

DATA FOUNDATION AND MODELING FRAMEWORK ANALYSIS FOR OPTIMIZING FARMLAND RESOURCE ALLOCATION AND CROP PLANTING STRUCTURE IN MOUNTAINOUS RURAL AREAS OF NORTH CHINA

WENJIN QIN*

Sydney Smart Technology College, Northeastern University, Qinhuangdao, 066004, China

*Corresponding author's email: jerryqsh2003@163.com

Abstract

This study develops crop-planting plans for fragmented farmland in mountainous rural areas of North China. The model covers grain crops, vegetables, legumes, and edible fungi, and accounts for differences in land conditions. It specifies planting area, expected and surplus yield, crop-plot suitability, planting decisions, and market fluctuation parameters for subsequent linear and intelligent optimization. The dataset covers 1,201 mu of open-field farmland and 20 mu of greenhouse land. Plots include flat dryland, slopes, terraces, conventional greenhouses, and smart greenhouses. Price ranges were represented by their mean values, and baseline sales volumes were reconstructed from the historical census year. Flat dryland and terraces carried most field-crop production, while greenhouses supported vegetables and edible fungi. Under stable conditions, the optimized profit was about 3.0 million yuan when surplus was discarded and about 15.0 million yuan when it was sold at 50% of the original price. Under market uncertainty, particle swarm optimization reached about 18.0 million yuan in the discounted-surplus case. The regression model produced an R^2 of 0.835 and an adjusted R^2 of 0.816. The resulting scenario-based planting plans account for land use, rotation, legume requirements, profit, and market uncertainty.

Key words: Crop Planting Optimization; Farmland Resource Allocation; Multi-objective Linear Programming; Genetic Algorithm; Uncertainty Modeling

Introduction

Mountainous parts of North China have fragmented plots, varied land types, and limited irrigation. These conditions affect which grain crops, vegetables, legumes, and edible fungi can be grown on each plot (Mardani Najafabadi *et al.*, 2019). Over a seven-year planning cycle, planting decisions must also meet rotation requirements and respond to changes in costs and demand (Dury *et al.*, 2012). Farmers often base decisions on experience and past planting patterns, which makes it difficult to consider these constraints together. Earlier studies have often emphasized one-season returns or single-crop patterns. They give less attention to multi-year rotation, crop-land suitability, greenhouse intensity, and limits on planting concentration. This study uses data from 50 plots to define planting areas, plot constraints, yield, sales scenarios, rotation rules, and economic returns in one optimization model.

The primary objective of this study is to develop an integrated decision-making framework for multi-year farmland resource allocation and crop-planting structure optimization in fragmented and heterogeneous mountainous farmland in North China. The study does not aim to propose a fundamentally new genetic algorithm or particle swarm optimization algorithm; rather, these optimization techniques are employed as solution approaches within the proposed agricultural planning framework.

The framework integrates crop-land suitability, multi-year crop rotation, legume requirements, planting-concentration limits, surplus-production treatment, and

market uncertainty into a unified optimization process. This integrated formulation is intended to reflect the practical decision environment of mountainous agricultural systems, where heterogeneous plot conditions, fragmented land resources, agronomic rotation requirements, and uncertain market conditions must be considered simultaneously.

The contribution of the study therefore lies in the integration and application of these interacting constraints and uncertainty factors within a multi-year crop-planning framework and in evaluating the resulting planting strategies under stable and uncertain market scenarios.

Previous research on agricultural crop planning has used several methodological approaches. Mathematical programming and linear programming have been widely applied to crop-area allocation and agricultural resource management because they provide transparent formulations for land, production, and resource constraints (Zeng *et al.*, 2010; Filippi *et al.*, 2017; Jiang, 2025). Multi-stage and multi-objective formulations have further extended these models to represent temporal decisions and competing economic or resource objectives (Groot *et al.*, 2012; Galán-Martin *et al.*, 2015). However, the effectiveness of such approaches depends strongly on how land heterogeneity, nonlinear relationships, and uncertainty are represented in the formulation.

Evolutionary and intelligent optimization methods have been increasingly used for crop-planning problems with large or complex search spaces. Genetic and hybrid evolutionary algorithms can provide flexible search mechanisms for complex agricultural planning problems (Sarker & Ray, 2009;

Li *et al.*, 2025), while NSGA-II has been applied to multi-objective crop allocation and spatial optimization (Krityakierne *et al.*, 2025; Wang *et al.*, 2025). Particle swarm optimization has also been used for crop-planting strategy optimization (Benini *et al.*, 2025; Zuo *et al.*, 2025). These methods provide useful computational search capabilities, but the optimization algorithm itself does not replace the need to formulate crop suitability, rotation, land-use, and operational constraints appropriately for the agricultural context.

Another important research direction concerns uncertainty in agricultural planning. Crop yield, sales volume, production cost, and market price can vary over time, making a single deterministic planting plan potentially sensitive to changes in economic and production conditions. Existing studies have addressed uncertainty through alternative optimization formulations, simulation, and scenario-based analysis (Xu *et al.*, 2025). Crop rotation and diversification have also been incorporated into agricultural planning to reflect agronomic, and soil-management requirements (Felsman *et al.*, 2025; Kümmerer *et al.*, 2025; Shilpa *et al.*, 2025; De Barros *et al.*, 2026). These strands of research demonstrate the importance of considering both economic optimization and agronomic constraints, but they do not necessarily address all of these factors simultaneously within a fragmented mountainous farmland setting.

Against this background, the research gap addressed in the present study is not the absence of individual optimization techniques, but the need for an integrated planning formulation that combines heterogeneous plot suitability, multi-year crop rotation, legume requirements, planting-concentration limits, alternative surplus-production treatments, and market uncertainty. The present study therefore builds on established mathematical and intelligent optimization methods and applies them within a unified framework tailored to fragmented mountainous farmland in North China.

Linear programming, intelligent optimization, orthogonal uncertainty tests, and regression analysis are used to generate planting plans for heterogeneous plots (Groot *et al.*, 2012; Li *et al.*, 2025). The model includes crop rotation, legume requirements, planting-concentration limits, and economic relationships among crops. Section 2 describes the data and methods. Section 3 presents the results under different scenarios. Section 4 discusses their agronomic implications and limitations, and Section 5 concludes the paper.

The main contributions of this study are threefold. First, it formulates a multi-year crop-planning framework that simultaneously considers heterogeneous plot conditions, crop-land suitability, crop rotation, legume requirements, and limits on planting concentration. Second, it incorporates alternative surplus-production treatments and market uncertainty into the planning process, allowing the robustness of planting decisions to be examined under different economic conditions. Third, it combines mathematical programming with intelligent optimization and regression-based analysis to generate and evaluate feasible planting strategies for the specific context of fragmented mountainous farmland. These contributions are intended to provide a practical planning framework for agricultural resource allocation rather than to claim a new optimization algorithm.

To further clarify the positioning of the present study relative to established optimization approaches, representative

methods used in crop-planning and agricultural resource-allocation studies are compared in Table 1. The comparison focuses on the methodological characteristics and application scope of these approaches rather than direct numerical performance, because different studies generally use different datasets, constraints, planning horizons, and objective formulations. This comparison helps clarify that the contribution of the present study lies primarily in integrating land heterogeneity, crop-land suitability, multi-year rotation, planting-concentration constraints, surplus treatment, and market uncertainty within a unified framework for fragmented mountainous farmland.

Beyond demonstrating the feasibility of the optimization framework, the present study examines the effects of alternative surplus-production treatments and market uncertainty on multi-year crop-allocation decisions. The results show that the treatment of surplus production can substantially change the resulting economic objective, while uncertainty in sales volume, yield, cost, and price affects the robustness and composition of the resulting planting plans. The study therefore uses the optimization framework not only to generate feasible planting allocations, but also to examine how interacting agronomic, land-use, and market constraints shape planning outcomes in fragmented mountainous farmland. These findings are interpreted within the specified study context and are not presented as universally generalizable relationships.

To complement the methodological comparison in Table 1, quantitative optimization results are also reported for the solution approaches investigated in the present study. Under the stable-market condition, the improved genetic algorithm converged to approximately 3.0 million yuan when surplus production was discarded and approximately 15.0 million yuan when surplus production was sold at 50% of the original price. Under the market-uncertainty setting, the PSO-based optimization reached approximately 18.0 million yuan in the discounted-surplus case. The corresponding convergence behaviour is reported in the Results section. These values provide quantitative evidence of the optimization outcomes obtained under the specified model formulation and scenarios. They are not interpreted as direct numerical benchmarks against published state-of-the-art studies, because the existing studies use different datasets, planning horizons, constraints, uncertainty assumptions, and objective formulations that do not permit a like-for-like comparison.

This comparison indicates that the present study's contribution lies primarily in integrating and applying these interacting planning factors within a mountainous, fragmented agricultural setting, rather than in proposing a fundamentally new optimization algorithm.

The present study does not claim that the individual modeling components are novel or that the proposed framework is the first of its kind. Instead, its contribution is defined by the integrated application of established modeling components to the specific conditions of fragmented and heterogeneous mountainous farmland. This distinction is important because crop suitability, rotation, optimization algorithms, and uncertainty modeling have each been investigated in previous studies, whereas their combined use in the planning context examined here provides the specific application contribution of this study.

Table 1. Methodological comparison of representative crop-planning optimization approaches and the present framework.

Approach	Typical application	Key strength	Limitation relative to present framework
Linear programming (Zeng <i>et al.</i> , 2010; Filippi <i>et al.</i> , 2017; Jiang, 2025)	Crop-area/resource allocation	Transparent optimization of constrained allocation	Limited treatment of complex nonlinear/uncertain search conditions
SA-GA / evolutionary methods (Sarker & Ray, 2009; Li <i>et al.</i> , 2025)	Complex crop-planning problems	Flexible search in large solution spaces	Computational solution method rather than agricultural framework itself
NSGA-II (Krityakime <i>et al.</i> , 2025; Wang <i>et al.</i> , 2025)	Multi-objective crop allocation	Handles competing objectives	Requires appropriate formulation and parameterization
PSO (Zuo <i>et al.</i> , 2025)	Crop-planting optimization	Efficient population-based search	Does not itself represent land-specific agronomic constraints
Present framework	Fragmented mountainous farmland	Integrates land suitability, rotation, legumes, concentration, surplus and market uncertainty	Application-specific; direct numerical benchmarking with published studies is limited by differences in datasets and problem formulations, while broader benchmark testing remains future work

Theoretical framework and analytical methods

Data preprocessing and model assumptions: The agricultural data were organized before model construction. The survey recorded plot area, plot type, and crop suitability (Galán-Martín *et al.*, 2015; Saracian *et al.*, 2025). The dataset includes 1,201 mu of open-field farmland and 20 mu of greenhouse land. Plots were classified as flat dryland, terraces, irrigated fields, slopes, conventional greenhouses, or smart greenhouses (Borthakur *et al.*, 2025). These classifications were used as constraints in the suitability assessment and planting-area allocation.

For production and economic variables, the midpoint of each price range was used to limit the influence of extreme values. Baseline sales volumes were reconstructed from the historical census year. The final dataset records crop type, plot type, planting area, yield per mu, cost, selling price, and expected sales volume.

The practical planning problem considered in this study is characterized by substantial heterogeneity in both land resources and production conditions. The study area contains multiple plot types, including flat dryland, terraces, irrigated fields, slopes, conventional greenhouses, and smart greenhouses, with different suitability for crop production. The fragmented nature of these plots means that crop allocation cannot be determined solely from expected economic return. Instead, planting decisions must simultaneously account for plot-specific suitability, available area, crop rotation, legume requirements, operational minimum areas, and limits on the spatial concentration of individual crops. The inclusion of these interacting constraints distinguishes the application framework from a simple crop-area optimization problem and provides the practical context for evaluating the proposed multi-year planning approach.

The model assumes stable agro climatic conditions, manageable planting operations, and predictable changes in demand. Extreme weather, biological threats such as pathogen outbreaks (Clement *et al.*, 2025), and unusual market disruptions are outside the scope of the model (Park *et al.*, 2025; Sakpal, 2026).

Linear programming model for crop planting optimization

Model establishment and objective function: The continuous variable $X_{i,j,t}$ denotes the cultivation area of the i -th crop in the j -th plot during the t -th year. Correspondingly, $E_{i,j,t}$ is defined as the amount (in jin, and

1 jin is 500 gram) of the i -th crop that goes beyond the anticipated sales volume produced from the j -th plot in the t -th year. $Z_{i,j,t}$ is a binary indicator that indicates whether crop i is cultivated in plot j during year t (1 for planting, 0 for not planting).

The objective function when the excess is unsold and wasted is shown in Formula (1) (Jiang, 2025):

$$\max \sum_{t=1}^7 \sum_{j=1}^{82} \sum_{i=1}^{41} [(X_{i,j,t} \cdot Y_{i,t} - \text{Excess}_{i,j,t}) \cdot P_{i,t} - X_{i,j,t} \cdot C_{i,t}] \quad (1)$$

The key variables in the model are: $P_{i,t}$, reflecting the market sales price for the i -th crop during year t ; $Y_{i,t}$, denoting the per-unit-area yield of the crop i in year t ; and $C_{i,t}$, representing the corresponding crop cost of the crop i in year t .

Meanwhile, when the surplus volume is disposed of at half its baseline price, the objective function is shown in Formula (2) (Filippi *et al.*, 2017; Jiang, 2025; Yu *et al.*, 2025):

$$\max \sum_{t=1}^7 \sum_{j=1}^{82} \sum_{i=1}^{41} [(X_{i,j,t} \cdot Y_{i,t} - \text{Excess}_{i,j,t}) \cdot P_{i,t} + \text{Excess}_{i,j,t} \cdot P_{i,t} \cdot 0.5 - X_{i,j,t} \cdot C_{i,t}] \quad (2)$$

Model constraints: Non-negativity of Excess: Since $E_{i,j,t}$ denotes the surplus over the expected sales, it must be non-negative.

$$\text{Excess}_{i,j,t} \geq 0 \quad \forall i, j, t \quad (3)$$

Upper Boundary for Excess: The amount that exceeds sales is calculated by subtracting the anticipated sales volume from the actual yield. This only applies if the actual production is greater than the forecasted demand. Herein, $S_{i,t}$ represents the anticipated sales volume.

$$\text{Excess}_{i,j,t} \geq X_{i,j,t} \cdot Y_{i,t} - E_{i,t} \quad \forall i, j, t \quad (4)$$

Minimum Plot Area for Crops: To ensure machinery accessibility, the planting area allocated to each chosen crop within a given plot is bounded by a minimum operational scale of 5 mu.

$$X_{i,j,t} \geq 5 \cdot Z_{i,j,t} \quad \forall i, j, t \quad (5)$$

Minimum Land Utilization Constraint: The area used for planting on a plot should be no less than 20% of the overall plot size. The overall size of the plot i is denoted by A_j .

$$X_{i,j,t} \geq 0.2 \cdot A_j \cdot Z_{i,j,t} \quad \forall i, j, t \quad (6)$$

Maximum Land Allocation Constraint: The total area of crops planted on each plot must not surpass the total plot area. A_j is defined as the overall size of the plot j .

$$\sum_{i=1}^{41} X_{i,j,t} \leq A_j \quad \forall j, t \quad (7)$$

Link between $X_{i,j,t}$ and $Z_{i,j,t}$: The binary variable $Z_{i,j,t}$ determines whether a plot is used for the i -th crop during year t . If $Z_{i,j,t} = 0$, then the corresponding planting area $X_{i,j,t}$ must be zero. If $Z_{i,j,t} = 1$, then $X_{i,j,t}$ may have a positive value, limited to the maximum possible area for the plot, denoted as M .

$$X_{i,j,t} \leq M \cdot Z_{i,j,t} \quad \forall i, j, t \quad (8)$$

$$Z_{i,j,t} \in \{0,1\} \quad \forall i, j, t \quad (9)$$

No Continuous Cropping: A plot land cannot be planted with the same crop for two consecutive years.

$$Z_{i,j,t} + Z_{i,j,t+1} \leq 1 \quad \forall i, j, t \in \{2024, \dots, 2029\} \quad (10)$$

Rotation Requirement for Leguminous Crops: Each plot must include at least one leguminous crop every three years in order to maintain soil fertility and avoid over-reliance on a single crop type (Felsman *et al.*, 2025; Kümmerer *et al.*, 2025; Shilpa *et al.*, 2025).

$$\sum_{k=0}^2 \sum_{i \in \text{Legumes}} X_{i,j,t+k} \geq 0.5 \quad \forall j, t \in \{2024, 2027, 2030\} \quad (11)$$

Planting Concentration Limit: Aiming to mitigate excessive spatial fragmentation, each crop category is restricted to a maximum of 10 plots per year to concentrate labor and logistical resources.

$$\sum_{j=1}^{82} Z_{i,j,t} \leq 10 \quad \forall i, t \quad (12)$$

Solution framework for stable and uncertain crop-planning scenarios: The formulated crop-planning problem was solved under stable and uncertain market conditions using established intelligent optimization techniques. The improved genetic algorithm was employed for the stable-market scenarios, while particle swarm optimization was used for the uncertain multi-objective scenarios. These algorithms serve as computational solution methods for the integrated crop-planning framework and are not presented as independent algorithmic innovations. Their purpose is to efficiently search for feasible planting configurations subject to the land-use, rotation, suitability, concentration, and market-related constraints defined in the model. The improved genetic algorithm was employed for the stable-market scenarios (Sarker & Ray, 2009; Krityakierne *et al.*, 2025), while particle swarm optimization was used for the uncertain multi-objective scenarios (Benini *et al.*, 2025; Zuo *et al.*, 2025).

Model solution based on improved genetic algorithm

(1) Multi-Matrix encoding and operators: The decision variable contains several linked dimensions. A multi-matrix encoding was used to represent it by separating the search space into smaller components. This makes the genetic search easier to manage while retaining the relationships among crop, plot, and year. $X_{i,j,t}$

Crossover can produce candidates that violate one or more constraints, so the algorithm repairs infeasible candidates before evaluation. Mutation uses random point changes: a year and plot are selected, and values are modified within the plot-area limit. Values above 0.8 are reset to zero. The repair step then restores compliance with all constraints (Li *et al.*, 2025).

(2) Repair strategy for infeasible solutions: In this study, several constraints are applied to correct infeasible solutions:

Constraint (1) Cultivation Restrictions for Specific Land Types: Rice must be planted only once per year on irrigated fields, flat dryland, terraced fields, or sloping fields.

Constraint (2) Compliance with Plot-Specific Crop Restrictions: If a plot is assigned an invalid crop type, such an allocation is removed.

Constraint (3) No Continuous Cropping: The subsequent cultivation instance is omitted if an identical crop is detected on a single plot across back-to-back annual cycles.

Constraint (4) Cultivation Limit by Plot: Any crop that is planted in more than 10 plots during the same year is adjusted to stay within this limit.

Constraint (5) Leguminous Crop Requirement: If a plot has not included a leguminous crop for three consecutive years, an area is allocated for such a crop in the third year.

Constraint (6) Minimum Area for Each Crop: If a specific crop occupies less than 20% of the plot area, it is increased to 30% of the plot's size, while also enacting a rule that restricts each plot from having more than 5 different crops in the same year.

Constraint (7) Maximum Plot Usage Limit: The total cultivation area for a plot cannot exceed the plot's available area.

The repair procedure applies the seven constraints in sequence. Second-season crops that violate the land-type rule are set to zero. Invalid crop-plot assignments and consecutive plantings are removed. Allocations exceeding the plot-count limit are reduced. If legumes are absent for three successive years, an area is assigned to a legume in the third year. Small crop areas are raised to 30% of the plot area, subject to a maximum of five crops per plot and year. If total planting area exceeds plot area, the excess is reduced proportionally.

(3) Model parameter setting: Table 2 outlines the specific parameter configurations established for the developed genetic algorithm.

Table 2. Genetic algorithm parameter setting table.

Parameter	Value
Number of iterations	800
Crossover rate	0.7
Mutation rate	0.3
Population size	30

Orthogonal uncertainty testing and multi-objective linear programming model

(1) Uncertainty test method: Agricultural parameters vary by crop. Sales volumes were allowed to change by about $\pm 5\%$ each year; wheat and corn were assigned annual growth of 5-10%. Yield per mu was allowed to vary by $\pm 10\%$, and planting costs were increased by about 5% each year. Grain prices were held relatively stable, vegetable prices were allowed to rise by about 5% per year, and edible-fungi prices were allowed to decline by 1-5% per year. Because the survey provided intervals rather than exact values, integer approximations were used in the optimization. Table 2 lists the parameter ranges.

Planting decisions are affected by price, sales, yield, and cost uncertainty. An orthogonal experimental design was used to examine the combined effects of the uncertain agricultural and market parameters while avoiding evaluation of every possible combination of parameter levels (Xu *et al.*, 2025). Seven uncertainty factors were considered: annual growth rate of wheat and corn sales volume, sales-volume fluctuation for other crops, per-mu yield variation, annual growth rate of planting cost, annual growth rate of vegetable price, annual reduction rate of edible-fungi price excluding morel, and annual reduction rate of morel price. Their allowable levels are specified in Table 3.

The 32 experimental groups in Table 4 represent 32 selected combinations of these uncertainty factors. Each row of Table 4 defines one uncertainty scenario by assigning a specific level to each of the seven factors. These values are then introduced into the corresponding annual sales-volume, yield, planting-cost, and selling-price calculations according to the uncertainty relationships specified in the model. The resulting parameter set defines the economic and production conditions under which the planting decision is evaluated.

For each experimental group, the multi-objective crop-planting model evaluates the corresponding planting allocation and objective values under the specified uncertainty conditions. The resulting performance measures across the orthogonal scenarios are then used in the PSO-based optimization, where mean profit and minimum profit are considered in evaluating candidate planting plans. Thus, the 32 groups function as representative uncertainty scenarios for evaluating the robustness of planting decisions rather than as 32 independent datasets or independent optimization algorithms.

The use of 32 representative combinations substantially reduces the number of uncertainty combinations that would otherwise be required if all possible parameter-level combinations were evaluated. Table 4 therefore provides the explicit parameter settings used for each experimental group, while Table 3 defines the allowable fluctuation levels from which these settings are specified. The design contains 32 runs and is shown in Table 4.

(2) Multi-objective linear programming model: The multi-objective model maximizes revenue and minimizes cultivation cost while allowing planting area, sales volume, selling price, and cultivation cost to vary across market scenarios.

In multi-objective linear programming model, the maximization of the objective function is derived from the subtraction of the cultivation cost from total sales receipts of the cultivated different, expressed as Formulas (13) and (14):

$$z_1 = \max \sum_{t=1}^T \sum_{i=1}^N (S_{i,t} \cdot P_{i,t} - C_{i,t} \cdot A_{i,t}) \quad (13)$$

$$z_1 = \max \sum_{t=1}^T \sum_{i=1}^N (S_{i,t} \cdot P_{i,t} - C_{i,t} \cdot A_{i,t}) \quad (13)$$

where z_1 and z_2 specify the profit across all orthogonal scenarios, T defines the specific annual cycle, N denotes the set of crop varieties, $S_{i,t}$ represents the market sales volume for crop i in the t -th year (with fluctuation range), $P_{i,t}$ is the corresponding selling price, $C_{i,t}$ represents the corresponding per-unit cultivation expense, and $A_{i,t}$ implies the allocated planting area of the i -th crop during the t -th year.

Planting area, sales volume, selling price, and crop cost vary with market conditions. Formulas (15)-(17) specify how these uncertainties enter the model.

$$S_{i,t} = SS_{2023}(i) \cdot (1 + Suc) \quad (15)$$

$$P_{i,t} = SP_{2023}(i) \cdot (1 + Puc) \quad (16)$$

$$C_{i,t} = SC_{2023}(i) \cdot (1 + Cuc) \quad (17)$$

where $S_{2023}(i)$ represents the sales volume of the i -th crop in 2023, $SP_{2023}(i)$ represents its corresponding, and $SC_{2023}(i)$ represents the associated crop costs. Suc , Puc and Cuc are used to describe the market fluctuation of sales volume, selling price and crop costs in the t -th year respectively.

Particle Swarm Optimization (PSO) algorithm for uncertain scenarios

(1) Model principle: For uncertain multi-objective crop planning, this study uses particle swarm optimization (PSO) (Benini *et al.*, 2025; Zuo *et al.*, 2025). Each candidate planting plan is represented as a particle in a continuous multidimensional search space. Its velocity is updated by Formula (18).

$$v_{id}(t+1) = \omega \cdot v_{id}(t) + a c_1 \cdot r_{n1} \cdot (p_{ib} - x_{id}(t)) + a c_2 \cdot r_{n2} \cdot (p_{gb} - x_{id}(t)) \quad (18)$$

The inertia weight controls velocity. The acceleration coefficients and random terms determine the relative influence of a particle's own best position and the best position found by the swarm. $\omega a c_1 a c_2 r_{n1} r_{n2} p_{ib} p_{gb}$
Position update formula of PSO algorithm particles:

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (19)$$

The PSO fitness function uses mean profit and minimum profit across the orthogonal experiments. Particle positions are updated from these objectives, and the swarm moves toward the best solution found during the search.

Table 3. Uncertainty fluctuation of crops.

Item	Fluctuation range
Sales volume fluctuation (excluding wheat and corn)	-5%, -4%, -3%, -2%, -1%, 0%, 1%, 2%, 3%, 4%, 5%
Annual growth rate of wheat and corn sales volume	5%, 6%, 7%, 8%, 9%, 10%
Per-mu yield	-10%, -9%, -8%, -7%, -6%, -5%, -4%, -3%, -2%, -1%, 0%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%
Annual growth rate of planting cost	5%
Annual growth rate of grain crop price	0%
Annual growth rate of vegetable price	5%
Annual reduction rate of edible fungi price (excluding morel)	1%, 2%, 3%, 4%, 5%
Annual reduction rate of morel price	5%

Table 4. Orthogonal experimental scheme.

Test number	Annual growth rate of wheat and corn sales volume	Sales volume fluctuation (excluding wheat and corn)	Per-mu yield	Annual growth rate of planting cost	Annual growth rate of vegetable price	Annual reduction rate of edible fungi price (excluding morel)	Annual reduction rate of morel price
1.	5%	-5%	-10%	5%	5%	1%	5%
2.	6%	-4%	-9%	5%	5%	2%	5%
3.	7%	-3%	-8%	5%	5%	3%	5%
4.	8%	-2%	-7%	5%	5%	4%	5%
5.	9%	-1%	-6%	5%	5%	5%	5%
6.	10%	0%	-5%	5%	5%	1%	5%
7.	5%	1%	-4%	5%	5%	2%	5%
8.	6%	2%	-3%	5%	5%	3%	5%
9.	7%	3%	-2%	5%	5%	4%	5%
10.	8%	4%	-1%	5%	5%	5%	5%
11.	9%	5%	0%	5%	5%	1%	5%
12.	10%	-5%	1%	5%	5%	2%	5%
13.	5%	-4%	2%	5%	5%	3%	5%
14.	6%	-3%	3%	5%	5%	4%	5%
15.	7%	-2%	4%	5%	5%	5%	5%
16.	8%	-1%	5%	5%	5%	1%	5%
17.	8%	-4%	-7%	5%	5%	1%	5%
18.	9%	-3%	-6%	5%	5%	1%	5%
19.	10%	-2%	-5%	5%	5%	3%	5%
20.	5%	-1%	-4%	5%	5%	4%	5%
21.	6%	0%	-3%	5%	5%	5%	5%
22.	7%	1%	-2%	5%	5%	1%	5%
23.	8%	2%	-1%	5%	5%	2%	5%
24.	9%	3%	0%	5%	5%	3%	5%
25.	10%	4%	1%	5%	5%	4%	5%
26.	5%	5%	2%	5%	5%	5%	5%
27.	6%	-5%	3%	5%	5%	1%	5%
28.	7%	-4%	4%	5%	5%	2%	5%
29.	9%	-3%	5%	5%	5%	3%	5%
30.	10%	-2%	-7%	5%	5%	4%	5%
31.	8%	-4%	-6%	5%	5%	5%	5%
32.	8%	-3%	-5%	5%	5%	1%	5%

Correlation-based regression modeling and evaluation method: This section examines relationships among crop sales volume, selling price, cost, and crop substitutability. Pearson correlations quantify linear associations. Sales volume is the dependent variable, and total production cost, selling price, cultivated area, crop cost, yield per mu, and unit sales price are the independent variables. Formula (20) gives the regression equation. $YX_1X_2X_3X_4X_5X_6$

For the regression analysis, 58 complete observations were extracted from the original agricultural dataset, with observed sales volume used as the dependent variable and the corresponding economic and production variables used as explanatory variables.

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6 + \varepsilon \quad (20)$$

where β_0 is the constant intercept, while the coefficients $\beta_1, \beta_2, \dots, \beta_6$ indicate the specific contribution of each

independent variable to the dependent variable, and ε represents the random error term.

The regression model allows expected sales volume to vary with selling price, crop cost, yield per mu, unit sales price, and planting area. This provides an additional input for the uncertainty model and allows planting plans with and without these economic relationships to be compared.

The regression model was evaluated with MSE, RMSE, MAE, MAPE, SMAPE, SSE, and R2. R2 was used to summarize the proportion of variation in sales volume explained by the model. These measures were used to quantitatively characterize the in-sample regression fit and to provide supplementary information for interpreting the economic relationships represented in the dataset and their use in the subsequent uncertainty analysis.

Results and Analysis

Stable market optimization results: The stable-market analysis uses two surplus treatments. In Scenario 1, output above expected sales is discarded. In Scenario 2, surplus output is sold at 50% of the original price. The improved genetic algorithm was run for both cases.

Figure 1 gives the Scenario 1 convergence curve. Profit rose quickly at the beginning of the run and changed little after about 750 iterations. The final objective value was approximately 3,000,000.

Figure 2 gives the Scenario 2 convergence curve, where surplus output is sold at a 50% discount. The curve became

nearly flat after about 580 iterations, with a final value of about 15,000,000. This run stabilized earlier than Scenario 1.

The convergence curves indicate stabilization of the objective values under the tested genetic-algorithm settings. However, convergence to a stable objective value alone does not constitute a mathematical proof that the obtained solution is the global optimum. The reported solutions are therefore interpreted as converged feasible planning solutions under the specified algorithm parameters and model constraints, rather than as mathematically proven global optima.

Selling surplus at a discount produced a higher objective value than discarding it. The handling of surplus therefore changes both the planting plan and its economic result.

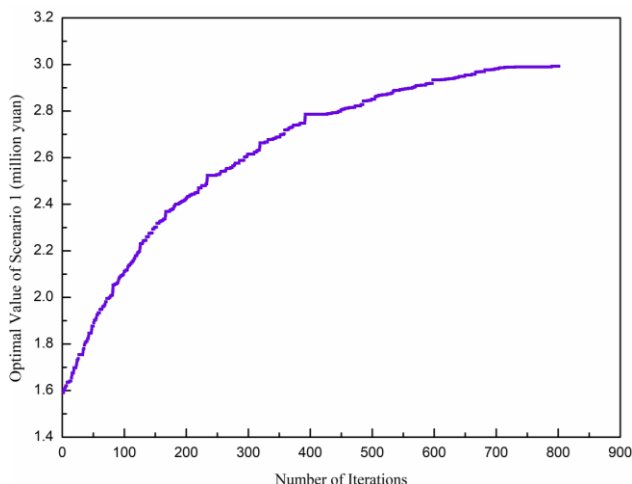


Fig. 1. Convergence curve of scenario 1.

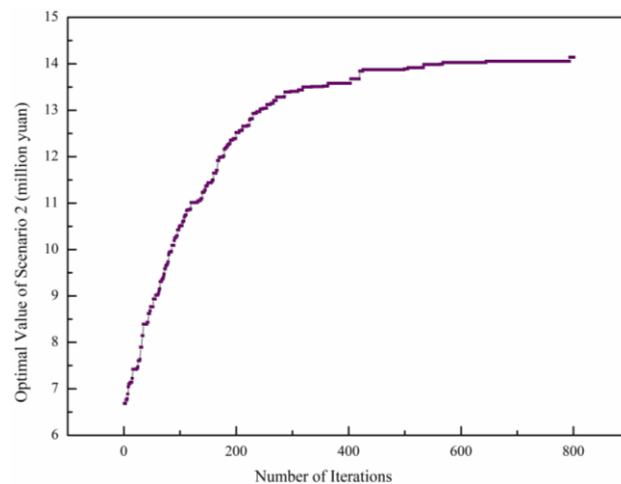


Fig. 2. Convergence curve of scenario 2.

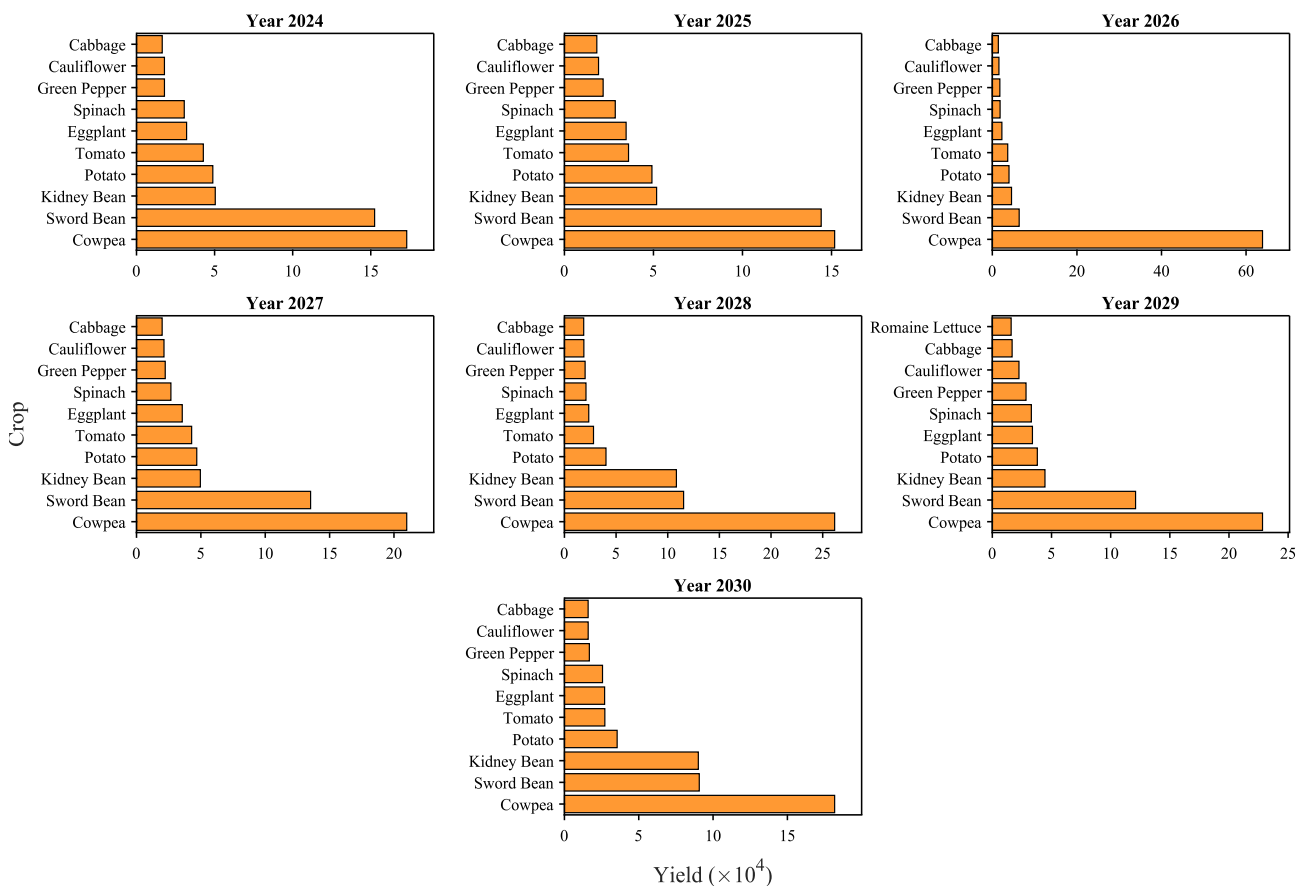


Fig. 3. Model-derived total output of the top ten crops in scenario 1.

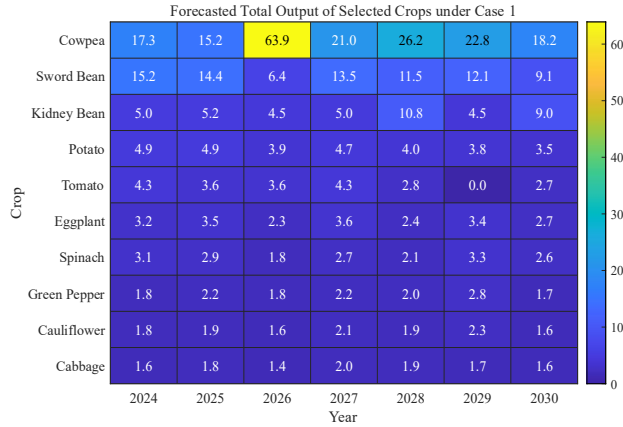


Fig. 4. Heatmap of model-derived total output of the top ten crops in scenario 1.

Figure 3 shows the model-derived output allocation for the ten selected crops under Scenario 1 for the 2024–2030 planning horizon. The model allocates more area to several dominant crops over the planning period. Cowpea, sword bean, and kidney bean have high output because the model requires legumes at least once every three years to meet rotation and soil-fertility requirements (Galán-Martín *et al.*, 2015; Saracian *et al.*, 2025). Profit, rotation, and land constraints jointly determine the Scenario 1 allocation.

Figure 4 maps the model-derived output allocation for the ten selected crops in Scenario 1 over the seven-year planning horizon. Cowpea has the highest output and peaks in 2026. Sword bean and kidney bean also remain relatively high. Cabbage, cauliflower, green pepper, spinach, eggplant, tomato, and potato have lower output. The allocation is therefore weighted toward legumes, while output for the other crops changes less over time.

Figure 5 shows planting area by plot in Scenario 1. Only plots with positive allocations are displayed, so the labels differ by year. Flat dryland and terraces account for

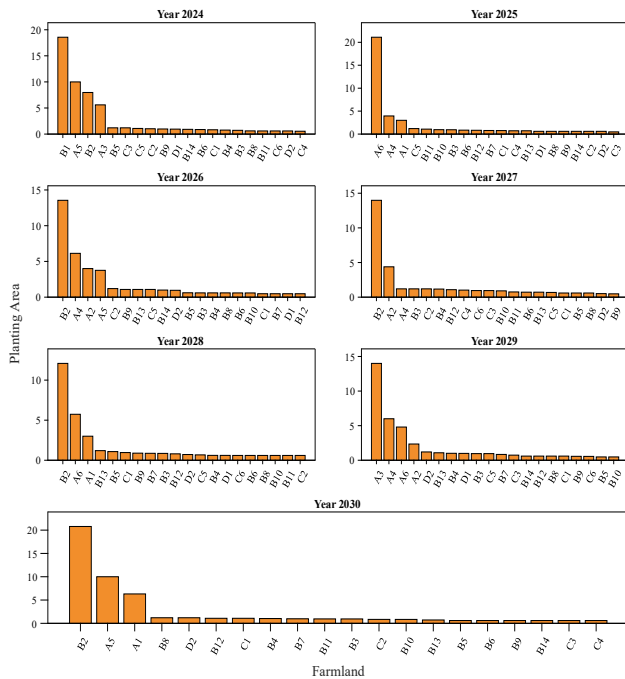


Fig. 5. Model-derived planting area of the top 20 selected plots in scenario 1.

the largest allocated areas, followed by slopes and irrigated fields. Greenhouse area is limited, which restricts the contribution of greenhouse plots to total profit.

Figures 6 and 7 show output and planting area allocations for Scenario 2. Cowpea again has the largest planted area, while sword bean and kidney bean vary across years. Flat dryland and terraces remain the main land types used in the allocation.

Optimization results under market uncertainty: For the uncertainty analysis, sales volume, yield per mu, crop cost, and selling price were allowed to vary. A multi-objective model and PSO were used to identify planting plans across these conditions.

Figures 8 and 9 show the convergence curves for the two uncertainty scenarios. In Figure 8, the two trajectories are similar. Improvement slows as the iteration count increases, and the objective values approach a stable range.

In Fig. 9, the objective value becomes nearly stable after about 400 iterations and remains close to 18,000,000. Subsequent oscillation is small.

Across the two scenarios, the convergence curves reach stable values with limited variation near the end of the search. This suggests that the tested PSO settings produced consistent solutions for the specified constraints. As with the genetic-algorithm results, the observed stabilization indicates convergence of the tested runs but does not by itself establish global optimality. Demonstrating global optimality would require exhaustive enumeration or a mathematically derived optimality bound, which is beyond the scope of the present heuristic optimization framework.

Figure 10 shows output under uncertain Scenario 1. Cowpea output declines in 2028, while legume output rises in the following year. This pattern reflects the requirement to include legumes at least once every three years. Figure 11 shows that flat dryland receives the largest area allocation, followed by terraces.

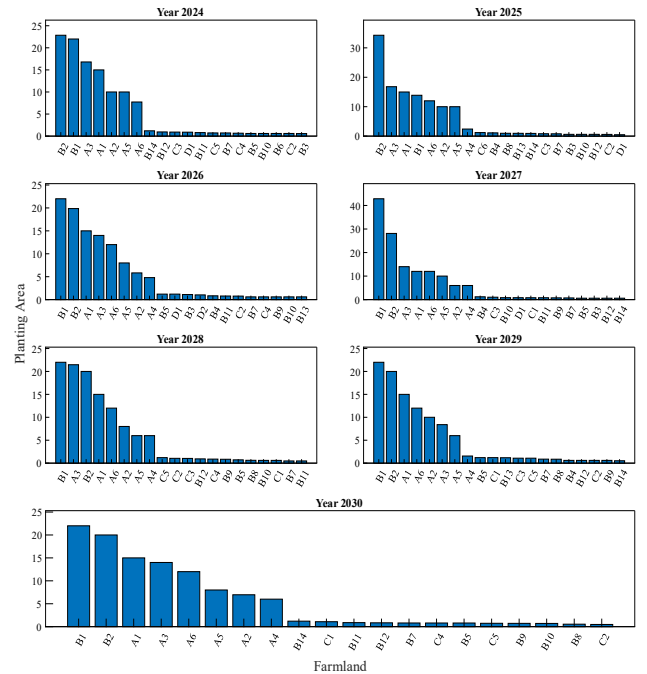


Fig. 7. Model-derived planting area of the top 20 selected plots in scenario 2.

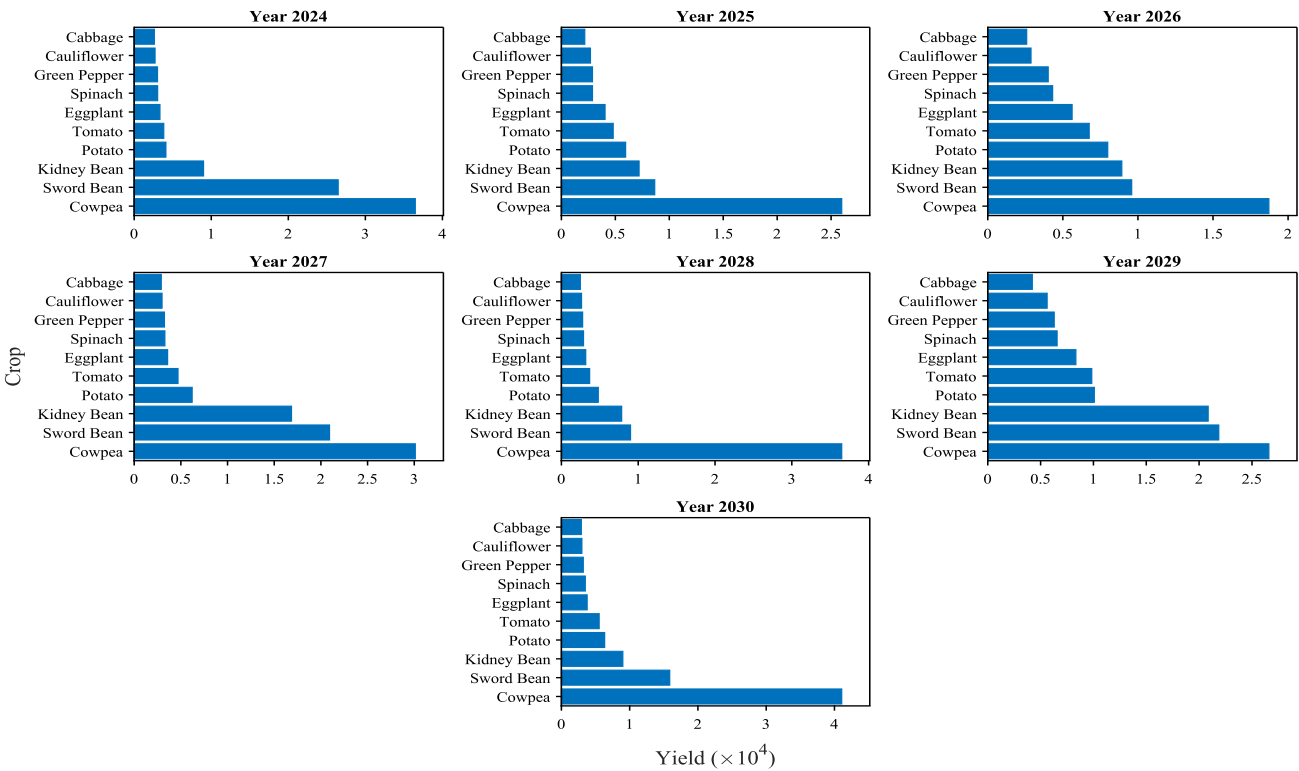


Fig. 6. Model-derived total output of the top ten crops in scenario 2.

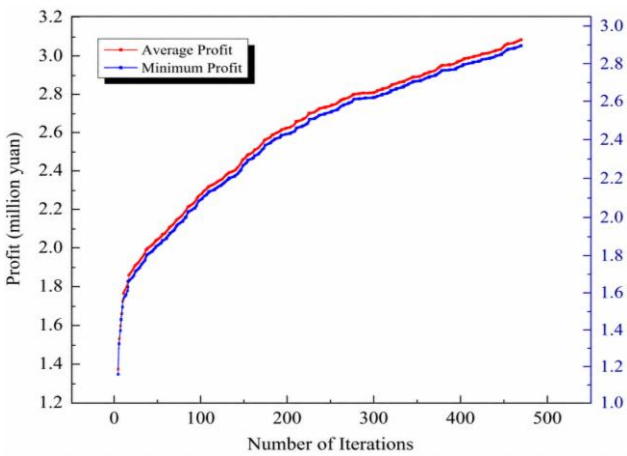


Fig. 8. The trajectory of convergence for scenario 1.

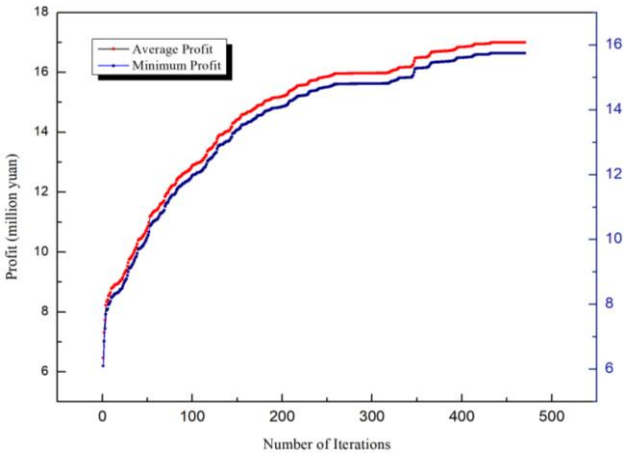


Fig. 9. The trajectory of convergence for scenario 2

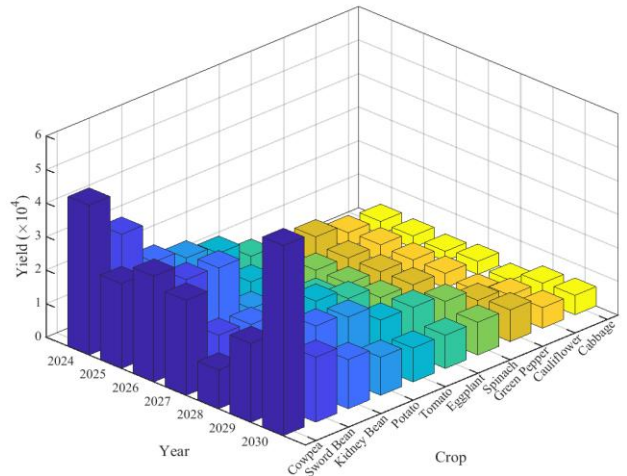


Fig. 10. Model-derived total output of selected crops in scenario 1.

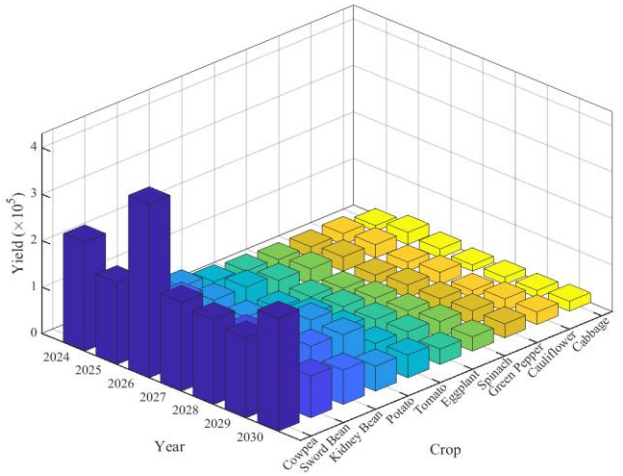


Fig. 12. Model-derived total output of selected crops in scenario 2.

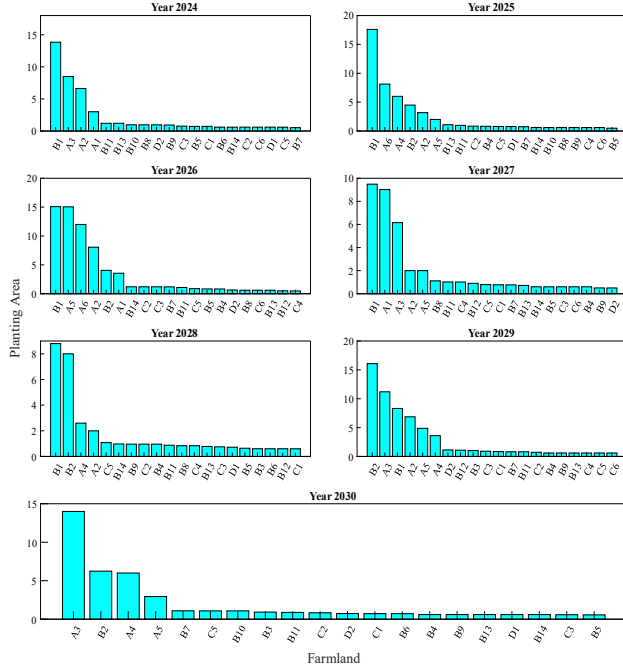


Fig. 11. Model-derived planting area of different plots in scenario 1.



Fig. 14. Correlation heatmap of per-mu yield, planting cost and unit sales price.

Figure 12 shows selected-crop output under uncertain Scenario 2. Cowpea output declines in 2024 and 2027 and rises in the subsequent years. Figure 13 shows that flat dryland and terraces receive most of the allocated planting area.

Correlation analysis and improved model evaluation: Correlation analysis and multiple linear regression were used to examine how crop-related economic variables relate to planting decisions. Figure 14 reports correlations among yield per mu, unit sales price, and planting cost.

Yield per mu and planting cost had a moderate positive correlation ($r = 0.624$). Higher-yield crops in the dataset tended to have higher costs, which may reflect greater use of inputs such as fertilizer, pesticides, labour, or production technology.

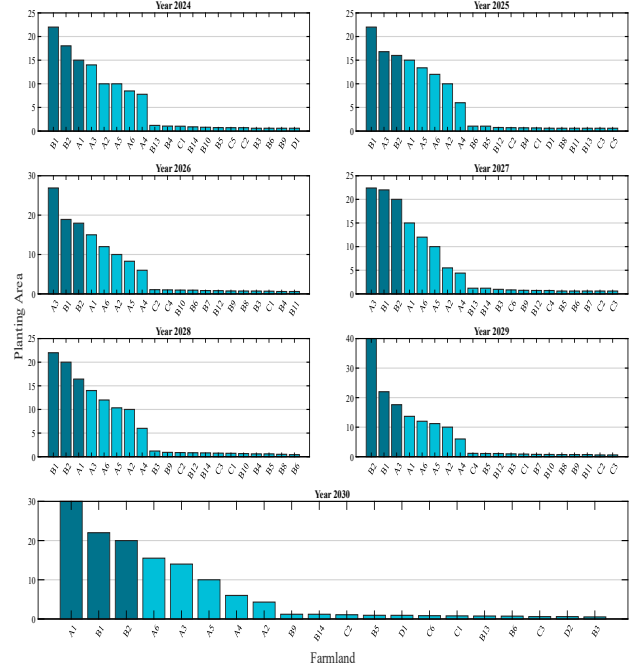


Fig. 13. Model-derived planting area of different plots in scenario 2.

Yield per mu and unit sales price were almost unrelated in the data ($r = -0.08$). Unit price therefore did not show a clear linear response to changes in yield.

The association between yield per mu and unit price was negligible.

Planting cost and unit sales price had a positive correlation of 0.40. In this dataset, crops with higher input costs also tended to have higher selling prices.

Figure 15 presents the regression fitting results based on 58 original observations ($n = 58$). The observed values represent the sales-volume observations in the regression dataset, while the predicted values are the corresponding values estimated by the fitted regression model. The regression analysis produced an R^2 of 0.835 and an adjusted R^2 of 0.816. Figure 15 was regenerated directly from these original regression observations without adding or reconstructing additional plotting points.

Figure 16 shows that the objective value changed little after about 250 iterations and remained close to 3,500,000.

Figures 17 and 18 show total output and planting area allocation. Total output rises in a three-year pattern. Most allocated area is flat dryland or terraced land, and flat-dryland area is lower in 2026 than in the other years.

The regression model was evaluated using the 58 complete observations used for model estimation. The resulting MSE was 143,050,473.26, RMSE was 11,960.37, MAE was 7,783.66, MAPE was 115.04%, SMAPE was 68.03%, and SSE was 8,296,927,449.14. The model produced an R^2 of 0.835 and an adjusted R^2 of 0.816. These metrics provide a quantitative assessment of the in-sample regression fit to the observed sales-volume data. The relatively high MAPE should be interpreted cautiously because percentage-based error measures can be sensitive to observations with relatively small sales volumes. These regression metrics do not constitute independent validation of the multi-year planting allocation results.

$$\beta_0 = 5185.191\beta_1 = 1.393\beta_2 = 0.104\beta_3 = -275.077\beta_4 = -4.805\beta_5 = 1.7\beta_6 = -640.801$$

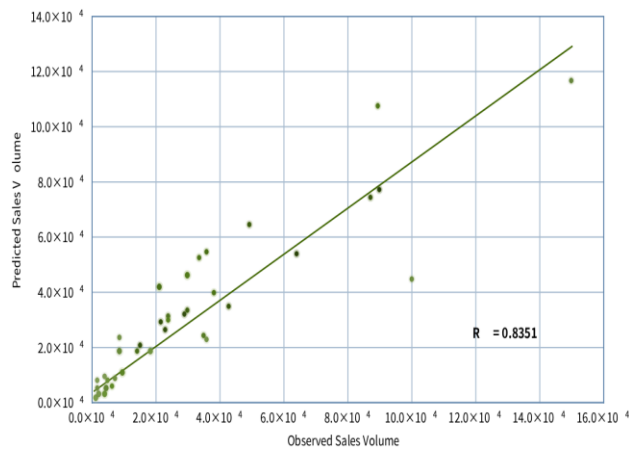


Fig. 15. Regression fitting results based on observed and predicted values ($n = 58$).

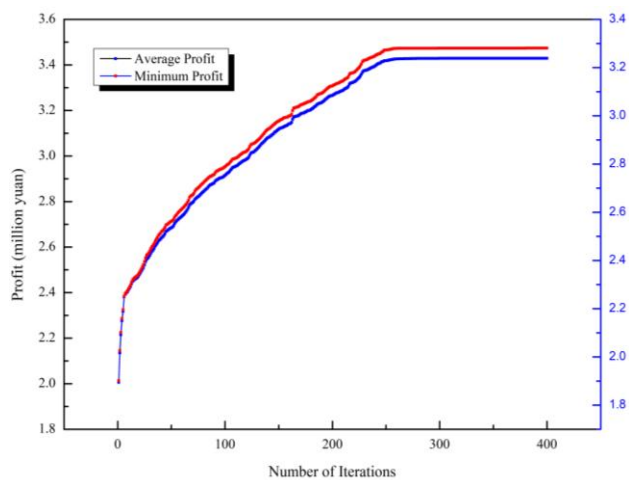


Fig. 16. Convergence curve.

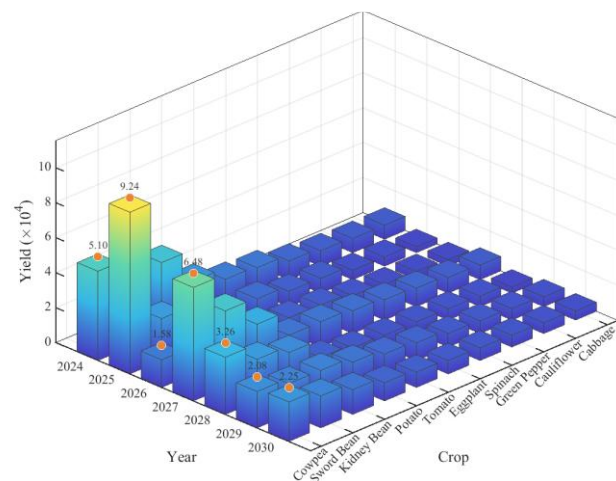
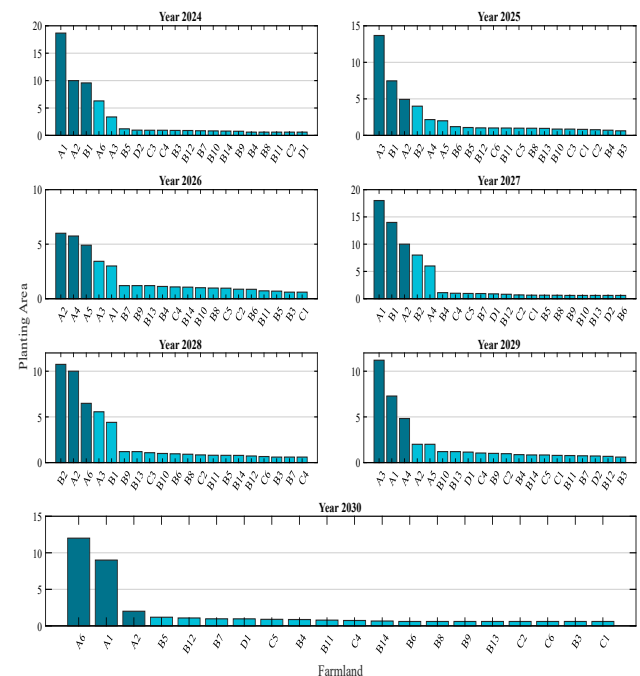


Fig. 17. Model-derived total output of selected crops.

Discussion

The results suggest that crop allocation should vary by land type, crop suitability, and expected economic return rather than be uniform across plots. Rotation constraints, particularly those for legumes, prevent the model from selecting only the most profitable short-term options. The model nevertheless has important limits. It assumes stable

climatic conditions, which does not account for the effects of extreme weather (Park *et al.*, 2025) or biological threats (Clement *et al.*, 2025) on crops such as edible fungi. It also simplifies the marginal costs associated with fragmented plots, and does not fully incorporate micro-level agronomic configurations, such as the specific impacts of plant density and row spacing on final yield (Zuo *et al.*, 2024). It is important to distinguish the findings derived from the present study from conclusions that have already been established in previous agricultural optimization research. Previous studies have demonstrated the usefulness of crop-land suitability assessment, crop rotation constraints (Dury *et al.*, 2012), evolutionary optimization methods (Sarker & Ray, 2009; Li *et al.*, 2025), and uncertainty analysis in crop-planning decisions (Xu *et al.*, 2025) in agricultural planning. The findings of the present study, however, arise from the integrated application of these factors within the fragmented and heterogeneous mountainous farmland setting examined here. In particular, the present results show how crop-land suitability, multi-year rotation requirements, legume constraints, planting-concentration limits, surplus-production treatment, and market uncertainty jointly influence the resulting crop and plot allocation under the specified planning scenarios. These findings are therefore specific to the modeled planning context and should not be interpreted as universally established relationships.



planning solutions for the specified scenarios rather than reconstructing or forecasting historical planting behaviour.

Independent observations of planting areas, yields, sales volumes, prices, and production costs for the corresponding planning years were not available for out-of-sample validation. Accordingly, the reported results should be interpreted as conditional planning solutions rather than predictions with established forecast accuracy. Future work should validate the framework against independent observed data as they become available and should evaluate its performance using out-of-sample error and robustness measures.

A further methodological limitation is that the convergence of the heuristic optimization procedures does not provide a formal guarantee of global optimality. The convergence curves demonstrate stabilization of the objective values for the tested parameter settings, but they do not establish that no better feasible solution exists. Global-optimality verification would require exhaustive search, a valid optimality bound, or an exact optimization procedure for the full problem, which is computationally beyond the scope of the present study. Accordingly, the reported solutions are interpreted as high-quality converged feasible solutions for the specified planning scenarios rather than formally proven global optima.

Future work could add stochastic weather, d market processes and dynamic measures of soil fertility (Borthakur *et al.*, 2025; De Barros *et al.*, 2026) when scheduling legume rotations.

Conclusion

This study converts data on cultivated land, crop suitability, and planting constraints into a crop-planning optimization model for mountainous rural North China. The model uses land use, crop distribution, cost, and market information to generate multi-year planting plans. Flat dryland and terraces receive most of the field-crop allocation, whereas greenhouse land is allocated mainly to vegetables and edible fungi. Legumes remain important because they satisfy the rotation constraint and support soil-fertility management. The model can be extended by incorporating climate and market information as it becomes available.

Declaration of Conflicting Interests: The author(s) declared no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

Data Sharing Agreement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding: The author(s) received no financial support for the research, authorship, and/or publication of this article.

AI usage Disclosure: This article utilized AI solely for English language polishing and editing, with no other applications.

References

- Benini, M., P. Detti and L. Nerozzi. 2025. Optimization models and algorithms for sustainable crop planning and rotation: an arc flow formulation and a column generation approach. *Omega*, 135: 103320.
- Borthakur, M., B. Bora, N. Dutta, D. Pegu, S. Bhattacharjya, R. Saikia and K. Saikia. 2025. Studies on soil physicochemical properties of different crop ecosystems for future cropping and soil management strategies in Jorhat District, India. *J. Exp. Agric. Int.*, 47(7): 557-565.
- Clement, W.J.J., K. Kalpana, K.E.A. Aiyathanan, M. Ramakrishnan, A. Kandan, K. Manonmani, I. Yesuraja, K.G. Sabarinathan, M.L. Mini, M. Shanthi, J. Rajangam and A. Punitha. 2025. Exploring the perilous nature of Phytophthora: insights into its biology, host range, detection, and integrated management strategies in the fields of spices and plantation crops. *Plant Pathol. J.*, 41(2): 121-139.
- De Barros, B., J.K.D. Martins, T.S. Duque, V.P. Sallin, A.A. Fernandes, H.C.J. Dos Santos, L.F.T. De Menezes, F.S. Machado, J.E.S. Paes, A.L. Nascimento and F.R. Pires. 2026. Ecological restoration of oil-impacted areas in restinga: soil management strategies and native species planting techniques. *Ecol. Eng.*, 222: 107792.
- Dury, J., N. Schaller, F. Garcia, A. Reynaud and J.E. Bergez. 2012. Models to support cropping plan and crop rotation decisions. A review. *Agron. Sustain. Dev.*, 32(2): 567-580.
- Felsman, J.H., A. Mobli, J.J. Nunes, R.P. DeWerff, D.E. Stoltenberg, C.A. Baxter and R. Werle. 2025. Termination strategies for high biomass cereal rye cover crop in soybean planting green systems. *Agrosyst. Geosci. Environ.*, 8(3): e70163.
- Filippi, C., R. Mansini and E. Stevanato. 2017. Mixed integer linear programming models for optimal crop selection. *Comput. Oper. Res.*, 81: 26-39.
- Galán-Martín, A., C. Pozo, G. Guillén-Gosálbez, A. Antón Vallejo and L. Jiménez Esteller. 2015. Multi-stage linear programming model for optimizing cropping plan decisions under the new Common Agricultural Policy. *Land Use Policy*, 48: 515-524.
- Groot, J.C.J., G.J.M. Oomen and W.A.H. Rossing. 2012. Multi-objective optimization and design of farming systems. *Agric. Syst.*, 110: 63-77.
- Jiang, J. 2025. Optimization research on crop planting strategies based on linear programming model. *Agric. For. Econ. Manag.*, 8(1): 93-99.
- Krityakierne, T., P. Sinpayak and N. Khiripet. 2025. GIS spatial optimization for agricultural crop allocation using NSGA-II. *Inf. Process. Agric.*, 12(2): 139-150.
- Kümmerer, R., D. Heuermann, F. Laidig, H.P. Piepho and B. Bauer. 2025. Strategies to develop simple multi-species cover crop mixtures to enhance aboveground biomass and quality. *Field Crops Res.*, 333: 110014.
- Li, C., Z. Su, Y. Nie, Z. Ye, J. Li, J. Wang, Z. Yang, X. Li, W. Zeng and Y. Chen. 2025. Scientific planning of dynamic crops in complex agricultural landscapes based on adaptive optimization hybrid SA-GA method. *Sci. Rep.*, 15: 28992.
- Mardani Najafabadi, M., S. Ziaee, A. Nikouei and M. Ahmadpour Borazjani. 2019. Mathematical programming model (MMP) for optimization of regional cropping patterns decisions: a case study. *Agric. Syst.*, 173: 218-232.
- Park, J., S.H. Lee, J. Lee, S.H. Wi, T.C. Seo, J.H. Moon and S. Jang. 2025. Growing vegetables in a warming world - a review of crop response to drought stress, and strategies to mitigate adverse effects in vegetable production. *Front. Plant Sci.*, 16: 1561100.
- Sakpal, A. 2026. Resilient strategies to mitigate negative heat stress consequences in seed production of field crops. *Asian Res. J. Agric.*, 19(1): 54-68.

- Saraeian, Z., C. Farrell and N.S.G. Williams. 2025. Using plant functional traits and life strategies of annual species to inform green roof plant selection. *Perspect. Plant Ecol. Evol. Syst.*, 69: 125900.
- Sarker, R. and T. Ray. 2009. An improved evolutionary algorithm for solving multi-objective crop planning models. *Comput. Electron. Agric.*, 68(2): 191-199.
- Shilpa, S., B.S. Lalitha, K.H. Nagaraj, K.N.K. Murthy, H.M. Jayadeva, M.S. Sannagoudar, R. Vasundhara, T.G. Amrutha, M.B. Nagangoudar, G. Raddy, R.M. Salmankhan, C. Kavyashree, H.P. Thimmaiah and S.B. Hosalli. 2025. Crop diversification strategies for yield enhancement in finger millet-based cropping systems. *Front. Agron.*, 7: 1644473.
- Wang, C., X. He, D. Yan and C. Shao. 2025. Research on optimal crop planting strategy based on NSGA-II algorithm. *Int. Core J. Eng.*, 11(4): 345-352.
- Xu, Y., X. Yang and J. Zhang. 2025. Optimization of crop planting strategy based on Monte Carlo simulation. *Acad. J. Math. Sci.*, 6(2): 41-47.
- Yu, W., W. Yi and D. Xie. 2025. Research on crop planting strategy problem based on multi objective dynamic linear programming model. *Int. J. New Dev. Eng. Soc.*, 9(2): 41-45.
- Zeng, X., S. Kang, F. Li, L. Zhang and P. Guo. 2010. Fuzzy multi-objective linear programming applying to crop area planning. *Agric. Water Manag.*, 98(1): 134-142.
- Zuo, T.Y., L.Y. Ji and Y.S. Liu. 2025. Optimal crop planting strategy based on particle swarm algorithm. *Eurasia J. Sci. Technol.*, 7(5): 17-20.
- Zuo, W., B. Wu, Y. Wang, S. Xu, M. Chen, F. Liang, J. Tian and W. Zhang. 2024. Optimal row spacing configuration to improve cotton yield or quality is regulated by plant density and irrigation rate. *Field Crops Res.*, 305: 109187.