

# Construction of Flowering Period Prediction Model based on Mathematical Modeling

Baotong Zhang<sup>1</sup>, Nuan Feng<sup>2</sup>, Siru Liu<sup>2</sup>, Peng Gao<sup>2</sup>, Wenling Yang<sup>2</sup>

<sup>1</sup> School of Liaoning University of Science and Technology Liaoning, Benxi 117004, China

<sup>2</sup> School of Liaoning University of Science and Technology Liaoning, Benxi 117004, China

---

## Abstract

In the context of the rainy weather during the Qingming Festival which may affect the outing for flower viewing, this article conducts an in-depth analysis and modeling on the topic of flower viewing during the Qingming Festival. By using historical data and information about flowers, a comprehensive optimization model is established to predict the weather and plan the flower viewing routes. We predict whether it will rain in a city by establishing a Holt-Winters seasonal model based on historical weather data (such as atmospheric temperature, atmospheric pressure, relative humidity, etc.) to predict whether it will rain during the Qingming Festival. We also establish a forecast model for the opening time and flowering period of flowers such as apricot blossoms and rape flowers in 2026, using the data of the initial flowering time of each year to build a prediction model for flowering prediction, thereby predicting the opening time and flowering period of flowers in 2026. Based on the weather forecast for the Qingming Festival in 2026 and the predicted flowering period, a self-guided tour plan for flower viewing during the Qingming Festival is drafted. By taking into account various factors such as the flowering time and period of flowers, the distance between cities, transportation costs, and weather conditions, an objective optimization model is established to plan the best travel route. This predicts the best route for flower viewing during the Qingming Festival in 2026 and provides a measure for the government to extend the "flower viewing economy" industrial chain, allowing the "flower viewing economy" to have an "extra-long flowering period". In light of the characteristics of the flowering period of different flowers, measures to extend the flowering period are provided, and corresponding measures and suggestions are given to the government and local flower viewing parks to extend the "flower viewing economy" industrial chain.

## Keywords

Holt-Winters; Seasonal Model; Probability of Rain; Flowering Prediction.

---

## 1. Research Background

The Qingming Festival, a natural solar term as well as an important traditional Chinese festival, carries the cultural memories and emotional attachments of the Chinese nation for thousands of years. During this time, plants sprout and flowers bloom, presenting a vibrant scene in nature. Frequent rainfall occurs during the Qingming Festival, mainly due to the convergence of cold and warm air (frontal rain) and convective precipitation caused by the unstable thermal conditions in the atmosphere in spring. The probability of rainfall in the middle and lower reaches of the Yangtze River (such as Hunan and Jiangxi) exceeds 50%, while in South China, it is mainly intermittent precipitation. In recent years, due to global warming, the number of rainy days in some areas (such as Hunan) has decreased but the intensity has increased, weakening the traditional feature of "drizzling rain" and

increasing the occurrence of strong convective weather. According to phenological patterns, the Qingming Festival (early April) falls in the transition period from mid-spring to late spring, with significant differences in the flowering periods of flowers in the north and south. In the north (such as Beijing), early spring flowers (mountain peach, cherry blossoms) are in full bloom from late March to mid-April, while peonies and tulips are in full bloom in mid to late April. In the south (such as Xiamen and Chongqing), cherry blossoms, azaleas, rape flowers, etc. reach their best viewing period from late March to early April. Historical landscapes in places like Beijing (Yuyuantan Cherry Blossoms, Jingshan Peonies), Luoyang (Peonies), and Wuyuan (Rape Flowers) attract tourists. However, the study of the rainfall phenomenon during the Qingming Festival can be conducted from multiple perspectives, including meteorological causes, regional differences, the impact of climate change, and cultural significance. Therefore, the refinement of climate models, especially in improving predictions of temperature and rainfall, has become a focus in related fields.

## 2. Model Assumption

To facilitate the establishment of the model and ensure its reliability, we first make some assumptions to make the model more complete and the prediction results more comprehensive and reasonable.

- 1) We assume that the historical meteorological data collected for the Qingming Festival is accurate and reliable. Based on the weather classification standard, the range of rainfall and duration of "rainy and misty" weather can be used to predict the weather conditions during the Qingming holiday in 2026.
- 2) In this study, we assume that the economic benefits generated by flower tourism are solely dependent on the duration of tourists' visits and the blooming period of flowers, without considering other potential sources of income such as transportation, accommodation, and dining. At the same time, the economic benefits of related industries are included in the calculation of flower tourism benefits at a certain proportion.
- 3) We further assume that the travel time between different cities is constant and not affected by external factors such as weather and traffic. The calculation of travel time only involves the speed of high-speed rail, boarding and alighting time, and the transfer time from the station to the scenic spot.
- 4) All benefits are calculated on an hourly basis, and there is a fixed proportion relationship between flower tourism benefits and related industry benefits. In addition, there is a linear correlation between the blooming period of flowers and the duration of tourists' visits.
- 5) Under rainy conditions, only tourists' visiting intentions are affected, resulting in a decrease in the number of tourists on rainy days, while other economic activities are not affected. The impact of rainfall on flower tourism benefits is adjusted by multiplying a specific coefficient.
- 6) Except for flower tourism and related industry benefits, other economic factors (including local government financial support, other tourism activities, etc.) are not considered to directly affect economic benefits in this analysis. Additionally, we assume that the local government can effectively promote policies to extend the blooming period of flowers and that such policies will significantly attract tourists after implementation.

## 3. Data Pre-processing

1) Based on the analysis of the dataset involved in this research, the data source is mainly from the Weather.com website. Since 1981, this website has systematically collected historical weather data and real-time weather forecast information for 241 countries around the world. The website's database is updated every three hours, with a total of eight updates per day. You can search for the weather conditions of the city you need on the website.

2) As the research focuses on the weather conditions during the Qingming Festival holiday, we have established a mathematical model on the basis of appropriate simplification. Here, we have selected as many years as possible of weather records on the day of the Qingming Festival to predict the weather conditions of the Qingming Festival in 2026, and then carry out the necessary data screening.

3) In the traditional Chinese festival system, the Qingming Festival undoubtedly holds a significant position. It is not only an important part of China's traditional festivals but also a key node in the "Twenty-Four Solar Terms" system. This festival also has alternative names such as the Tomb Sweeping Festival, the Spring Outing Festival, and the Third Month Festival. Its time is set on the 15th day after the Spring Equinox in the lunar calendar, which corresponds to around April 5th in the Gregorian calendar. It is not fixed on April 4th or April 5th.

4) Based on this study, literature related to the flower bloom prediction of five types of flowers was retrieved through a search on CNKI. This study collected the blooming time data and the initial information of five major flowers including apricot blossoms, rape flowers, rhododendrons, cherry blossoms, and peonies in the regions of East China, Central China, Southwest China, Northwest China and North China.

5) Based on the data collected in this study, the flowering times and the initial flowering periods of 25 types of flowering plants - apricot blossoms, rape flowers, rhododendrons, cherry blossoms, and peonies - in the eastern China region, central China region, southwest China region, northwest China region and north China region have all been recorded. This study conducted a comprehensive analysis and summary based on common climatic conditions and the flowering patterns of common flowers.

(1) Climate differences: Different climatic conditions in different regions will affect the flowering period and time.

(2) Temperature differences: In the spring, the temperature differences between the north and the south mainly manifest in temperature changes, weather phenomena and agricultural impacts, etc.

(3) Flowering sequence: Usually, at the same time each year, the temperature and sunlight conditions in the same region are similar, and it can be inferred that the flowering time of different regions will be different.

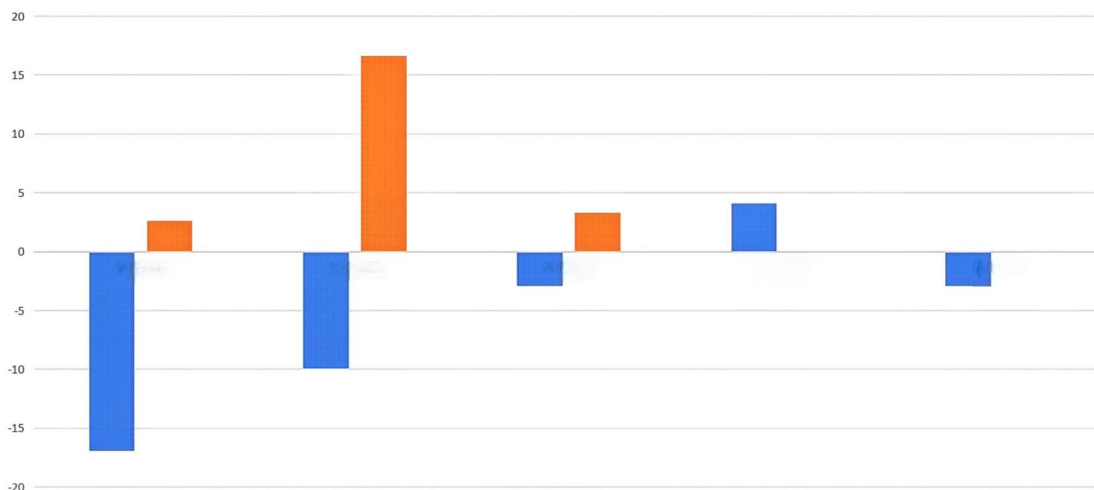


Figure 1. Comparison of Flowering Start Dates in 2025

Figure 1 shows the bar chart of flower blooming predictions for different regions (taking peach blossoms and rape flowers as examples). This chart uses different colors to represent different types of flowers, and each bar represents a different variable. The variables in the chart have different lengths, representing the different characteristics of the flowers growing in different regions. In general, this bar chart can help us analyze and compare different variables, allowing us to have a comprehensive understanding of the overall information.

## 4. Modeling

### 4.1 Establishment of the Rainfall Determination Model

For the establishment of the determination classification model for rainfall, feature selection is the first step. We need to analyze the connections between the classification data and each continuous data. This rainfall determination model will use the Spearman Rank Correlation Coefficient for feature selection, analyzing the data of different cities and selecting the features that are strongly correlated with the target variable. First, use the code to define the list of cities to be analyzed and set the naming rules for each city's data file. Then, process the data of each city one by one through a loop.

In the data preprocessing stage, the first task is to convert the time series in the dataset to a standard date-time format and further extract key time information such as year, month, day, and hour. Then, select several key feature columns for subsequent analysis. These features include, but are not limited to, temperature, pressure, and humidity, meteorological indicators. For the selected meteorological variables (temperature, pressure, humidity), sort their observations in ascending order and assign grade identifiers (1, 2, ..., n). If there are duplicate values in the observations, calculate the average of their grades. Finally, compare the features using the Spearman Rank Correlation Coefficient, which represents whether there is a certain phenomenon.

Through the analysis of each feature using the Spearman Rank Correlation Coefficient, the Spearman Rank Correlation Coefficient is a non-parametric statistical method used to measure the monotonic relationship (whether linear or not) between two variables. It is calculated based on the ranks (ordering) of the data rather than the original values and is suitable for continuous or discrete ordinal variables.

The specific formula is:

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

Among them,  $d_i$  is the calculation of the grade difference, and the specific formula is:

$$d_i = \text{rank } X_i - \text{rank } Y_i$$

Among them,  $X_i$  and  $Y_i$  represent the observations of two different categories, and  $n$  is the sample size. Their basic characteristics and value ranges are as follows:

$$-1 \leq r_s \leq 1$$

If  $r_s = 1$ , it indicates a completely positive monotonic correlation (with completely consistent ranks).

If  $r_s = -1$ , it indicates a completely negative monotonic correlation (with completely opposite ranks).

If  $r_s = 0$ , there is no monotonic relationship.

In addition, to assess whether the correlation is significant, for large samples ( $n > 30$ ), a significance test can be conducted using the t-statistic. The formula for calculating the t-value is:

$$t = r_s \sqrt{\frac{n-2}{1-r_s^2}} (\text{DOF} = n-2)$$

Then, the corresponding  $r_s$  index value is obtained through distributed computing. This  $r_s$  index value can be used to determine the significance level of the correlation. If the  $r_s$  index value is lower than the preset threshold, it can be determined that there is a significant correlation between the feature and the target variable. (It is worth noting that when the variable is a strict monotonic function (such as  $Y = e^X$ ),  $r_s = 1$ , but the Pearson coefficient may not be 1.)

## 4.2 Building the Basic Model

By reading the historical meteorological data of the Qingming Festival and converting the time column into a datetime object, and then extracting information such as year, month, day, and hour, data analysis is conducted. During the data preprocessing stage, it is checked whether there are any missing values. If there are missing values, linear interpolation is used to fill them to ensure data integrity and rationality.

In the feature engineering stage, T (temperature), U (humidity), and Ff (wind speed) were selected as feature indicators, and the target variable was whether there was rainfall. Model training used data from 2025 and approximately the previous 20 years. By filtering the data on the Qingming Festival day each year, historical records of each feature indicator on that date were obtained to establish a Gaussian process regression model.

To predict the meteorological characteristics of the Qingming Festival day in 2026, the Holt-Winters seasonal model (also known as the triple exponential smoothing method) was used for time series prediction. For each feature, data from the past few years (from 2025 to approximately the previous 20 years) was used to predict the T (temperature), U (humidity), and Ff (wind speed) values for each hour on the Qingming Festival day in April 5, 2026. The core idea of the Holt-Winters seasonal model (also known as the triple exponential smoothing method) is a popular method for predicting time series data with trends and seasonal variations. It is an extension of the Holt linear trend model, adding seasonal components.

The Holt-Winters model has three variants, as follows in the standard form:

- 1) Additive Model: Applicable when the seasonal variation does not change over time - Formula:  $y(t) = \text{Level} + \text{Trend} + \text{Seasonality} + \text{Error}$
- 2) Multiplicative Model: Applicable when the seasonal variation increases over time - Formula:  $y(t) = (\text{Level} + \text{Trend}) \times \text{Seasonality} + \text{Error}$
- 3) Seasonless Model: Also known as the Holt Linear Trend Model, it only contains the Level and Trend components.

The Holt-Winters model has three smoothing parameters:

$\alpha$  (alpha): Level smoothing parameter ( $0 \leq \alpha \leq 1$ )

$\beta$  (beta): Trend smoothing parameter ( $0 \leq \beta \leq 1$ )

$\gamma$  (gamma): Seasonality smoothing parameter ( $0 \leq \gamma \leq 1$ )

And a seasonal cycle parameter  $s$  (for example, for monthly data  $s = 12$ , for quarterly data  $s = 4$ ).

The mathematical equation of the Holt-Winters seasonal model varies depending on the type of seasonality (additive or multiplicative). The complete mathematical expression is as follows:

### ① Seasonal Model (Additive Seasonality)

This model is applicable when the amplitude of seasonal fluctuations does not change over time.

1) Level Equation:

$$l_t = \alpha \left( \frac{y_t}{s_{t-s}} \right) + (1 - \alpha)(l_{t-1} + b_{t-1})$$

Here,  $l_t$  represents the smoothed level at time  $t$ ,  $y_t$  is the actual observed value at time  $t$ ,  $s_{t-s}$  is the seasonal component of the previous seasonal cycle, and  $\alpha$  is the level smoothing parameter ( $0 \leq \alpha \leq 1$ ).

2) Trend Equation:

$$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$$

Here,  $b_t$  represents the smoothed trend at time  $t$ , and  $\beta$  is the trend smoothing parameter ( $0 \leq \beta \leq 1$ ).

3) Seasonal Equation

$$s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-s}$$

Among them,  $s_t$  represents the seasonal component at time  $t$ ,  $\gamma$  is the seasonal smoothing parameter ( $0 \leq \gamma \leq 1$ ), and  $s$  is the seasonal cycle length (for example, for monthly data,  $s = 12$ ).

4) Prediction equation (forward prediction by  $h$  steps):

$$\hat{y}_{t+h} = l_t + h \cdot b_t + s_{t-s+h}$$

Among them,  $\hat{y}_{t+h}$  represents the predicted value at time  $t+h$ , and  $h$  is the prediction step size.  $s_{t-s+h}$  is the seasonal component corresponding to the seasonal cycle (indicated by the subscript in the modulo operation).

## ② Multiplicative Seasonality Model

It is applicable to situations where the amplitude of seasonal fluctuations increases or decreases along with the trend.

1) Level Equation:

$$l_t = \alpha\left(\frac{y_t}{s_{t-s}}\right) + (1 - \alpha)(l_{t-1} + b_{t-1})$$

2) Trend equation: (The same as the additive model)

$$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$$

3) Seasonal equation:

$$s_t = \gamma\left(\frac{y_t}{l_{t-1} - b_{t-1}}\right) + (1 - \gamma)s_{t-s}$$

4) Prediction equation (h-step ahead prediction):

$$\hat{y}_{t+h} = (l_t + h \cdot b_t) \times s_{t-s+h}$$

key distinction:

**Table 1.** Comparison Table of Holt-Winters Model

Element	Additive Model	Multiplicative Model
Seasonal	Add and subtract seasonal components (changes in absolute values)	Account for seasonal components (proportional changes)
Applicable Scene	Seasonal fluctuations are stable (such as temperature data)	Seasonal fluctuations change along with the trend (such as sales data)

The Holt-Winters model is utilized to transform seasonality into rainfall probability prediction feature indicators, and the Gaussian process regression model is employed for prediction verification to forecast the occurrence probability of rainfall. The Gaussian process regression model defines the prior distribution of the function space (Gaussian process), updates the posterior distribution based on observed data, and thereby predicts the mean and confidence interval of new data points. Its core advantage lies in its flexible function fitting ability and natural probabilistic interpretation, making it suitable for small datasets and scenarios where quantifying prediction uncertainty is necessary. The core formula of the Gaussian process regression model is:

$$f(x) \sim \mathcal{GP}(\mu(x), k(x, x'))$$

Among them, Gaussian processes are an extension of infinite-dimensional Gaussian distributions, which are completely described by the mean function  $\mu(x)$  and the covariance function (kernel function)  $k(x, x')$ . The mean function refers to typically being set as a constant (such as zero mean) or a simple linear function. The covariance function (kernel function) refers to controlling the smoothness, periodicity, and other characteristics of the function (such as RBF kernel, Matern kernel).

After completing the training process, the logistic regression model acquires the ability to extract the regression coefficients for each feature. The sign and magnitude of the regression coefficients are the key factors determining the degree of influence of each feature on the prediction result. Specifically, if the regression coefficient is positive, it indicates that the improvement of this feature will increase the probability of rainfall occurrence; conversely, if the regression coefficient is negative, it means that the improvement of this feature will decrease the probability of rainfall occurrence.

By providing the training data  $X$  and  $y$ , the prediction distribution formula for the new input  $X^*$  can explain the influence of each feature:

$$\begin{aligned} \mu^* &= K(X^*, X)[K(X, X) + \sigma_n^2 I]^{-1}y \\ \Sigma^* &= K(X^*, X^*) - K(X^*, X)[K(X, X) + \sigma_n^2 I]^{-1}K(X, X^*) \end{aligned}$$

$\sigma_n^2$  is the variance of the total data rainfall, and  $K$  is the covariance matrix calculated by the kernel function. Here, the Gaussian process regression model provides prediction uncertainty, and this probability value indicates the possibility of a rainfall event occurring under specific meteorological conditions. With this probability value, the model can accurately predict the probability of rainfall occurrence for each hour.

This figure presents three subgraphs of the rainfall prediction analysis for the Qingming Festival in 2026. The top subgraph reveals the predicted results of rainfall probability, showing that the data is almost entirely concentrated within the 0 value range, suggesting that the probability of rainfall events is extremely low and no significant fluctuations are observed. The middle subgraph provides the predicted hourly rainfall amounts, with the predicted values being close to zero for most of the time. However, within specific hour intervals, the predicted values show significant fluctuations, which may indicate a relative increase in the rainfall probability during these periods. The bottom subgraph provides a detailed display of the rainfall probability, with several hours showing predicted

probability values approaching 1, indicating a higher likelihood of rainfall during these periods. Overall analysis shows that although the overall rainfall probability is low, the chart still indicates the possibility of rainfall events occurring in certain specific hour periods.

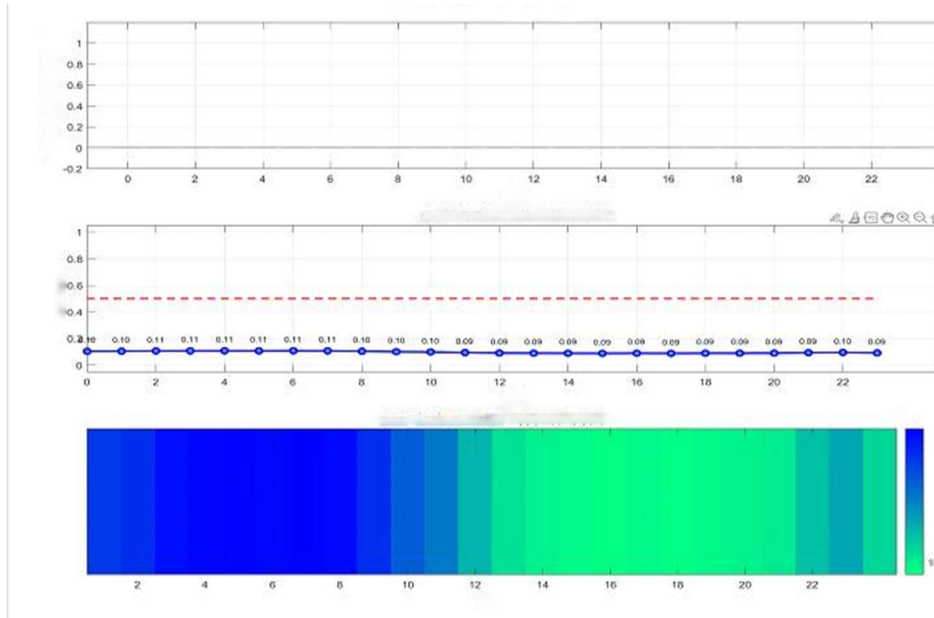


Figure 2. Three subplots of rainfall prediction during the Qingming Festival in 2026

### 5. Flowering Period Prediction Model

The initial flowering period of flowers refers to the starting point when plants begin to display flowers. This time point is usually recorded in the format of "month-day". We have constructed a prediction model for the initial period by analyzing historical data.

This research model takes the annual initial date of flower blooming as the input variable, aiming to predict the starting date of flowering in a specific year (2026).

Data transformation: To facilitate modeling, we convert the initial period data in the "month-day" format into the number of days in a year. If the initial date of the start period in a certain year is "MM.DD", we convert it into the Day Of Year (DoY) of that year, and the calculation formula is as follows:

$$\text{DayOfYear} = \text{Days of the year until MM.DD}$$

This enables us to express the dates as a continuous sequence of numbers, thereby facilitating the smooth conduct of time series regression analysis. Initial prediction model: We use historical "initial" data to train the regression model to predict the initial values for future years. Assuming the training data is  $(y_1, x_1), (y_2, x_2), \dots, (y_n, x_n)$  where  $y_i$  represents the year and  $x_i$  represents the corresponding initial value (converted into days). By fitting these data using the regression model, we obtain the fitting function  $f(x)$ :

$$f(\hat{y}) = f(x)$$

In this study, variable  $x$  is defined as the number of days until the starting time point, denoted by  $y$ , specifically for a specific future year (2026), and  $f$  represents the regression model.

During the regression process, the loss function used is the mean absolute error (MSE):

$$L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Among them,  $n$  represents the sample size, and  $y_i$  is the true value (Ground Truth) of the ( $i$ )th sample, while  $\hat{y}_i$  is the predicted value (Model Prediction) of the ( $i$ )th sample.  $|\cdot|$  denotes the absolute value operation to ensure that the error is a positive number. Then, through the reverse conversion formula, it is converted back to the "month-day" format to obtain the final initial prediction:

BeginPredictionMMDD = DayOfYear to MM.DD

The flowering period refers to the duration from the beginning of flowering until the flowers wither, usually expressed in terms of days. The prediction model for the flowering period is similar to that for the start period, but there are differences in the input and output contents. We use the recorded flowering periods in historical data (i.e., the number of days of the flowering period each year) to build a flowering period prediction model.

Flowering period data: The flowering period data includes the year and the duration of the flowering period in that year. Assuming the training data is  $(y_1, d_1), (y_2, d_2), \dots, (y_n, d_n)$ , where  $d_i$  represents the number of days of the flowering period for the flowers in each year, and  $y_i$  represents the year. We also use a regression model to fit the flowering period data and construct a flowering period prediction model  $g(y)$ :

$$\hat{d} = g(y)$$

Among them,  $\hat{d}$  represents the predicted number of flowering days,  $y$  represents the future year (for example, 2026), and  $g$  represents the regression model used. The objective of the regression model is to minimize the error between the predicted value and the actual value. Huber Loss can be used as the loss function:

$$L(y, \hat{y}) \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{if } |y - \hat{y}| \leq \delta \\ \delta|y - \hat{y}| - \frac{1}{2}\delta^2 & \text{otherwise} \end{cases}$$

The final predicted flowering period value  $\hat{d}$  represents the number of flowering days in 2026. For the prediction model, specifically, we use historical data to predict the start and duration of the flowering period of the flowers. We employed multiple models (including linear regression, quadratic polynomial regression, moving average, exponential smoothing, overall average, and the average of the last three years) for data prediction, and selected the model with the best performance from among them.

The figure shows that these charts are often used to verify the model's effectiveness, visually present the fitting data situation, assist in testing the model's assumptions and identify prediction deviations. At the same time, the band-shaped areas in the graphs may correspond to statistical probability intervals, which are used to quantify the credibility of the prediction results.

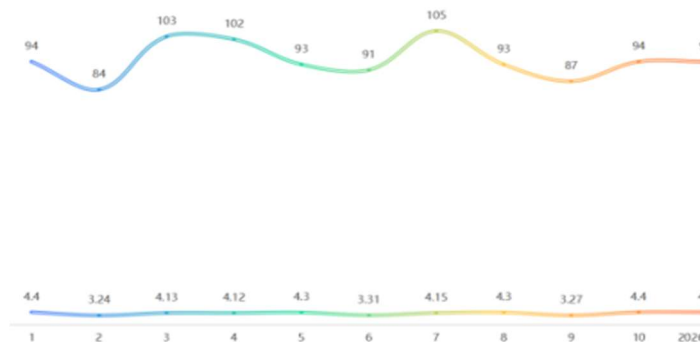


Figure 3. Prediction Chart for Peach Blossom Flowering Date

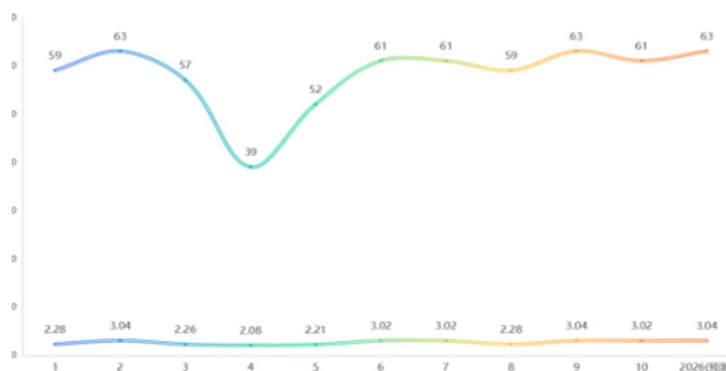


Figure 4. Prediction Chart for the Flowering Date of Rapeseed Flowers

## 6. Optimization Model

### 6.1 Model Construction

#### 6.1.1 Calculation of Overlap between Holiday Time and Flower Bloom Time

We know the starting and flowering periods of different flower varieties in different regions. Now, we need to calculate the overlapping time between the holiday period and the flower bloom time.

- (1) Assume that the start date of the Qingming holiday is April 4th and the end date is April 6th.
- (2) For the starting period and flowering period of different flower varieties, we will extract relevant data. Further calculate the overlapping time. The calculation process of overlapping time is as follows:
  - ① Assume that on the first day of the Qingming holiday, the flowers have already bloomed, and it is counted as the longest visiting time = [flowering start time, holiday end time]
  - ② Assume that on the last day of the Qingming holiday, the flowers are blooming, and it is counted as the shortest visiting time = [flowering end time, holiday end time]

#### 6.1.2 Calculation of Visiting Time:

- (1) Longest visiting time. Assume that the starting visiting time is April 4th and the ending visiting time is April 6th,  $T_{ij} = (6 - 4 + 1) * 24 = 72$  hours.

The actual visiting time (ending visiting time - starting visiting time + 1) \* 24 = actual visiting time.

- (2) Flowering time in each city. Some reference information is shown in Table 2 (taking apricot blossoms as an example)

**Table 2.** Flowering Time of Apricot Blossoms in 2025

Flowers	Regions	a particular year	Initial stage	Flowering period (rounded to two decimal places)
Apricot blossoms	East China Region	2025	3.89	6.43
	Central China Region		6.04	6.22
	Northwest China Region		4.1	84
	Southwest China Region		6.11	14.89
	North China Region		6.11	9.45

According to Table 2, it can be known that the blooming time and flowering period of different varieties of flowers in each region, and the blooming time can be calculated.

6.1.3 Calculation of Distance and Costs across Cities



**Figure 5.** Map of Latitude, Longitude and Altitude Information for Each City

According to Figure 5, the latitude and longitude as well as the altitude information of each city can be obtained, from which the distances between the cities can be calculated. A route map for the best flower-viewing route can be drawn.

6.1.4 The Strategy Also Takes into Account the Impact of Rain on the Trip.

The rainfall situation during the holiday is obtained by reading the rainfall status data. The total duration of the holiday is 72 hours (i.e., 3 days), and the program will calculate the number of rainfall hours for each city during the holiday:

$$t_r = \sum_{t=1}^{72} \text{Rainy condition } (t)$$

In this study, the rainfall status (t) is defined as a binary variable indicating whether rainfall occurred during a specific period, where 1 represents the occurrence of rainfall and 0 indicates no rainfall event. Then, the percentage of rainfall is calculated:

$$P_r = \frac{\text{Number of hours of rainfall}}{72} \times 100\%$$

In the field of Geographic Information System (GIS), the precise calculation of distances between cities is of great significance for spatial analysis. This study employed the Haversine Formula for the calculation of spherical distances. This formula is based on the spherical coordinate system and can accurately calculate the shortest distance between any two points on the Earth's surface. By inputting the geographical coordinates (longitude and latitude) of the two points, the Haversine Formula can provide an accurate spherical distance value. The specific calculation formula is as follows:

$$S_{\text{activity}} = 2 \times R \times \arcsin \left( \sqrt{\sin^2 \left( \frac{\Delta \phi}{2} \right) + \cos(\phi_1) \cdot \cos(\phi_2) \cdot \sin^2 \left( \frac{\Delta \lambda}{2} \right)} \right)$$

Among them:

- The radius of Earth is approximately 6371 kilometers.
- $\phi_1, \phi_2$  represent the latitudes of two points.
- $\lambda_1, \lambda_2$  represent the longitudes of two points.
- $\Delta \phi$  and  $\Delta \lambda$  represent the differences in latitude and longitude between the two points.

#### 6.1.5 Scoring Constraints

(1) In the quantitative assessment of the ornamental value of urban flowers, the first step is to clearly identify the types of flowers that are open during the holiday period. Then, based on the overlapping period of the flower opening period and the holiday, calculate the duration of flower viewing. The scoring also takes into account the flowering quality of the flowers. If the middle stage of the flowering period is taken as the best viewing time, the closer the time is to this best viewing time, the higher the score obtained. The specific scoring is expressed by the following formula:

$$G_{\text{activity}} = \text{Effective viewing duration} \times \left( 1 - \frac{\text{The number of days until the best viewing period arrives}}{\text{Duration of flowering period} / 2} \right)$$

If there are multiple flower varieties in bloom during the holiday period, the scores of each plant will be added to the total score of the city. Subsequently, the city score will be adjusted according to the rainfall conditions during the holiday. The score of the rainy city will be affected by specific factors, and the specific degree of the impact can be calculated according to the following formula:

$$G_{\text{change}} = \text{score} \times (1 - \text{Percentage of rainfall} \times 0.8)$$

The objective function can be expressed as:

$$\text{Maximize: } Z = \sum_{i=1}^N (\text{Flower Score}_i \times \text{Time Factor}_i)$$

Among them:

Flower Score  $i$  represents the flower viewing score of the  $i$ -th city; Time Factor  $i$  is an adjustment factor designed to take into account the travel time and visiting time of the city. The goal is to maximize the total score of flower viewing while optimizing the ratio between the score and the time consumed.

## (2) Decision Variables

The decision variables include the selection of travel routes and the duration of stay in each city. They are defined as follows:

- $x_{ij} \in \{0,1\}$ , indicating whether to travel from city  $i$  to city  $j$

$$x_{ij} = \begin{cases} 1, & \text{If one travels from city } i \text{ to city } j \\ 0, & \text{If one does not travel from city } i \text{ to city } j \end{cases}$$

$t_i \geq \text{min\_visit\_time}$  indicates the total duration of stay in city  $i$ . This duration includes the time spent visiting the flowers as well as the total time required for traveling to and from the station.

$y_i \in \{0,1\}$ , representing whether city  $i$  is visited:

$$y_i = \begin{cases} 1, & \text{If you visit city } i \\ 0, & \text{If one does not visit city } i \end{cases}$$

$-R_i$  assigns a score to the flowers of the  $i$ -th city, which is based on the overlapping period of the blooming seasons of various flowers and the adjusted comprehensive calculation result considering weather factors.

## (3) Other Constraints

- Each city can only be visited once:

$$y_i \in \{0,1\} \forall i$$

- Each city must abide by the minimum stay time regulations:

$$t_i \geq \text{min\_visit\_time} \forall i$$

Make sure to complete the trip within the specified time frame, with the duration not exceeding 72 hours.

$$\sum_{i=1}^N \sum_{j=1}^N t_{ij} x_{ij} + \sum_{i=1}^N t_i y_i \leq T_{\text{total}}$$

### 6.1.6 Complete Optimization Model

By combining the aforementioned goals and constraints, this problem can be transformed into a more understandable form.

Objective function:

$$\text{Maximize: } Z = \sum_{i=1}^N (\text{Flower Score}_i \times \text{Time Factor}_i)$$

Constraints:

1) Total time constraint:

$$\sum_{i=1}^N \sum_{j=1}^N t_{ij}x_{ij} + \sum_{i=1}^N t_i y_i \leq 72$$

2) Urban visit restrictions:

$$\sum_{j=1}^N x_{ij} = y_i \forall i$$

3) Formula for calculating travel time:

$$t_{ij} = \frac{D_{ij}}{V_{\text{train}}} + T_{\text{boarding}} \times 2 \forall i, j$$

4) Flower scoring constraints:

$$R_i = f(\text{flower scores, overlap days, weather factor}) \forall i$$

5) Each city is allowed to be visited only once:

$$y_i \in \{0,1\} \forall i$$

6) The duration of stay in each city must be no less than the minimum required duration as stipulated:

$$t_i \geq \text{min\_visit\_time} \forall i$$

7) The duration of the trip shall not exceed 72 hours:

$$\sum_{i=1}^N \sum_{j=1}^N t_{ij}x_{ij} + \sum_{i=1}^N t_i y_i \leq 72$$

The solution of this optimization model can be achieved through various methods such as linear programming, integer programming, or heuristic algorithms.

## 6.2 Solution Process

The artificial bee colony algorithm (ABC) is a swarm intelligence optimization algorithm based on the foraging behavior of bees, proposed by Karaboga in 2005, and is often used to solve continuous optimization problems (such as function optimization, parameter tuning).

The core idea of the artificial bee colony algorithm is to simulate the division of labor and cooperation mechanism of a bee colony when searching for nectar sources (i.e., the optimal solution). The specific steps are as follows:

## 6.3 Calculation of Travel Scores

For each city, the score of its flower viewing value is based on the comprehensive assessment of the following factors:

- (1) The overlapping days of flower blooming during the holiday.
- (2) The best viewing condition (the closer to the midpoint of the flowering period, the better the viewing effect).
- (3) Weather influence (adjust the score according to the percentage of rainfall; the more rainfall, the lower the score).

The total score of the city will directly affect the priority of the greedy algorithm in selecting the city. The city with a higher score should be given priority for visit.

#### **6.4 Calculation of Travel Time**

The travel time (by high-speed rail) from one city to another is fixed. By calculating the distance matrix and the speed of high-speed rail, the travel time between each pair of cities can be accurately determined. For each city, the city with the shorter travel time should be selected to optimize the itinerary.

#### **6.5 Initialization**

Initialize the score values of all cities to zero. Set a total time limit, for example, a 72-hour holiday duration. Initialize the current set of visited cities. Select the city with the highest score as the starting point of the journey.

Select the next city

Based on the score of the current city, search for the next city with the highest score and calculate the travel time from the current city to the target city.

If the time resources permit, one can continue to explore the next city; otherwise, the selection process should be terminated.

If a city is selected, add it to the list of visited cities and reduce the remaining time (covering travel and visiting time).

#### **6.6 Update Total Score**

Update the total score after visiting a city and adjust it according to the weather conditions of that city. If the city has rainfall during the holiday, reduce the score accordingly based on the percentage of rainfall time.

#### **6.7 Termination Conditions**

- If there are no more accessible cities or all remaining time has been exhausted, terminate.

### **7. Report on the Extension of the Flower Appreciation Economy Chain**

#### **7.1 Background and Problem Description**

The arrival of spring brings a scene of blooming flowers. People's demand for flower appreciation is increasing day by day. The flower appreciation economy is a multi-industry composite economic form centered on flower viewing, integrating agricultural planting, tourism services, cultural creativity, and leisure experiences. The extension of the flower season refers to extending the industrial chain from single flower cultivation to upstream and downstream, forming a "planting - processing - service - derivative" full industrial chain value system.

#### **7.2 Measures to Extend the Flower Season**

##### **1) Apricot Blossoms**

The characteristics of apricot blossoms are that they bloom in early spring, and the single flower's flowering period is short, which is easily affected by late spring cold.

In response to this flower characteristic, the following measures to extend the flowering period are proposed:

- ①. When apricot blossoms are in bloom, cover them with plastic sheds with cold-proof fabric to avoid damage to apricot blossoms caused by low temperatures.<sup>[6]</sup>
- ②. One month before the apricot blossoms bloom, keep the soil moist to avoid drought-induced flower bud shedding.<sup>[7]</sup>
- ③. After the blossoms, appropriate chemical spraying can be done on apricot blossoms to extend the flowering period.

## 2) Rapeseed Flowers

The characteristics of rapeseed flowers are that they bloom in spring, and the group's flowering period is 20-30 days, suitable for planting in cool climates.

In response to this flower characteristic, the following measures to extend the flowering period are proposed:

- ①. Choose to plant in high-altitude areas, which can naturally extend the flowering period of rapeseed flowers.<sup>[3]</sup>
- ②. During the bolting period, appropriate mepiquat chloride can be sprayed to prevent the collapse of rapeseed flowers.<sup>[4]</sup>
- ③. Maintain the field humidity at 60%-70%, to prevent rapeseed flowers from dehydration and death due to drought.<sup>[5]</sup>

## 3) Rhododendrons

The characteristics of rhododendrons are that they bloom in spring, with a single flower's flowering period of 10-15 days, suitable for growing in acidic soil.

In response to this flower characteristic, the following measures to extend the flowering period are proposed:

- ①. During the flowering period of rhododendrons, appropriate shading nets can be added to shade and avoid burning of petals.
- ②. Appropriate ferrous sulfate solutions can be sprayed to adjust the soil pH value to achieve the most suitable soil for rhododendron growth.
- ③. During the flowering period, residual flowers can be built regularly to reduce nutrient loss.

## 4) Cherry Blossoms

The characteristics of cherry blossoms are that they bloom in spring, with a single flower's flowering period of 10-15 days, and are sensitive to temperature.<sup>[2]</sup>

In response to this flower characteristic, the following measures to extend the flowering period are proposed:

- ①. In early spring, spray cold water on the tree crown to delay the flowering speed.
- ②. During the flowering period, appropriate balanced fertilizers of nitrogen, phosphorus, and potassium can be applied to extend the flowering period.

## 5) Peonies

The characteristics of peonies are that they bloom in spring, with a single flower's flowering period of 7-10 days, and are not suitable for growing in hot and humid conditions.<sup>[1]</sup>

In response to this flower characteristic, the following measures to extend the flowering period are proposed:

- ①. During the flowering period, set up a rain shelter to prevent the flower edges from getting waterlogged and rotting.
- ②. When the temperature during the flowering period is higher than 25 degrees Celsius, add reflective boards to lower the temperature.

### 7.3 Conclusion and Suggestions

- (1) Strengthen publicity efforts: The government can produce different publicity signs for different varieties of flowers to attract tourists, thereby extending the economic industrial chain.
- (2) Provide scene innovation: The government can use different varieties of flowers to create a unique light show, attracting tourists to come and watch, or use flowers to provide wedding photography services to extend the economic "industrial chain".
- (3) Produce cultural and creative products: Local governments can jointly create a variety of cultural and creative products themed on unique flowers with local attractions; for example: refrigerator stickers, creative backpacks, etc.

Through these measures, the government can not only extend the flowering period but also extend the economic "industrial chain" to achieve economic growth.

## 8. Model Summary

### 8.1 Model Advantages

- (1) Comprehensive: It takes into account factors such as the blooming time of flowers, regional rainfall predictions, economic benefits, and the time required for activities, providing a comprehensive prediction of rainfall and flower viewing.
- (2) Efficient itinerary planning: By considering the duration of the flower season, prioritizing flower viewing, and designing a minimal overlap path map, the itinerary becomes efficient and time-saving.
- (3) Precise flower season prediction: It combines data such as temperature and altitude in different regions to predict the optimal flower season, avoiding missing the best viewing period or encountering flower withering.
- (4) Visual output: The model combines charts and provides theoretical support, making the model output more three-dimensional and clear.

### 8.2 Model Disadvantages

- (1) The model has limitations: Data updates are not timely. Even if the peak time of visitor flow is missed, due to the large number of people using the same model, congestion may occur.
- (2) Low fault tolerance rate: The data model's prediction relies on historical data. It may encounter unexpected situations, leading to inaccurate conclusions.
- (3) Delayed data update: Obtaining data on the flow of people in remote areas is relatively difficult, which may affect the accuracy of the data.
- (4) Unable to handle complex situations: Most of the factors in the model are linearly connected. In the face of complex and variable situations, they may be connected by non-linear factors. However, due to the model's issues, it is unable to handle complex situations.

## Acknowledgments

Project number: xjclx202504 1025145 Textbook Project for 2025 - System Integration Design of Industrial Robots.

## References

- [1] Yue Gaofeng, Wang Liping, Liu Li, et al. Research on Mudan Flowering Prediction Based on Meteorological Conditions [J]. Journal of Jiangxi Agricultural Sciences, 2020, 32(02): 105-109. DOI: 10.19386/j.cnki.jxnyxb.2020.02.19.
- [2] Chen Daopo. Correlation Analysis and Forecast of Cherry Blossom Bloom Period with Meteorological Factors [J]. Horticulture and Seedling, 2023, 43(10): 50-52. DOI: 10.16530/j.cnki.cn21-1574/s.2023.10.019.

- [3] Li Chunhui, Zhang Xiaofang, Cai Zhe, et al. Improvement of Wuyuan Rapeseed Flowering Period Forecast Model Based on Effective Accumulated Temperature Method [J]. Chinese Agricultural Meteorology, 2024, 45(03): 281-292.
- [4] Gao Ping, Xu Min, Kong Weicai, et al. Research on the Forecast Model of Rapeseed Flowering Period Based on Optimal Correlation Analysis [J]. Journal of Marine Meteorology, 2021, 41(03): 77-83. DOI: 10.19513/j.cnki.issn2096-3599.2021.03.008.
- [5] Feng Minyu, Kong Ping, Hu Ping, et al. Establishment of Rapeseed Flowering Period Forecast Model Based on Pre-flowering Phenology and Grey Correlation Analysis Method [J]. Chinese Agricultural Meteorology, 2021, 42(11): 929-938.
- [6] Yang Yanling, Pan Chenliang, Chen Yanli, et al. Research on the Variation Characteristics and Forecast Method of Hami Oasis Peach Flowering Period [J]. Chinese Agricultural Science Bulletin, 2023, 39(14): 105-109.
- [7] Song Xiaoxin, Wu Mitai, Zhang Guodong, et al. Research on Meteorological Forecast of Rapeseed Flowering Period in Xinyuan County [J]. Desert and Oasis Meteorology, 2021, 15(06): 103-107.