Research on Optimization of Crop Planting Strategies and Multi-Factor Analysis Based on Edge Computing and Cloud Computing

Qi Liu *

School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China, 100091

* Corresponding Author Email: 22723011@bjtu.edu.cn

Abstract. This study aims to address the challenges in agricultural development by combining communication technologies with multi-factor analysis to optimize crop planting strategies and improve agricultural production efficiency. Leveraging edge computing and cloud computing technologies, this research achieves real-time data collection, transmission, and intelligent processing of agricultural data, constructing a comprehensive optimization model. By collecting historical sales data of 41 crops and calculating the Pearson correlation coefficient, the study identifies a strong complementary relationship between soybeans and black beans, as well as substitution relationships between wheat and corn, and between Chinese cabbage and bok choy. Furthermore, the Manhattan distance is introduced to quantify the yield differences between crops, and complementary benefit and substitution loss modules are incorporated into the objective function, effectively coordinating the allocation of planting areas for different crops. Through an analysis of the price elasticity of demand for grain, vegetable, and edible fungus crops, it is found that grain crops exhibit the weakest price sensitivity, while edible fungi show the strongest sensitivity. The most significant changes in crop yields occur within the price fluctuation range of -10% to 10%. This study provides a scientific basis for optimizing agricultural planting strategies, promotes the development of smart agriculture, and offers strong support for the efficient allocation of agricultural resources and the maximization of economic benefits.

Keywords: Edge computing, cloud computing, crop planting strategies, complementary-substitution relationships, Pearson correlation coefficient.

1. Introduction

Agriculture serves as the cornerstone of human society, directly impacting national food security and economic development [1]. Improving agricultural production efficiency and optimizing crop planting strategies have become critical issues in the agricultural sector. The rapid development of communication technologies has brought new opportunities to agriculture. In crop planting, edge computing and cloud computing technologies play a vital role by enabling in-depth analysis and mining of large-scale data, providing robust support for agricultural decision-making.

In recent years, numerous scholars have conducted extensive research on agricultural planting optimization. For example, Zhang Qiaoying found that although the agricultural industry in Yuzhong County, Gansu Province, has achieved certain results, there are still issues such as the need to improve the quality of germplasm resources and the monotony of rural industrial structures. She proposed optimization measures such as enhancing modern seed industry quality and promoting the efficiency of specialty industries [2]. Zhao Yang and colleagues identified problems in the development of the planting industry, including significant disparities in investment, imbalanced planting ratios, and lagging resource recycling. They suggested scientifically planning planting development paths, formulating adjustment strategies based on regional resources, cultivating new types of planting management entities, and improving industrial chains [3].

However, existing research has notable limitations. On the one hand, most studies have constraints in data collection and the application of communication technologies, focusing primarily on data transmission while lacking in-depth research on data processing and analysis. On the other hand, many studies overlook the complementary-substitution relationships among crops and fail to consider the impact of crop prices on sales.

To address these issues, this study aims to resolve the shortcomings of existing research by combining communication technologies with multi-factor analysis to optimize crop planting strategies. The specific goals are to leverage edge computing and cloud computing technologies to achieve real-time data collection, transmission, and intelligent processing of agricultural data. By collecting historical sales data, planting areas, and yields of 41 crops in 2023, the study uses the Pearson correlation coefficient to precisely analyze the complementary-substitution relationships among crops. It introduces complementary benefit and substitution loss modules to explore their impact on optimal planting strategies. Additionally, based on price elasticity theory, crops are categorized into grain, vegetable, and edible fungus classes, and the price sensitivity of each category is analyzed in detail. This comprehensive approach examines the impact of market price fluctuations on the sales of different crops.

2. Application of Communication Technologies in Agricultural Planting

In smart agriculture, wireless sensor networks (WSN) are responsible for collecting environmental data from various agricultural fields, such as temperature, humidity, light intensity, and soil moisture. After preliminary processing, these data are transmitted to edge computing nodes and cloud computing platforms via 5G communication technology for further analysis and decision-making support. Edge computing and cloud computing play crucial roles in this process, ensuring efficient data processing and real-time responsiveness.

2.1. Edge Computing Platform

The edge computing platform, as a computing node close to the data source, is responsible for real-time data processing, local decision support, and collaboration with the cloud computing platform. With its low-latency and high-responsiveness characteristics, the edge computing platform ensures real-time and efficient agricultural production [4]. Deployed at the edges of farmlands or inside greenhouses, the platform processes environmental data collected by sensors in real time. Sensor networks transmit data to edge computing nodes using low-power wireless communication protocols such as ZigBee and LoRa. The edge computing nodes first perform preliminary data processing, such as noise reduction and filtering, to ensure data accuracy and reliability.

2.1.1 Kalman Filter Algorithm

Edge computing nodes use the Kalman filter algorithm to fuse and denoise data from multiple - source sensors [5]. The Kalman filter fuses the input and output observation data of the system through prediction and update steps to make an optimal estimate of the system state. Its core formulas are as follows:

Prediction step:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \tag{1}$$

$$P_k^- = AP_{k-1}A^T + Q \tag{2}$$

where $\hat{x}_k^-$ is the predicted state value, A is the state- transition matrix, B is the control- input matrix, u_{k-1} is the control input, P_k^- is the predicted error covariance matrix, and Q is the process noise covariance matrix

Update step:

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{3}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \tag{4}$$

$$P_k = (I - K_k H)P_k^- \tag{5}$$

where K_k is the Kalman gain, H is the observation matrix, z_k is the observed value, R is the observation noise covariance matrix, $\hat{x}_k$ is the updated state estimate, and P_k is the updated error covariance matrix

The Kalman filter algorithm can effectively fuse data from multiple- source sensors such as temperature, humidity, illumination, and carbon dioxide concentration, accurately estimate the environmental state, and assist in planting decisions.

2.1.2 Wavelet Transform Algorithm

The edge computing nodes also use the wavelet transform

algorithm to denoise and extract features from the illumination data [6]. The wavelet transform analyzes the signal at different scales to separate the noise from the real- signal and retains the information about illumination changes.Its core formula is:

$$W(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt \tag{6}$$

where $W(a,b)$ is the wavelet coefficient, $f(t)$ is the original signal, $\psi_{a,b}(t)$ is the wavelet basis function, and a and b are the scale and translation parameters respectively.

2.1.3 Local Decision Support and Environmental Control

The edge computing platform is not only responsible for data processing but also capable of making local decisions based on real-time data, ensuring rapid response and efficiency in agricultural production. The edge computing platform monitors environmental parameters in real-time, automatically adjusting the environmental conditions of greenhouses and fields, optimizing irrigation and fertilization control, and ensuring the stability of the crop growth environment.

In smart greenhouses, edge computing nodes can automatically adjust the environmental parameters within the greenhouse based on data such as temperature, humidity, and light intensity. For instance, when the temperature exceeds the set threshold, the edge computing node can automatically activate ventilation equipment or adjust shading curtains to maintain the stability of the crop growth environment. Similarly, when light intensity is insufficient, the edge computing node can automatically turn on supplementary lighting equipment to ensure that crops receive adequate light.

In arid fields, edge computing nodes can automatically control the operation of irrigation systems based on data from soil moisture sensors. For example, when soil moisture falls below the set threshold, the edge computing node can activate irrigation equipment to ensure that crops receive sufficient moisture. Likewise, edge computing nodes can also automatically control the operation of fertilization equipment based on soil nutrient data, optimizing fertilizer usage. By monitoring soil moisture and nutrient conditions in real-time, the edge computing platform can achieve precise irrigation and fertilization, reduce resource waste, and improve crop yield and quality.

2.2. Application of Cloud Computing Platforms in Smart Agriculture

The cloud computing platform serves as the core for data processing and decision support, undertaking tasks such as large-scale data storage, advanced analysis, and global decision-making. By integrating data transmitted from edge computing nodes, the cloud computing platform utilizes various models and algorithms for in-depth analysis, providing scientific decision-making support for agricultural production [7].

2.2.1 Data Storage and Management

The cloud computing platform first stores sensor data from edge computing nodes in a cloud-based database. This data includes various environmental parameters such as temperature, humidity, light intensity, soil moisture, and carbon dioxide concentration. Through the data management platform, users can access historical data at any time for trend analysis and forecasting.

2.2.2 Data Storage Architecture

The cloud computing platform adopts a distributed storage architecture to ensure high availability and scalability of data. The data storage architecture typically includes the following layers:

Data Collection Layer: Responsible for receiving data from edge computing nodes.

(1) Data Storage Layer: Stores data in a NoSQL database.

(2) Data Management Layer: Provides functions such as data querying, backup, and recovery to ensure data security and traceability.

2.2.3 Advanced Analysis and Decision Support

The cloud computing platform employs various advanced algorithms and models to conduct in-depth analysis of the data, offering scientific decision-making support for agricultural production. By leveraging communication technologies, the platform collects rich environmental data. Based on this data, the platform uses the DSSAT model to predict crop yield per unit area [8][9], the linear regression model to evaluate planting costs [10], and the ARIMA time series model to forecast market prices, generating key parameters required for optimizing planting strategies [11].

(1) Application of the DSSAT Model

The DSSAT model is used to predict crop yield per unit area. The cloud computing platform dynamically adjusts model parameters by integrating historical data and real-time environmental data, providing accurate yield predictions.

Data Input: The platform collects real-time environmental data (e.g., temperature, humidity, light intensity, soil moisture) from edge computing nodes and combines it with historical data for model input.

Model Calculation: The platform uses the DSSAT model to predict yields, generating crop growth curves and yield distribution maps.

Decision Support: Based on the prediction results, the platform provides farmers with planting recommendations, such as adjusting irrigation and fertilization levels, to optimize crop yields.

(2) Application of the Multiple Linear Regression Model

The multiple linear regression model is used to assess planting costs. The cloud computing platform analyzes historical planting data to build a cost prediction model, helping farmers optimize resource allocation.

Data Input: The platform collects data on planting area, labor costs, fertilizer inputs, and machinery usage from edge computing nodes.

Model Calculation: The platform uses the multiple linear regression model to predict costs, generating cost distribution maps and optimization suggestions.

Decision Support: Based on the prediction results, the platform provides farmers with cost control recommendations, such as adjusting labor allocation and optimizing fertilizer usage, to reduce planting costs.

(3) Application of the ARIMA Time Series Model

The ARIMA time series model is used to predict market prices. The cloud computing platform analyzes historical market price data to build a price prediction model, helping farmers develop sales strategies.

Data Input: The platform collects historical price data from market sources and combines it with real-time market information for model input.

Model Calculation: The platform uses the ARIMA time series model to predict prices, generating price trend charts and volatility analysis.

Decision Support: Based on the prediction results, the platform provides farmers with sales recommendations, such as adjusting sales timing and selecting sales channels, to maximize profits.

3. Analysis of Complementary-Substitution Relationships Considering Potential Gains and Losses

The complementary relationship between crops is observed when consumers tend to purchase one crop and simultaneously choose to purchase another, as the two complement each other. The substitution relationship, on the other hand, is observed when consumers typically select one crop over the other, and the two are mutually exclusive.

3.1. Model Setup Considering Complementary-Substitution Relationships

3.1.1 Model Construction

Since complementary or substitution relationships exist among certain crops, this study calculates the correlations between 41 different crops based on their historical sales data and market conditions using the Pearson correlation coefficient (The historical sales data for the 41 crops primarily comes from the China Agricultural Yearbook published by the National Bureau of Statistics and publicly available market data from the Ministry of Agriculture and Rural Affairs). The Pearson correlation coefficient is a statistical measure used to quantify the degree of linear correlation between two variables [12]. As shown in Figure 1, the red color tones represent complementary relationships, while the blue color tones represent substitute relationships. The results show that there is a strong complementary relationship between soybeans and black beans, and a strong substitution relationship between wheat and corn, as well as between napa cabbage and bok choy.

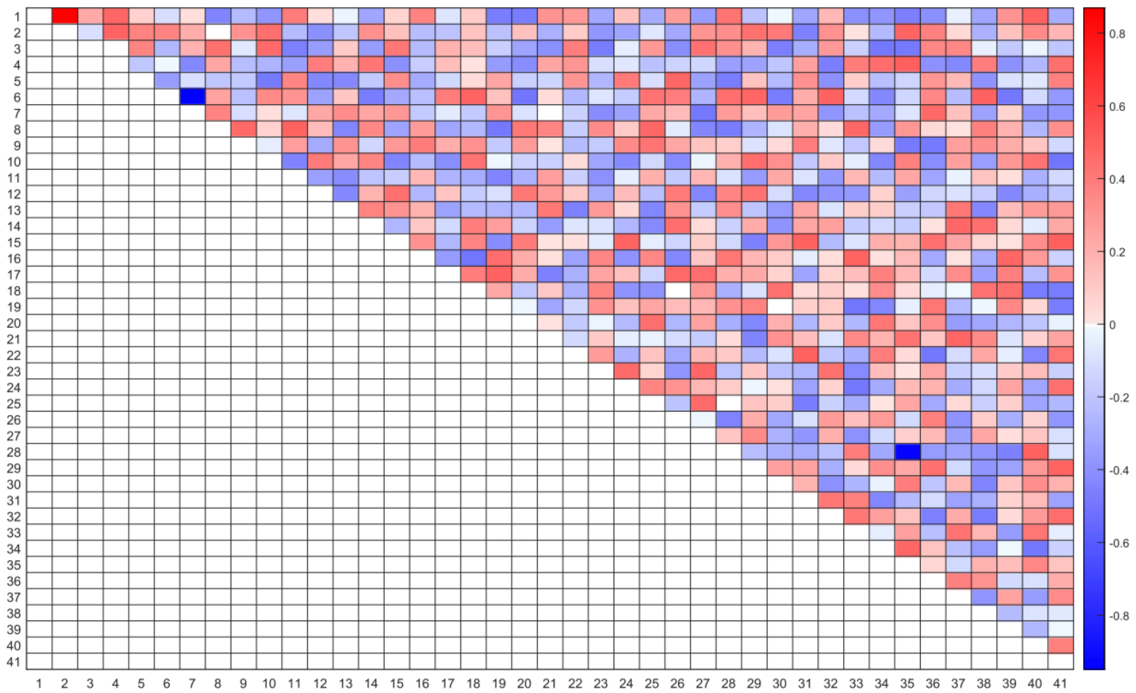

Figure 1. Analysis of crop complementary-substitution relationship

Notes: The sequence from 1 to 41 in the figure represents the following crops in order: Soybeans, Black beans, Red beans, Green beans, Lima beans, Wheat, Corn, Millet, Sorghum, Foxtail millet, Buckwheat, Pumpkin, Sweet potatoes, Oats, Barley, Rice, Cowpeas, Sword beans, Kidney beans, Potatoes, Tomatoes, Eggplants, Spinach, Green peppers, Cauliflower, Cabbage, Oil lettuce, Bok choy, Cucumbers, Lettuce, Chili peppers, Water spinach, Yellow heart cabbage, Celery, Napa cabbage, White radish, Red radish, Yellow mushrooms, Shiitake mushrooms, White mushrooms, and Morel mushrooms.

3.1.2 Model Mechanism

To characterize the potential benefits or losses brought by the complementary-substitute relationships between crops, for example, when there is a large difference in the yields of crops with complementary relationships, it reduces consumer purchasing desire, thus leading to a decrease in revenue; or when crops with substitute relationships have similar yields, consumers may prefer one, resulting in the surplus of the other. The Manhattan distance, also known as the taxi distance or city block distance, is a metric used to measure the distance between two points in geometric space [13]. This paper introduces the Manhattan distance to characterize the yield difference between two crops and incorporates a new module in the objective function based on their complementary-substitute relationship coefficients as follows:

$$\beta\left\|x_{i,j_1}^{k}-x_{i,j_2}^{k}\right\|\gamma_c \tag{7}$$

In the above-mentioned formula, $x_{i,j}^{k}$ represents the area of the j-th crop planted on the i-th plot in the k-th year. β represents the complementary - substitute coefficient of crops. A positive coefficient indicates a complementary relationship between the two, and the smaller the yield distance, the more favorable it is for the objective; a negative coefficient indicates a substitute relationship between the two, and the larger the yield distance, the more favorable it is for the objective. γ_c is the unit loss or gain, and this parameter reflects the impact of the complementary - substitute relationship on the model results.

3.1.3 Result Analysis Considering Complementary-Substitution Relationships

This study is based on a comprehensive development scenario, where complementary benefits and substitution losses modules are introduced to construct a new model. Other parameters are consistent with the comprehensive development scenario (the comprehensive development scenario refers to taking the average change of various parameters to reflect future balanced development. Specifically, in this scenario, the expected annual growth rate of wheat and corn sales is 7.5%, and the expected sales of other crops will change by approximately ±5% relative to 2023; the expected average annual price increase for vegetables is 5%; the expected annual price decrease for edible fungi is about 3%; planting costs increase by an average of 5% annually; the per-acre yield changes by ±10% each year; and the price of morel mushrooms decreases by 5% annually). The three pairs of crops with the strongest complementary-substitution relationships were selected, and the planting yield differences for these three pairs of crops were compared under the two models. The results are shown in Figure 2.

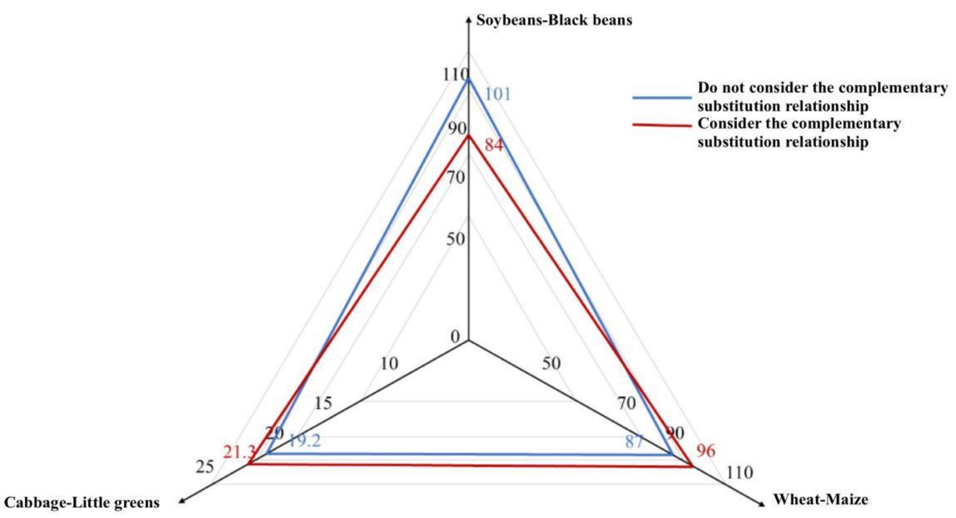

Figure 2. Planting Area Distance for Three Sets of Crops Under the Influence of Complementary-Substitution Relationships

The results in Figure 2 show that after considering the complementary-substitution relationships, the planting area difference between soybeans and black beans, which have a complementary relationship, decreased from 101 to 84. Meanwhile, the planting area differences for napa cabbage-bok choy and wheat-corn, which have substitution relationships, increased by 2.1 and 9, respectively. This indicates that by introducing the complementary benefits and substitution loss modules in the model, the distribution of planting areas for different crops was effectively coordinated.

4. Analysis of the Sales Price-Expected Sales Volume Relationship Considering Crop Price Elasticity

4.1. Price Elasticity Analysis for Different Types of Crops

The 41 crops are divided into three categories: cereals, vegetables, and edible fungi. The relationship between the sales price and expected sales volume of different types of crops, i.e., price elasticity, is analyzed.

The demand price elasticity of cereals, vegetables, and edible fungi refers to the sensitivity of the market demand for these agricultural products to price changes, i.e., the ratio of the percentage change in demand to the percentage change in price. Different agricultural products have different demand elasticities, which are generally represented by the demand elasticity coefficient. The following is a specific analysis of the demand price elasticity for these three types of agricultural products:

(1) Cereals: As a staple food, the demand price elasticity of cereals is generally small or inelastic. Cereals are the foundation of daily diets, and their basic demand remains relatively stable regardless of price changes. Specifically, when the price of cereals rises, although consumers may reduce some non-essential cereal consumption, the overall demand will not decrease significantly. Conversely, when the price of cereals falls, consumers will not significantly increase their purchase of cereals.

(2) Vegetables: The demand price elasticity of vegetables is relatively large. As an important part of the daily diet, the demand for vegetables is influenced by factors such as consumer dietary habits and health awareness. When vegetable prices rise, consumers may look for substitutes or reduce their consumption of vegetables. On the other hand, when vegetable prices fall, consumers may increase their vegetable purchases.

(3) Edible Fungi: Edible fungi, as a special agricultural product, are difficult to fully replace in many cases due to their unique taste, nutritional value, and specific roles in certain dishes. When the price of edible fungi rises, consumers have limited substitute options, which increases their sensitivity to price changes, i.e., the demand increases elasticity.

In summary, cereals, as a staple food, have a small demand price elasticity; vegetables have a relatively large demand price elasticity; and edible fungi have the largest demand price elasticity. The demand price elasticity for these three types of crops is represented in Figure 3. In the figure, the vertical axis represents price, and the horizontal axis represents expected sales volume.

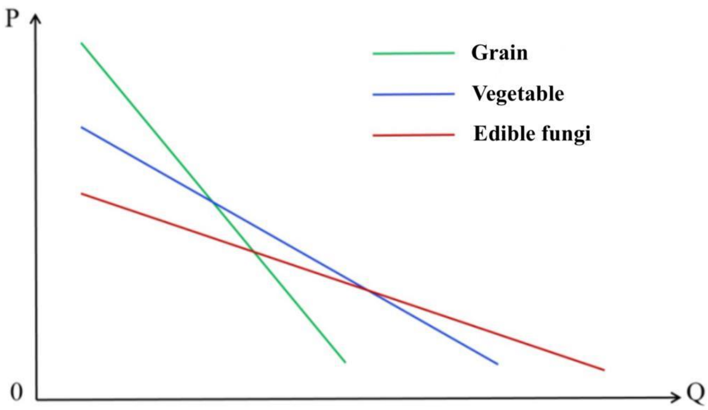

Figure 3. Schematic diagram of the price elasticity of the three types of crops

4.2. Scenario Design and Result Analysis Considering Price Elasticity

Based on the analysis of price elasticity for the three types of crops, this study designed seven parameters as shown in Table 1. Starting from the comprehensive development scenario (where price changes are 0), the prices of crops varied from -30% to 30%. The changes in sales volume for different types of crops were set according to their different price elasticities.

Table.1. Parameter Design Considering Price Elasticity

Price changes	Sales volume change of grain	Sales volume change of vegetable	Sales volume change of edible fungi category
-30%	5%	15%	20%
-20%	3%	8%	15%
-10%	2%	5%	8%
0	0	0	0
10%	-2%	-5%	-8%
20%	-3%	-8%	-15%
30%	-5%	-15%	-20%

The analysis was conducted by selecting the yield change ranges for three types of crops under different price change amplitudes as indicators, and the results are shown in Figure 4. From the perspective of different crop types, cereals exhibit the weakest sensitivity (-2.6% to 2.2%), while edible fungi show the strongest sensitivity (-8.1% to 9.3%). From the price change interval perspective, when the price change is in the range of -10% to 10% (red area), the yield change for all types of crops is the largest. The ±10%-20% range (yellow area) comes second, and the ±20%-30% range (green area) shows the weakest yield change. Contrary to intuitive perception, when the price of agricultural products changes within a small range, the yield is more sensitive. This may be because agricultural products are necessities, and when the price increases significantly, the minimum yield must be reached. However, when the price increase is small, it will not exceed the maximum yield.

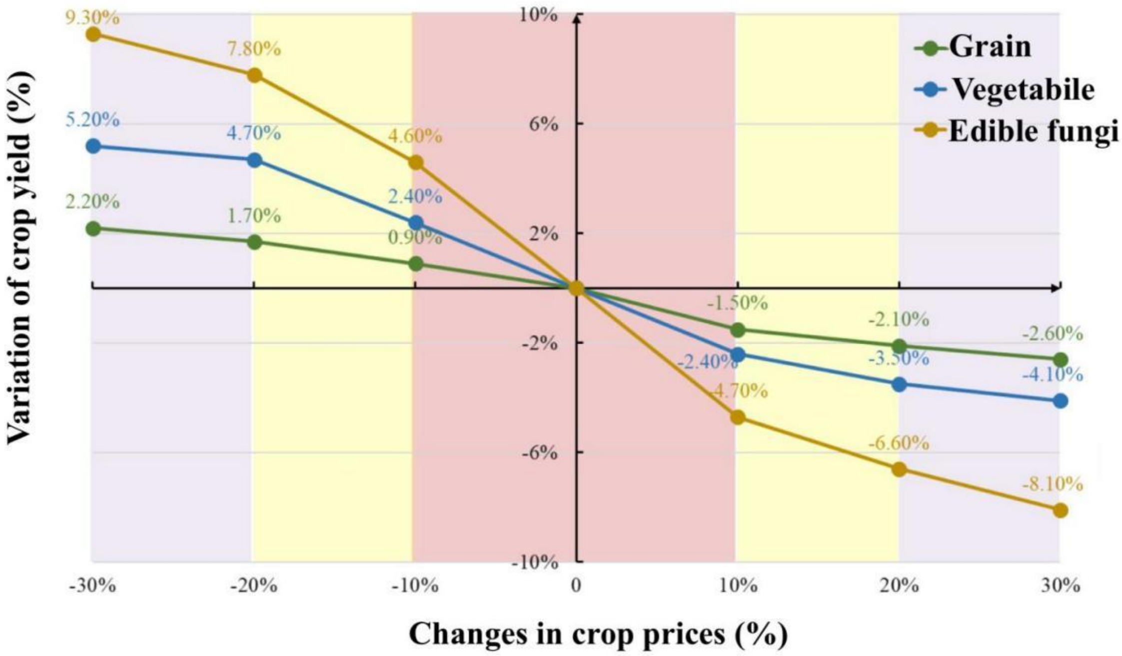

Figure 4. The range of variation in yield of the three class crops affected by price

5. Conclusions

This study successfully constructs a comprehensive optimization model by deeply integrating edge computing and cloud computing technologies, combined with multi-factor analysis, to analyze the complementary-substitution relationships among crops and the impact of price elasticity on planting strategies. A scientific agricultural planting optimization scheme is proposed. The research results indicate a strong complementary relationship between soybeans and black beans, while substitution relationships exist between wheat and corn, as well as between Chinese cabbage and bok choy. By introducing the Manhattan distance and complementary benefit-substitution loss modules, the allocation of planting areas for different crops is effectively coordinated, further enhancing the efficiency of agricultural resource allocation. Additionally, through the analysis of price elasticity for grain, vegetable, and edible fungus crops, it is found that grain crops exhibit the weakest price sensitivity, while edible fungi show the strongest sensitivity. Moreover, within the price range of -10% to 10%, the yield changes of various crops are most significant.

The innovation of this study lies in the close integration of communication technology with agricultural planting strategy optimization. By fully leveraging edge computing and cloud computing technologies, real-time data collection, transmission, and intelligent processing of agricultural data are achieved, significantly improving the accuracy and practicality of the model. At the same time, this study is the first to organically combine the complementary-substitution relationships among crops with market price fluctuations, comprehensively considering crop planting benefits, market demand changes, and price volatility. This provides agricultural practitioners with highly accurate and flexible decision-making support.

Future research could further expand the breadth and depth of data sources, extensively covering crop data from more regions and longer time spans. By integrating advanced artificial intelligence algorithms and big data technologies, in-depth mining and analysis of massive agricultural data can be conducted to continuously improve the planting strategy optimization model. This will help better address the emerging trends and challenges in the agricultural industry, steadily advance the modernization of agriculture, and provide strong support for the efficient allocation of agricultural resources and the maximization of economic benefits.

References

[1] Wang, X. In-depth study and implementation of General Secretary Xi Jinping's important discourses and instructions on "Three Rural Issues" and solidly moving towards the goal of building an agriculturally strong district [N]. Xinjiang Daily (Chinese), 2025-02-08(003).

[2] Zhang, Q. The current situation, problems, and optimization measures of agricultural industry development in Yuzhong County, Gansu Province, under the background of rural revitalization [J]. Agricultural Engineering Technology, 2024, 44(19): 15-16.

[3] Zhao, Y., & Zhao, X. Research on agricultural industry structure optimization under the rural revitalization strategy [J]. Southern Agriculture, 2021, 15(29): 38-39.

[4] Pu Jin'an, Bai Huafei, Sun Shifeng, et al. Design and Implementation of a Greenhouse Monitoring System Based on Edge Computing and Agricultural IoT [J]. Southern Agriculture, 2024, 18(17): 245-251.

[5] Zhang Fanyu, Yang Yunyi, Kang Jun, et al. Edge Computing Architecture Design for Intelligent Waste Disposal Systems [J/OL]. Journal of Tianjin University of Technology, 1-8 [2025-02-23].

[6] Zhao Heng, Miao Kun, Chen Lei, et al. Research on Power IoT Networking Solutions Based on Edge Computing [J]. Microcontrollers & Embedded Systems, 2021, 21(10): 7-11.

[7] Chang Jun. Design and Implementation of a Smart Agricultural Sensor Network on a Cloud Computing Platform [J]. Wireless Internet Technology, 2017, (15): 31-32.

[8] Wu Yuanyuting. Evaluation of the Effects of Nitrogen Application Rate and Film - mulching Cultivation on Wheat Yield in the Loess Plateau Using the DSSAT Model [D]. Northwest A&F University, 2024.

[9] Zhang Puhan. Estimation of Lanzhou Lily Yield Based on the DSSAT Model [D]. Northwest Normal University, 2024.

[10] Suo Panpan. Analysis of the Impact of Temperature on Wheat Yield Based on the Functional Regression Model [D]. Henan University, 2024.

[11] Li Moying. Time - series Analysis Model Based on Multi - source Data [J]. Information Technology, 2025, (01): 112 - 118 + 125.

[12] Cheng Juanjuan. An Empirical Study on the Relationship Between Scientific Research and Teaching in Universities—Based on Pearson Correlation Coefficient Analysis [J]. Chinese University Science and Technology, 2022, (10): 46-52.

[13] Bu Yijun, Xu Jian. Solving Problems Related to Manhattan Distance from a Dynamic Perspective [J]. Mathematics Teaching, 2022, (02): 28-31.