivation and selling price. We conducted a cluster analysis to get the preliminary substitutability, and there is a negative correlation between the crops that are substitutable for each other in terms of market competition and planting competition. Through the literature search, it was found that there are also some crops planted at the same time or every other season successive planting will make the return reduced, specifically soybean and wheat, sorghum and millet, tomatoes and cucumbers, peppers and eggplants, corn and sorghum, peppers and cauliflower, and so on.

Through a search of the literature, many crops can be planted in a reasonable combination of planting, resulting in a "complementary effect", specifically corn, pumpkin and beans, corn and beans, beans and potatoes, etc. In addition, there are legumes and roots. In addition, the roots of legumes contain rhizobacteria, which fix nitrogen. Soils containing rhizobacteria from legume crops favor other crops

Therefore, we categorize the relationship between crops into strong positive correlation and strong negative correlation. Strong positive correlation means that there is a complementarity between the crops and growing one crop helps to increase the yield of another crop. A strong negative correlation means that the crops are mutually exclusive and growing one crop will result in a decrease in the yield of the other crop. On this basis, this paper introduces the Spearman correlation coefficient, for two crops X and Y, its Spearman correlation coefficient is calculated as:

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{9}$$

where d_i is the difference between the rankings of crop X and crop Y in the i-th sample, and n is the number of samples.

The Spearman correlation coefficient takes the range of [-1,1], which $r_s = 1$ means that the rankings of two crops are completely positively correlated, $r_s = -1$ indicating that the direction of change of Crop X and Crop Y is the same in each sample; it means that the rankings of two crops are completely negatively correlated, $r_s = 0$ indicating that the direction of change of Crop X and Crop Y is opposite in each sample; and it means that there is no correlation of the rankings between two crops.

The Spearman correlation coefficients we obtained for each crop are shown in Figure 4 below

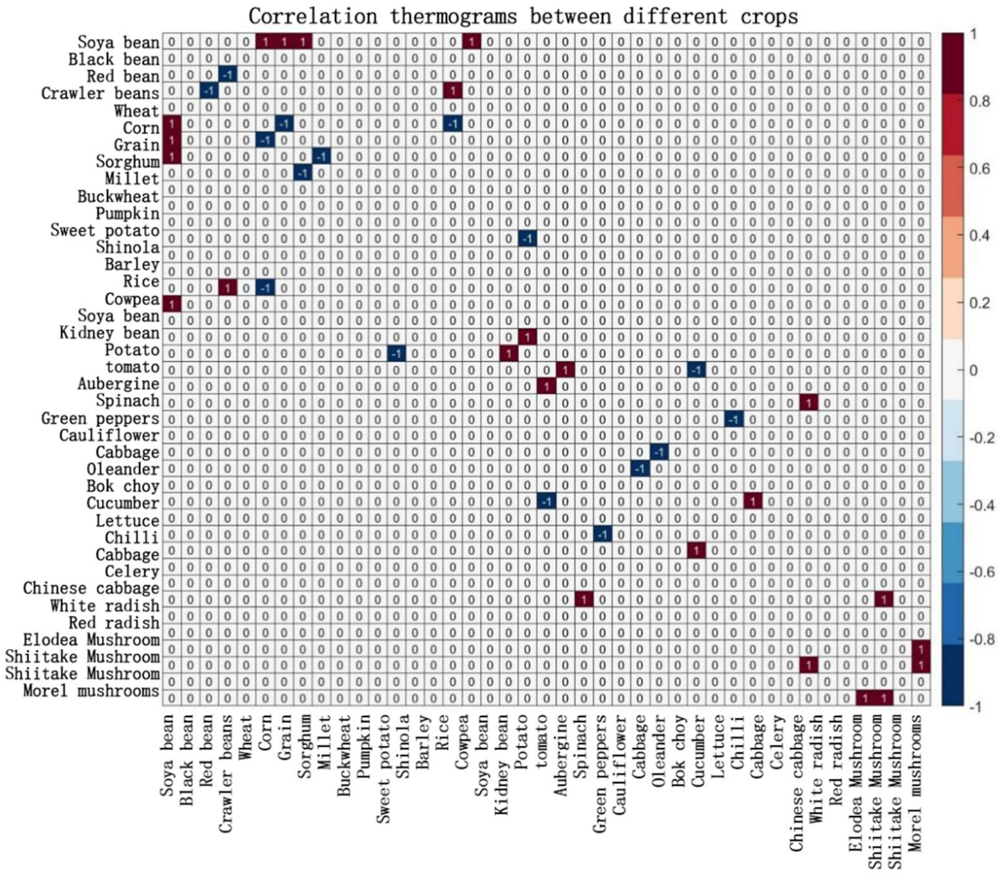

Figure 4. Spearman's correlation coefficient for each crop

The elasticity correlation between the sales volume, the sales price and the planting cost, specifically, the more supply, the market oversupply, the lower the price, the less supply, the lack of commodities in the market, the higher the price, the two are negatively correlated; the expected increase in the sales volume can be reflected to a certain extent in the increase in planting, the increase in planting may lead to a unit of area of the land rent, machinery costs and other fixed costs will decline, but with the expansion of fertilizer, seeds, labor and other variable costs may increase; the planting cost affects the sales price, but the actual sales price depends more on market competition and supply and demand. But with the expansion of planting scale, fertilizer, seeds, labor and other variable costs may increase; planting costs affect the sales price, but the actual sales price depends more on the market competition and supply and demand when the market competition is fierce, despite the increase in planting costs, the sales price may not necessarily be synchronized with the increase, and may even decline, resulting in a narrower profit margin. On the contrary, if there is a shortage of supply, the selling price can rise sharply without depending entirely on changes in costs.

In order to specifically analyze the relationship between expected sales volume and sales price and planting cost, we analyze them using the OLS regression model. OLS regression is a statistical method used to estimate the parameters of a linear regression model. Its goal is to find the line of best fit by minimizing the sum of squared errors, which minimizes the gap between observed and predicted values. Since the expected sales volume is not simply linearly related to sales price and planting cost, but elastically related. So this paper takes the logarithm of the expected sales volume and sales price, planting cost three, so that it is as close as possible to the linear model. The formula is as follows:

$$Y = a\ln C + b\ln N + c \tag{10}$$

For linear relationships of class $y = \beta_0 + \beta_1 x + \epsilon$, OLS minimizes the gap between predicted and actual values for all data points by finding a set of parameters β_0 and β_1. The formula is expressed as:

$$RSS = \sum_{i=1}^{n} (y_i - y_i^*)^2 = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2 \tag{11}$$

The final solution is $a = 1.0177$ $b = 1.0775$ $c = -0.601$. Solving this chapter yields a total profit of ￥25,733,174.65 from 2024 to 2030, which is a significant decrease in total profit compared to crop correlation and substitutability, based on the basis of whether or not to consider crop correlation and substitutability, and comparing the yields of various crops from 2024 to 2030, as shown in Figure 5. It is found that the yield of most of the crops has decreased, and the decrease of vegetables is especially obvious, which may be due to the inclusion of strong positive correlation and strong negative correlation between crops, the negative correlation of various crops is too strong resulting in a decrease in returns.

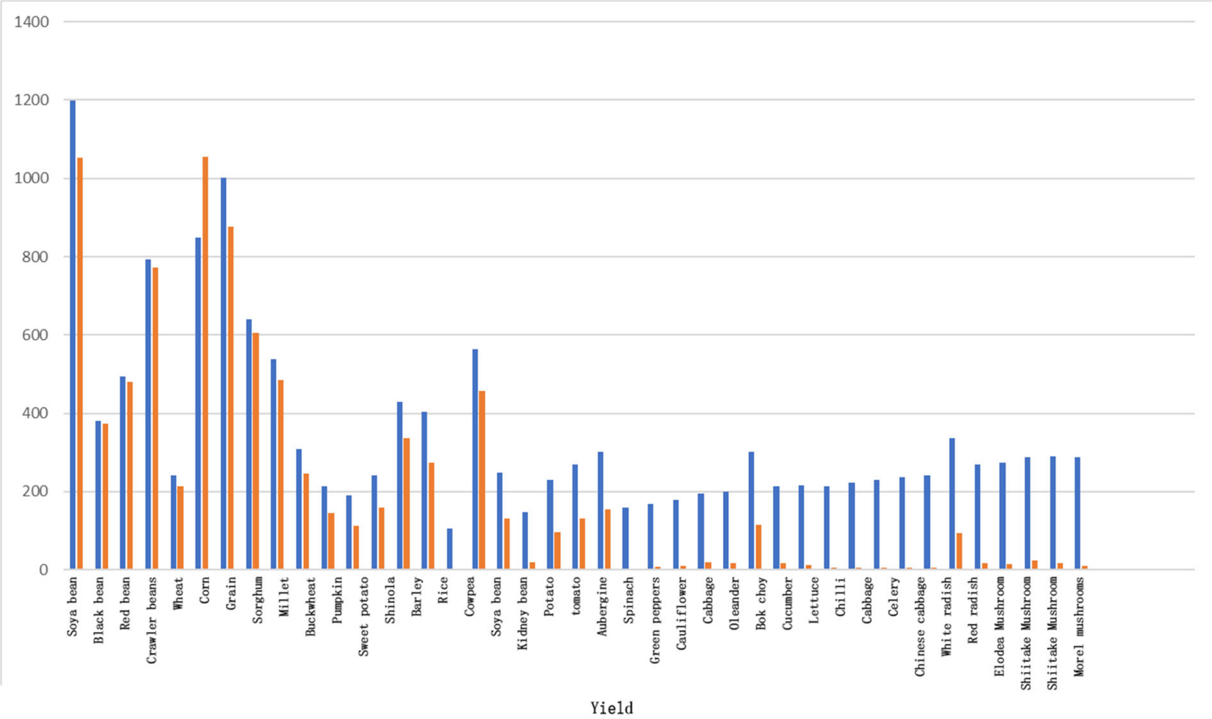

Figure 5. Whether or not to consider crop correlation versus substitutable crop yields

5. Conclusions

In this paper, the planting optimization of crops in rural areas is thoroughly studied through integer planning and stochastic planning models. Based on the analysis of the current agricultural background and modeling theory, a planting optimization model with comprehensive consideration of market uncertainty is constructed, and the optimal planting plan is proposed. The results show that the substitutability and complementarity between crops have an important impact on the overall planting efficiency, and reasonable planting planning can effectively improve the economic returns.

References

[1] Cao Jiangqiu, Scientific Understanding of Strong Agriculture from Multi-Dimensional Perspective (Economic Daily News, Aug. 24, p. 10).

[2] Yao Cunliu. Research on multi-stage portfolio decision-making based on stochastic planning and decision rule [D]. Lanzhou University of Technology, 2021.

[3] Zhou Xue, Zhao Yongman,Wang Xiangyang. Research on the equipping of agricultural machinery of 105th regiment based on mathematical planning method[J]. Xinjiang Agricultural Mechanization,2018,(01):36-39.

[4] Li M M. Research on branch bounding algorithm for three types of integer programming problems[D]. Northern Nationalities University,2023.

[5] Shoukui Si, Xijing Sun, Mathematical Modeling Algorithms and Applications 3rd Edition, Beijing: National Defense Industry Press, 410-411, 2021:18-19.

[6] Wang Wei. Research on fractal characteristics of several agricultural futures prices in China [D]. Sichuan:Southwest University of Finance and Economics,2020:11-12.

[7] WANG Xiaoming, ZHANG Zhang. Research on ecological optimization of crop cultivation [J]. Agricultural Science, 2019(5): 45-50.

[8] Xu Weichao. A review of correlation coefficient research [J]. Journal of Guangdong University of Technology,2012,(3): 12-17.