Journal of Computer Science and Artificial Intelligence
ISSN: 3078-8242 (Print), ISSN: 3078-2309 (Online) | Vol. 3, No. 1, 2025

Improved Genetic Algorithm-Driven Multi-Objective Optimization Model for Sustainable Tourism Development

Shuhuan Gao *

Xidian University, Xi'an, China

* **Corresponding author**: 24009300040@stu.xidian.edu.cn

Abstract: This paper constructs a multi-objective optimization model and focuses on its application in optimizing tourism resource allocation based on genetic algorithm. Firstly, the study determines the weights by establishing a multi-objective function system containing multiple indices, combining the hierarchical analysis method (AHP) and entropy weighting method, and solves the optimal solution by genetic algorithm. Among them, the genetic algorithm adopts Latin hypercubic sampling to generate the initial population and balances the global search and local convergence by dynamically adjusting the replication strategy of indices, and finally obtains the optimal solution under the constraints. Secondly, the study constructs a nonlinear model based on the Sigmoid function, and the analysis is further verified by the Shapley value. Finally, through parameter optimization and constraint design, the practical application value of the model on multiple decision variables is ensured to provide a scientific basis for sustainable tourism policy formulation.

Keywords: Multi Objective Optimization; Genetic Algorithm; Hierarchical Analysis; Latin Hypercube.

1. Introduction

In the field of sustainable tourism research [1], traditional methods mostly rely on the hierarchical analysis method (AHP) [2] or single-objective optimization model, which is difficult to take into account the dynamic balance, and the subjective weight assignment is easy to lead to result bias. Aiming at this limitation, this paper proposes a comprehensive model integrating multi-objective optimization and intelligent algorithms, aiming to provide a scientific decision-making tool for tourism resource allocation.

In this paper, firstly, a four-dimensional evaluation system is constructed, and subjective and objective weights are determined through the combination of hierarchical analysis method and entropy weight method [3]. Secondly, the improved genetic algorithm is introduced for global optimization [4], the population is initialized through Latin hypercubic sampling [5], and the replication strategy of dynamically adjusted indices balances global search and local convergence. In addition, the study innovatively constructs a nonlinear model based on the Sigmoid function [6], and verifies the key contribution to the integrated benefits through Shapley value analysis, which enhances the adaptability of the model. Compared with traditional research, this paper enhances the objectivity of evaluation through the fusion of subjective and objective weights, and introduces cultural factors to expand the model application scenarios, providing a quantifiable decision-making framework for multi-dimensional sustainable tourism planning.

2. Analysis of Sustainable Tourism

In this section, tourism revenue, infrastructure stress index, environmental impact index and social satisfaction index are used as state variables that directly affect the economic, environmental and social aspects of tourism development.

2.1. Model Building

(1) Plot the images of the four sets of actual data points separately and perform various linear and nonlinear regressions by comparing the cumulative residuals to determine the type of function with the best fit.

(2) Establish a multi-objective function

Maximize tourism revenue:

$$R\left(x_1,x_2,x_3\right)=r_1x_1\left(1+r_2x_2\right)\left(1-r_3x_3\right) \tag{1}$$

r_1 denotes the average spending per tourist (USD/person), which is a comprehensive value derived from statistical analysis of historical data, covering the average spending of tourists on accommodation, food and beverage, shopping, entertainment, etc. r_2 denotes the impact of tax rate adjustment on tourism revenue. r_3 denotes the impact of tax rate adjustment on tourism revenue. r_4 denotes the impact of tax rate adjustment on tourism revenue. r_5 denotes the impact of tax rate adjustment on tourism revenue.

$$\mathrm{I}\left(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3\right)=\mathrm{m}_1\mathrm{x}_1+\mathrm{m}_2\mathrm{x}_2+\mathrm{m}_3\mathrm{x}_3 \tag{2}$$

m_1 as the number of tourists increases, so does the demand for facilities such as transportation and accommodation, the value of which depends on the current state of the infrastructure and sensitivity to changes in the infrastructure. m_2 Indicates the effect of adjustments to tourism taxes and fees. m_3 Indicates the effect of adjustments to tourism taxes and fees on tourism revenue.

$$\mathrm{E}\left(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3\right)=\mathrm{c}_1\mathrm{x}_1+\mathrm{c}_2\ln\left(1+\left|\mathrm{x}_2\right|\right)+\mathrm{c}_3\mathrm{x}_3+\mathrm{d}_1[3] \tag{3}$$

Maximizing Social Satisfaction Index:

$$\mathrm{S}\left(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3\right)=-\mathrm{k}_1\mathrm{x}_1+\mathrm{k}_2\left(1-\mathrm{x}_2\right)+\mathrm{k}_3\mathrm{x}_3+\mathrm{b}_2[4] \tag{4}$$

K_1 Indicates the negative impact of increased tourist numbers on social satisfaction. K_2 Reflects the effects of tax and fee adjustments on social satisfaction. K_3 reflects the positive impact of investing in environmental protection, improving infrastructure and developing community programs on social satisfaction.

(3) After forward and normalization, the four state variables are multiplied by the corresponding weights and summed to obtain the main function.

$$\begin{aligned} Z = w_1 \frac{R(x_1,x_2,x_3)-R_{\min}}{R_{\max}-R_{\min}} - w_2 \frac{I(x_1,x_2,x_3)-I_{\min}}{I_{\max}-I_{\min}} \\ -w_3 \frac{E(x_1,x_2,x_3)-E_{\min}}{E_{\max}-E_{\min}} + w_4 \frac{S(x_1,x_2,x_3)-S_{\min}}{S_{\max}-S_{\min}} \end{aligned} \tag{5}$$

$$\text{s.t.} \begin{cases} \left| \dfrac{x_{1s}}{\sum_{s=1}^{n} x_{1s}} - \dfrac{1}{n} \right| \le \alpha \\ w_1 + w_2 + w_3 + w_4 = 1 \\ E(x_1,x_2,x_3) \le E_{\max} \\ I(x_1,x_2,x_3) \le I_{\max} \\ x_1 < x_{1\max} \\ 0 < x_2 < x_{2\max} \\ x_{3\min} < x_3 < x_{3\max} \end{cases} \tag{6}$$

α denotes the maximum deviation of the proportion of visitors from the average proportion in each season, and s denotes the season.

2.2. Model Solving

2.2.1. Eigenvalue method for determining exponential weights

Table 1. Judgment Matrix

	R	I	E	S
R	1	3	5	7
I	1/3	1	3	5
E	1/5	1/3	1	3
S	1/7	1/5	1/3	1

As shown in Table 1, we construct the judgment matrix U, calculate the maximum eigenvalues of the judgment matrix U and the corresponding eigenvectors W. After calculating the eigenvectors, we normalize them to obtain the approximate values of the weight variables of the four states.

$$W = \frac{w_i}{\sum_{j=1}^{n} w_j} \tag{7}$$

$$CI = \frac{\lambda_{\max} - n}{n-1} \tag{8}$$

The consistency index is calculated using the above formula, and the average stochastic consistency index RI is about 0.90 when $n = 4$. The consistency ratio is calculated.

$$CR = \frac{CI}{RI} \tag{9}$$

We calculate $CR = 0.0892 < 0.1$, and the calculated weights are valid considering that the judgment matrix has consistency, and the basic judgment is reasonable. In addition, the weights were calculated using Hierarchical Analysis of Hierarchy (AHP) and Entropy Weighting Methods, respectively, and the results were found to be similar. This indicates that the weight distribution points to similar results from multiple dimensions, including subjective judgments and objective data characteristics.

2.2.2. Parametric Solution

To solve for the parameters of a linear function, the least squares method is used, e.g., infrastructure pressure index (1, 2, 3) = 1 1 + 2 2 + 3 3. Here the variables are written in matrix form.

$$X = \begin{bmatrix} x_{11} & x_{21} & x_{31} \\ x_{12} & x_{22} & x_{32} \\ \vdots & \vdots & \vdots \\ x_{1n} & x_{2n} & x_{3n} \end{bmatrix}, \vec{m} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}, \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{10}$$

$$\mathrm{X}^{\mathrm{T}}\mathrm{X}\overrightarrow{\mathrm{m}} = \mathrm{X}^{\mathrm{T}}\vec{\mathrm{y}} \tag{11}$$

Where y denotes the dependent variable, the infrastructure stress index.

Since $X^T X$ is invertible, we directly transform the equation into $\vec{m} = (X^T X)^{-1} X^T \vec{y}$ to solve for the parameter vector, and ultimately, we obtain the parameter values estimated by the least squares method in the function. $\mathrm{I}(\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3)$

For nonlinear functions, we use the maximum likelihood estimation method to find the values of parameters, such asR $(x_1, x_2, x_3) = r_1 x_1 (1 + r_2 x_2)(1 - r_3 x_3)$. By organizing the data, we obtain n sets of independently and identically distributed observations $(x_{1i}, x_{2i}, x_{3i}, y_i), i = 1,2,\cdots,n$, wherey y_i represents the corresponding observed values. We ultimately need to maximize the probability of these observed values occurring and find the corresponding parameter values. Since our data is discrete, we first need to construct the parameters of the parametersHere, we ignore the effects of random noise and assume $y_i = R(x_{1i}, x_{2i}, x_{3i})$, then the likelihood function is.

$$L(r_1, r_2, r_3) = \prod_{i=1}^{n} r_1 x_{1i} (1 + r_2 x_{2i})(1 - r_3 x_{3i}) \tag{12}$$

To facilitate calculation, we take the logarithm of the likelihood function.

$$\ln L(r_1, r_2, r_3) = \sum_{i=1}^{n} \left[\ln(r_1) + \ln(x_{li}) + \ln(1 + r_2 x_{2i}) + \ln(1 - r_3 x_{3i}) \right] \tag{13}$$

Here, r_I represents the average expenditure per visitor, which we obtain directly through weighted historical data. We only need to take the partial derivatives with respect to r_2 and r_3. Then set the partial derivatives to zero, obtaining a system of equations:

$$\begin{cases} \dfrac{\partial \ln L\left(r_1, r_2, r_3\right)}{\partial r_2} = \displaystyle\sum_{i=1}^{n} \dfrac{x_{2i}}{1 + r_2 x_{2i}} = 0 \\ \dfrac{\partial \ln L\left(r_1, r_2, r_3\right)}{\partial r_3} = -\displaystyle\sum_{i=1}^{n} \dfrac{x_{3i}}{1 - r_3 x_{3i}} = 0 \end{cases} \tag{14}$$

2.2.3. Genetic algorithm to solve the optimal solution

In terms of the algorithm, genetic algorithm is used, and we have made several improvements to the traditional genetic algorithm, mainly including the following points.

First, when the initial population is obtained by sampling, Latin Hypercube Sampling (LHS) is used to ensure that the initial population is uniformly distributed in the entire parameter space, which allows for a wider search of the parameter space and increases the number of searchable parameters.

As shown in Figure 1. The initial upward trend of the adaptive curve reflects the explorability of the algorithm to continuously improve the quality of the solution.

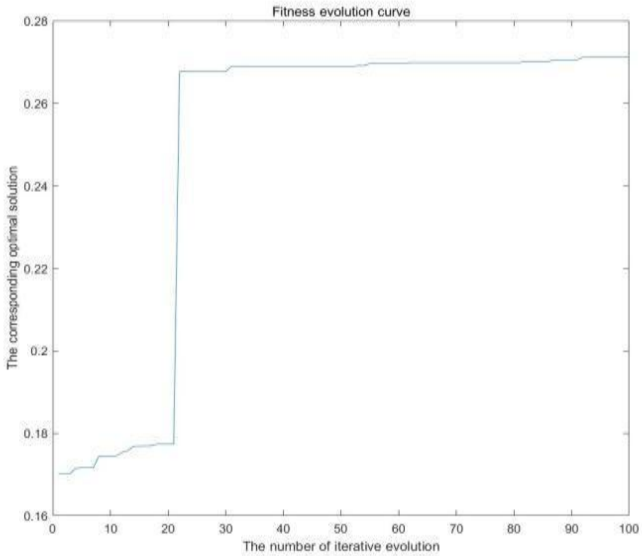

Figure 1. Fitness change curve

Secondly, during the copying operation, it's not just fitness that isusedas an evaluation criterion, but rather the value of an exponential function of fitness to characterize it. The exponent x is dynamically adjusted according to the iteration count. A smaller exponent at the beginning reduces the differences between individuals, further reducing the differences in selection, maintaining species diversity, andpreventing the algorithm from converging too early to a local optimum, which would affect the results. Now, it's necessary to search towards individuals with higher fitness, so $x = 2$ is used. Alarger exponent amplifies the differences between individuals with different fitness levels, and individuals with higher fitness have a greater probability of being copied, which accelerates the convergence of the entire population towards the optimal solution.

The specific implement ation process of the algorithm is as follows.

(1) Initialize parameters: Determine the basic parameters of the genetical gorithm, such as population size, individual length, crossover probability, and mutation probability. At the same time, set the range of decision variables (number of tourists, tax and fee adjustment ratios, proportion of support for environmental protection projects as a share of additional revenue), and the number of iterative evolution cycles. Define a coefficient structure that includes various coefficients related tototal tourism revenue, infrastructure pressure, environmental impact, and social satisfaction, as wellas the maxima and weights of various functions. These parameters and coefficient slay the foundation for subsequent calculations.

(2) Generate the initial population using Latin Hypercube Sampling: Use the lhs design function for Latin Hypercube Sampling to create an initial population gene_body consisting of multiple individuals, encoded in binary. Then, use the bintodec function to convert the binary encoding to decimal, obtaining decgene_body corresponding to actual variable values, preparing, preparing for the calculation of function values.

(3) Calculate the function values of the initial solutions and find the optimal solution: Substitute decgene_body into the calobjvalue function to Calculate the function values of the initial solutions and find the optimal solution. The objective function comprehensively considers factors such as total tourism revenue, infrastructure pressure, environmental impact, and social satisfaction, while checking constraint conditions. values of individuals that exceed the constraint range are set to a very small value. Use the max function to find the optimal solutionof the initial generation, recording its function value and corresponding value. The objective function values of individuals that exceed the constraint range areset toavery small value.

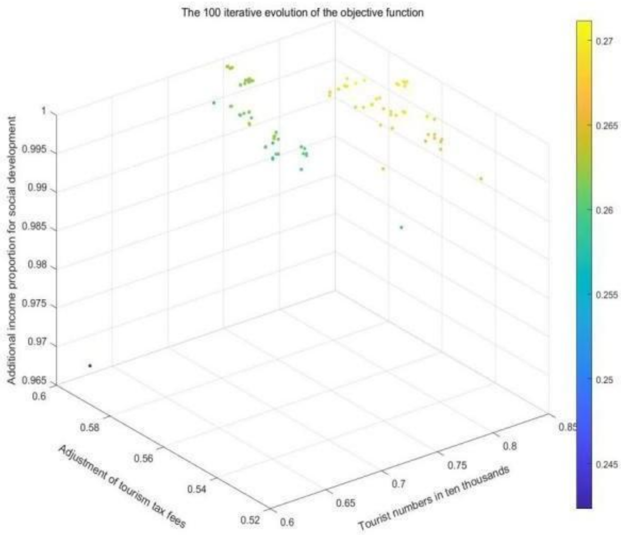

Figure 2. Distribution diagram of the optimal solutions in each generation

(4) Iterative evolution: In each iteration, first convert the previous generation's population gene_body to decimal decgene_body, calculate the function values fx and fitness values fit value. During the copy operation, use the roulette wheel strategy, calculating the selection probability based on dynamically adjusted exponents for individual replication. The crossover operation is performe donaneven number of individuals by randomly selecting crossover sites to exchange gene segments; for an odd number of individuals, first process the even number of individuals, and then add the remaining individuals back to the new population. The mutation operation randomly changes individual genes at the mutation probability. Calculate the function values and fitness values of the new population, compare the fitness of the new and old populations, and update the population. In each iteration, plot

the evolutionary graphs of the objective function, revenue function, infrastructure pressure function, environmental impact function, and social satisfaction function, and record the optimal solution of each generation.

As shown in the Figure 2, after the final iteration, the algorithm identifies the global optimum from the recorded best solutions of each generation, and outputs the position of the optimal solution (0.8490, 0.5999, 0.9974) with the corresponding best solution value of 0.2726.

3. Adaptation Analysis

3.1. Model Change

We introduced a new indicator, the Arts and Culture Tourism Attractiveness Index T, to measure the tourism attractiveness of city because of arts events and historic site preservation. To this index, several variables are added: time t (reflecting the cumulative effect of the arts), the number of music and cultural events organized per year n, and the investment in the preservation of historic and cultural sites x4. Based on a Sigmoid function, we construct a nonlinear relationship between these variables and the index:

L represents the maximum value of attractiveness, C_1 is the impact coefficient of the number of music and cultural events on the indicator T, and C_2 is the impact coefficient of investment in the protection of historical and cultural sites on the indicator T.

$$T\left(t,n,x_4\right)=\begin{cases}\dfrac{L}{1+e^{-k\left(t-t_0\right)+C_1\ln(n-1)+C_2x_4}}, & n\neq 0, x_4>x_{4\min}\\ 0, & n=0, x_4<x_{4\,\min}\end{cases} \tag{15}$$

3.2. Assessing the Impact of Destination Choice on The Importanceofvarious Measures Based on Shapley Values

The Shapley value is a concept in cooperative game theory used to measure the contribution of each participant to the total payoff. Here, we consider the five indicators R, I, E, S, and T as participants, and the comprehensive optimization indicator Z as the total payoff. Let $\mathrm{N}=\{\mathrm{R},\mathrm{I},\mathrm{E},\mathrm{S},\mathrm{T}\}$ be the set of all indicators. For each indicator $i\in N$, its Shapley value

$$\varphi_i=\sum\nolimits_{S\subseteq N\setminus\{i\}}\frac{|S|!(n-|S|-1)!}{n!}[Z(S\cup\{i\})-Z(S)] \tag{16}$$

Where $|\mathrm{S}|$ is the number of elements in the subset S, and n is the total number of indices.

When calculating the Z -values for each subset, due to the presence of multiplication terms between two indicators in our model, we cannot simply set unrelated terms outside the subset to zero, as this would also set the coefficients of the corresponding terms within the subset to zero.

After calculations, we found that the Shapley values for the indicators of Total Tourism Revenue (R) and Cultural Tourism Attractiveness (T) are larger, leadin gus to believe that measures should be taken to enhance cultural tourism attractiveness to increase total tourism revenue.

4. Conclusion

In this paper, a multi-objective optimization model is constructed for sustainable tourism development decision-making, and a dynamic balance of multiple aspects is achieved through the integration of improved genetic algorithm and multi-dimensional evaluation system. Firstly, the exploratory ability of the genetic algorithm is significantly enhanced by initializing the population through Latin hypercubic sampling and dynamically adjusting the selection pressure; secondly, the Shapley value method is introduced to quantify the contribution of each index to the comprehensive benefits, and it is found that the marginal contribution of the cultural tourism attractiveness index (T) and the tourism revenue (R) is the highest; finally, the model portrays the dynamic evolution law of the attractiveness through the nonlinear Sigmoid function, which enhances the explanatory power of complex socio-economic factors. Finally, the model portrays the dynamics of attractiveness through the nonlinear Sigmoid function, which enhances the explanatory power of complex socio-economic factors. The limitations of the study are that the impact of sudden public events on tourism demand is not considered, and the model parameters rely on the fitting of historical data. In the future, a dynamic weight adjustment mechanism can be introduced to optimize the frequency of parameter update with real-time data.

References

[1] SUN Yuting, DUAN Xiaoyun. Sustainable Tourism Management and Resource Protection Strategies [J]. Western Tourism, 2024, (20):13-15.

[2] Zhang Xiaorui. Research on the construction of evaluation index system of rural tourism resources based on improved hierarchical analysis [J]. Western Tourism, 2024, (21):42-44.

[3] Cui Yali. Research on Ecological Safety of Tourist Cities Based on AHM Model and Entropy Right Method [J]. Journal of Guiyang College (Natural Science Edition), 2024, 19(02): 43-47. DOI: 10.16856/j.cnki.52-1142/n.2024.02.007.

[4] Xu Xiangrong. Research on travel recommendation algorithm based on improved genetic algorithm with multi-source heterogeneous data [D]. Xi'an University of Technology, 2023. DOI: 10.27398/d.cnki.gxalu.2023.001863.

[5] Li G.S., Zhang X., Wen Q., et al. Evolutionary alignment optimization of Latin hypercube design method [J]. Journal of National University of Defense Technology, 2024, 46(03):150-157.

[6] ZHANG Bingyan, CHEN Xiaomei, ZHONG Bo. Filter-x quaternion least-mean-square algorithm for Sigmoid transform [J]. Noise and Vibration Control, 2024, 44(02):57-62.