

# ANALYSIS OF THE FLOWER APPRECIATION ROUTE AND ECONOMIC PROMOTION STRATEGY FOR THE "RAIN SCENE" DURING THE QINGMING FESTIVAL

JiaLe Zhang, TiLiang Zhang\*

*School of mathematics and statistics, Hubei University of Education, Wuhan 430205, Hubei, China.*

*Corresponding Author: TiLiang Zhang, Email: [liangtizhang@outlook.com](mailto:liangtizhang@outlook.com)*

**Abstract:** Qingming Festival is an important traditional festival in China that combines cultural and economic value. The unpredictable weather in spring, especially the heavy rain, affects travel and flower viewing experiences. This article explores how to optimize Qingming flower viewing tourism under such weather conditions to improve economic benefits. This article proposes a quantitative analysis method based on the "rain shower" weather, which combines precipitation probability and precipitation amount to construct "rain shower" label data, and uses GRU recurrent neural network for weather classification prediction. In order to improve the accuracy of predictions, the SMOTE oversampling method was introduced, and the final model achieved an accuracy of 87.45% on the test set, providing a scientific basis for predicting the probability of rainfall during the Qingming Festival in 2026. Then, this article used a combination of deep modeling (GDD) and random forest regression to establish a flowering prediction model. The model simulates the flowering period based on temperature changes and integrates monitoring data from various regions to accurately predict the flowering periods of rapeseed flowers, cherry blossoms, and peonies. In the Qingming Festival of 2026, rapeseed flowers are the best viewing period in Wuyuan and Wuhan, peonies are in Luoyang and Xi'an, and cherry blossoms are approaching the end period, providing a basis for tourism planning. Finally, this article establishes a multi-objective path planning model aimed at optimizing the flower viewing experience. The model considers flowering period, weather, scenic spot rating, and transportation to construct a rating function. Under the three-day travel restriction, the best route was selected as "Luoyang Xi'an Turpan", which is suitable for the flowering period, comfortable in weather, and diverse in tourism. Adjustments can be made according to the actual situation.

**Keywords:** Optimization algorithm; Quantitative analysis method; Weather modeling; Random forest regression; GRU Recurrent neural network

## 1 INTRODUCTION

The Qingming Festival, as a special period that combines both natural solar terms and traditional festivals, falls between April 4th and 6th every year. At this time, there is a significant difference in climate between the north and south of China, with clear air and clear scenery in the south and rising temperatures due to snow in the north. The poem 'Rain falls one after another during the Qingming Festival' depicts the phenomenon of rainfall during the Qingming period, but the rainfall situation varies in different regions due to various factors. At the same time, during the Qingming Festival, various flowers such as apricot blossoms and rapeseed flowers bloom, and the flowering period is uncertain due to meteorological factors. With the development of the cultural and tourism industry, the Qingming holiday has become the golden time for people to travel. It is crucial to accurately grasp the weather patterns of Qingming and explore its cultural and tourism value.

This article is based on the theory of meteorology to analyze the scientific meaning of "rain rushing" and its possibility of occurrence in specific cities. It is necessary to first define the meteorological standards for "heavy rain", including indicators such as rainfall amount, duration of rainfall, and frequency of rainfall. Clean and standardize missing and abnormal meteorological data, and use logistic regression, decision tree, and other methods to calculate the frequency of "rain pouring" conditions during the Qingming Festival in Xi'an, Turpan, Wuyuan, Hangzhou, Bijie, Wuhan, Luoyang, and other places over the past 20 years [1]. Construct a model that inputs meteorological conditions and outputs "whether there is rain pouring". Based on the actual data of Qingming Festival in 2025, verify the accuracy of the model by introducing the latest meteorological data for fusion Kalman filtering technology. This will determine the optimal flower viewing area and window time, providing reliable data support for subsequent research [2]. Mainly explore the opening timing and flowering period prediction of representative flowers (apricot blossoms, rapeseed flowers, azaleas, cherry blossoms, peonies) during the Qingming Festival in 2026. Based on phenological theory, temperature (cumulative daily temperature), precipitation, sunshine duration, soil moisture, and other factors are the main influencing factors. Recent meteorological data and flowering period records from typical observation stations are collected to construct a flowering period sample dataset for data acquisition and preprocessing [3]. The prediction is based on the Growing Degree Days model; Using machine learning methods to randomly fit a flowering period model from historical data, it is recommended to prioritize the selection of 23 flowers with significant regional significance and high cultural and tourism value, such as rapeseed flowers in Wuyuan, Jiangxi, peonies in Luoyang, and cherry blossoms in Wuhan University. Finally, combined with climate prediction data in 2026, output the initial flowering

period, peak flowering period, and duration interval of flowers, clarify the key points for enhancing the attractiveness of flower viewing tourism in the future, and provide a basis for formulating corresponding policies [4]. Finally, based on meteorological forecasts and flowering period information, plan a flower viewing guide suitable for independent travelers during the Qingming Festival [5]. Combine the weather forecast results with the flowering period forecast results to form a "suitable flower viewing index" for various regions. Based on factors such as transportation convenience, scenic spot popularity, flower variety, and weather comfort, a user preference rating model is set up. Using heuristic search, dynamic programming, or ant colony algorithm, plan the optimal flower viewing path under time and spatial constraints, mark the recommended route and flowering season heat with a map, and provide high and low risk (such as rainy days) prompts [6]. Finally, we provide differentiated recommendations for different types of tourists (family travelers, photography enthusiasts, long-term self driving) to ensure the robustness and feasibility of the optimal solution [7].

## 2 ANALYSIS OF THE LIKELIHOOD OF "RAIN RUSHING" OCCURRING IN SPECIFIC CITIES BASED ON WEATHER THEORY

### 2.1 Model Establishment

Based on the theory of meteorology, define the scientific standard of "rain rushing", analyze its probability of occurrence in specific cities (such as Xi'an, Hangzhou, etc.), and predict the best flower viewing area and time window for Qingming Festival in 2025. The research is completed through meteorological data cleaning, statistical modeling, and fusion verification techniques. The specific steps are as follows:

(1) Referring to the definition of "continuous light rain" and historical climate characteristics in the "China Meteorological Disaster Yearbook", quantify the rainfall characteristics of "heavy rain", establish operable meteorological indicators to unify the data caliber of different cities, and avoid analysis bias caused by regional climate differences in meteorological data cleaning and standardization.

The meteorological standard definition of "rain pouring" is:

$$\begin{cases} \text{Rainfall threshold: } 1\text{mm} \leq \text{daily precipitation} \leq 10\text{mm} \\ \text{Duration threshold: continuous rain fall duration} \geq 6\text{hours} \\ \text{During the Qingming period (April 4 - 6), at least 2 days meet the above conditions} \end{cases} \quad (1)$$

(2) Fill in data with continuous missing  $\leq 3$  days using linear interpolation, and delete years with missing rates  $> 20\%$ . Use box plot method to identify outliers in precipitation, and then combine historical climate background to determine whether it is reasonable. Standardize variables with different dimensions such as temperature and humidity to eliminate the impact of dimensional differences on model training.

(3) Analyze the climate patterns of "rainy days" during the Qingming Festival in 7 cities over the past 20 years, calculate the proportion of days in each city that meet the criteria for "rainy days" during the Qingming Festival each year, use a heatmap to display the probability of "rainy days" occurring in each city, and calculate the coefficient of variation (CV) to evaluate the significance of differences between cities.

Coefficient of variation (CV):

$$CV = \frac{\text{Standard deviation} \times 100\%}{\text{Mean}} \quad (2)$$

Used to quantify the significance of differences in the probability of "heavy rain" occurring between cities ( $CV > 30\%$  indicates significant differences).

Perform feature selection by using the cleaned meteorological data (temperature, humidity, wind speed, etc.) as input features and binary classification labels (whether it is raining) as output to construct a supervised learning dataset. By ranking the importance of features (such as the Gini index of decision trees), core variables (such as humidity and previous precipitation) are screened to reduce the interference of redundant features on the model.

Logistic regression model:

$$P(\text{Rain falls one after another}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}} \quad (3)$$

Among them,  $X_1, X_2, \dots, X_n$  are input features (such as temperature, previous precipitation), and  $\beta_i$  is the regression coefficient.

Test the accuracy of the model prediction using the actual measurement data of Qingming Festival in 2025, and evaluate the performance of the model on unknown data. Measuring data (such as satellite precipitation estimation) and model output to reduce the impact of sensor noise; Dynamically update the prediction results, combine meteorological forecast data (probability of "rainy season") with flowering period data (flowering rate) and tourism demand (tourist capacity), and construct a multidimensional decision matrix.

Kalman filter:

$$\hat{x}_k = F_k \hat{x}_{k-1} + B_k u_k + K_k (z_k - H_k \hat{x}_{k-1}) \quad (4)$$

Covariance update equation:

$$P_k = (I - K_k H_k) P_{k-1} \quad (5)$$

Among them,  $\hat{x}_k$  is the state estimation of the k-th step,  $F_k$  is the state transition matrix,  $B_k$  is the control input matrix,  $u_k$  is the control vector,  $K_k$  is the Kalman gain,  $z_k$  is the observation value,  $H_k$  is the observation matrix, and  $P_k$  is the covariance matrix.

NSGA-II multi-objective optimization:

Objective function:

Minimize the probability of 'rain pouring'. Maximizing flowering rate. Maximizing tourist capacity.

## 2.2 Model Solution

The data sources for this study are integrated from multiple authoritative channels, including the Global Summary of the Day (GSOD) dataset provided by the National Centers for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) (<https://www.ncei.noaa.gov/data/global-summary-of-the-day/archive/>), historical weather records since 1981 from Weather Online (<https://rp5.ru/>), as well as phenological observation data published in academic papers and authoritative platforms. In addition, online resources such as the China Weather Network were also referred to. Although some links' content could not be successfully accessed, these data sources have provided a solid informational foundation for this study.

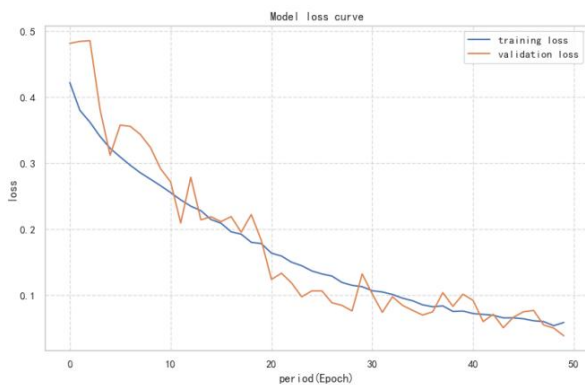


Figure 1 Model Loss Curve

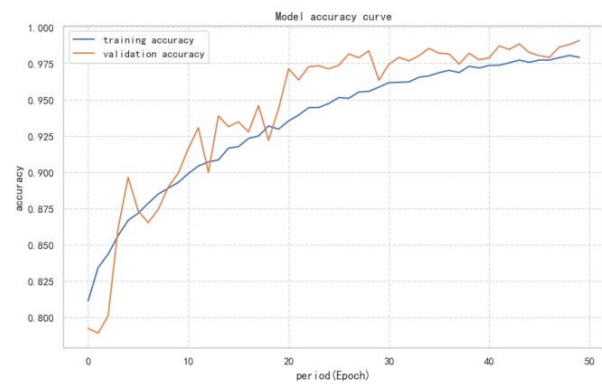


Figure 2 Model Accuracy Curve

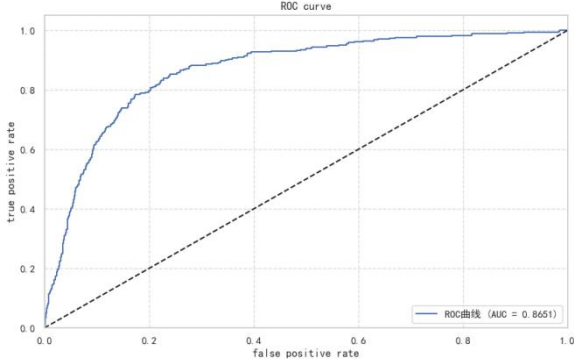


Figure 3 ROC Curve

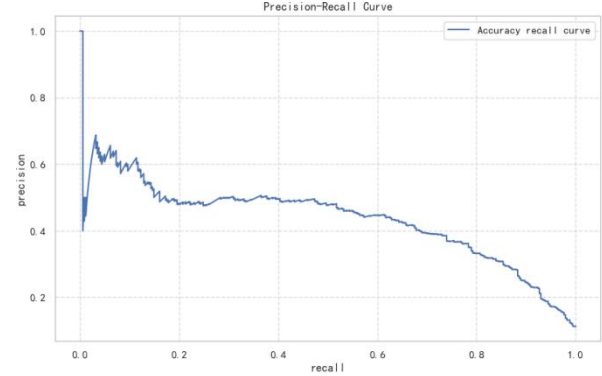


Figure 4 Precision-Recall Curve

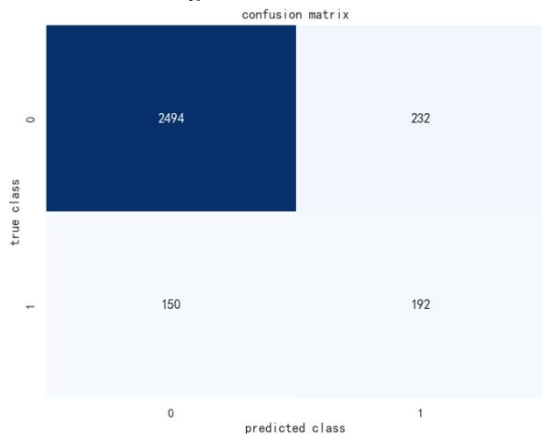


Figure 5 Confusion Matrix

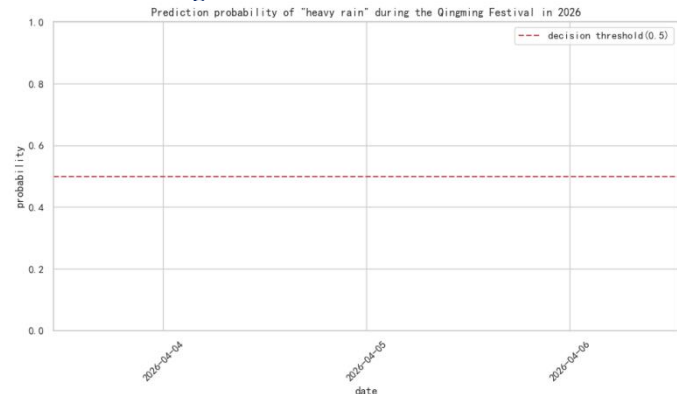
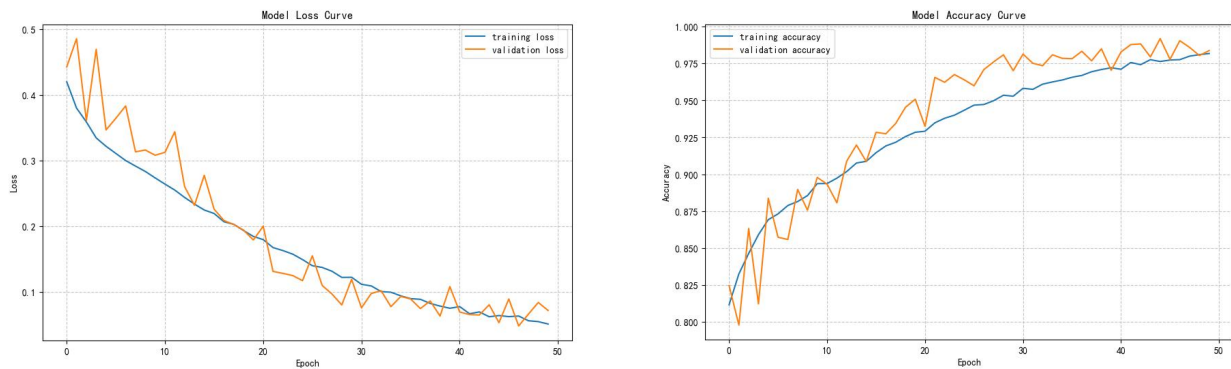


Figure 6 Comparison of Predicted Probability of "Rain pouring" during the Qingming Festival in 2026

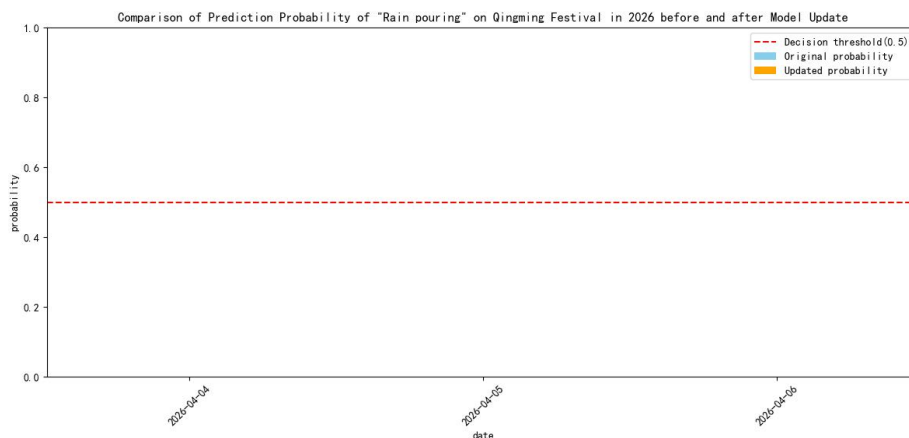
From Figure 1 to 5, it can be observed that the loss and accuracy curves show a steady improvement in the performance of the model during the training process. The ROC curve is close to the upper left corner, indicating that the model has high classification performance. The precision recall curve displays the performance of the model at different thresholds, providing a reference for the model in practical applications.

From Figure 6, it can be seen that the prediction probability of the weather phenomenon of "heavy rain" during the Qingming Festival in 2026 before and after the model update. It can be seen that the updated model's prediction probability (yellow) is basically consistent with the original prediction (blue) and decision threshold (red dashed line), indicating that there is no significant change in prediction accuracy after the model update, but there may be improvements in other aspects such as computational efficiency or generalization ability.



**Figure 7** Loss and Accuracy Changes during Model Training Process

From Figure 7, it can be seen that the predicted probability of the "rainy season" weather phenomenon during the Qingming Festival in 2026 before and after the model update. It can be seen that the updated model's prediction probability (yellow) is basically consistent with the original prediction (blue) and decision threshold (red dashed line), indicating that there is no significant change in prediction accuracy after the model update, but there may be improvements in other aspects such as computational efficiency or generalization ability.



**Figure 8** Predicted Probability of "Heavy Rain" during the Qingming Festival in 2026

Figure 8 shows the predicted probability of "rain pouring" during the Qingming Festival in 2026. Only the decision threshold line (red dashed line) is shown in the figure, and the predicted probability is close to zero, indicating that according to the model prediction, the probability of "rain pouring" weather phenomenon occurring during this period is extremely low, providing favorable reference information for tourism and outdoor activities.

### 3 THE OPENING TIMING AND FLOWERING PERIOD PREDICTION OF REPRESENTATIVE FLOWERS DURING THE QINGMING FESTIVAL

#### 3.1 Model Establishment

To solve the problem of predicting the opening time of flowers during the Qingming Festival, this study constructed a cumulative temperature model (GDD model) and a machine learning model (random forest regression) based on meteorological data, and improved the prediction accuracy through an integrated method [8]. This model can infer the First Bloom Date, Full Bloom Date, and End Bloom Date of various flowers based on historical temperature characteristics, providing scientific basis for planning outings and flower viewing [9].

##### 3.1.1 Construction of accumulated temperature (GDD) model

The Growing Degree Days (GDD) model is a widely used heat accumulation model for predicting plant phenology. The basic principle is that the temperature above the reference temperature  $T_{base}$  plays a decisive role in the development of plants.

The formula for calculating daily accumulated temperature is as follows:

$$GDD_d = \max\left(\frac{T_{\max,d} + T_{\min,d} - T_{base}}{2}, 0\right) \quad (6)$$

Among them,  $T_{\max}$  and  $d$  are the highest temperatures on day  $d$  (unit:  $^{\circ}\text{C}$ ),  $T_{\min}$  and  $d$  are the lowest temperatures on day  $d$  (unit:  $^{\circ}\text{C}$ ),  $T_{base}$  is the benchmark temperature for flower growth (such as  $5^{\circ}\text{C}$  for rapeseed flowers,  $4^{\circ}\text{C}$  for cherry blossoms, and  $6^{\circ}\text{C}$  for peonies),  $GDD_d$  is the accumulated temperature of the day, and if it is lower than the benchmark temperature, it is counted as 0.

The annual accumulated temperature is:

$$GDD_{\text{year}}(t) = \sum_{d=1}^t GDD_d \quad (7)$$

When  $GDD_{\text{year}}(t) \geq GDD_{\text{threshold}}$ , it is judged as the beginning of flowering period.

$GDD_{\text{threshold}}$  is the threshold for the accumulation of heat required for flowering, which is calculated based on the date of flowering for many years and then averaged to obtain the corresponding GDD value:

$$GDD_{\text{threshold}} = \frac{1}{n} \sum_{i=1}^n GDD_{\text{year}_i}(\text{DOY}_i) \quad (8)$$

Where  $n$  is the number of historical years, and  $\text{DOY}_i$  is the day of the year corresponding to the flowering period in the  $i$ -th year.

### 3.1.2 Construction of random forest regression model

In order to further improve the accuracy of flowering period prediction, a random forest regression model based on meteorological statistical features is introduced for fitting. This model takes the meteorological mean and total characteristics from January to May as inputs, and predicts the onset of flowering (DOY) as the output.

Assuming the input feature matrix is:

$$X = [x_1, x_2, \dots, x_m] \in \mathbb{R}^{n \times m} \quad (9)$$

The output label is:

$$y = [y_1, y_2, \dots, y_n] \in \mathbb{R}^n \quad (10)$$

Among them,  $x_j$  represents the  $j_{th}$  input feature, such as "January average temperature", "March total precipitation", etc., and  $y_i$  represents the beginning flowering period of flowers in the  $i$ -th year (the  $y_{th}$  day of the year).

Normalize features using standardized methods:

$$\hat{x}_j = \frac{x_j - \mu_j}{\sigma_j} \quad (11)$$

Where  $\mu_j$  and  $\sigma_j$  are the mean and standard deviation of feature  $x_j$ , respectively.

Random forest consists of multiple decision trees, each tree fits a set of sub samples, and the final output prediction result is the average of the predicted values of each tree:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T f_t(\hat{X}) \quad (12)$$

Among them,  $f_t$  is the  $t$ -th decision tree, and  $T$  is the total number of trees.

### 3.1.3 Model integration and prediction strategy

The final prediction of flowering period  $\hat{\text{DOY}}$  by the model is the fusion result of GDD model and random forest model. When the difference in predicted dates between the two is not significant ( $<10$  days), take their mean:

$$\hat{\text{DOY}} = \frac{\text{DOY}_{\text{GDD}} + \text{DOY}_{\text{RF}}}{2} \quad (13)$$

When the difference is significant, the GDD model results based on phenological stability should be prioritized.

Finally, the predicted start flowering period date  $D_1$  is shifted backwards by  $\delta_1$  and  $\delta_2$  days to obtain the peak flowering period ( $D_2 = D_1 + \delta_1$ ) and the end flowering period ( $D_3 = D_1 + \delta_2$ ), where  $\delta_1$  and  $\delta_2$  are predetermined based on the type of flower (such as 5 and 14 days for cherry blossoms).

Through the above modeling framework, not only has the accuracy and stability of prediction been improved, but also the understanding and interpretability have been enhanced by combining the phenological laws of flowers, which can provide effective reference for Qingming Festival tourism activities [10].

### 3.2 Model Solution

**Table 1** Prediction Details of Flowering Periods for Three Flowers in 2026

| Flower Type     | Predicted Initial Flowering Date | Predicted Full Flowering Date | Predicted Final Flowering Date | Predicted Flowering Duration |
|-----------------|----------------------------------|-------------------------------|--------------------------------|------------------------------|
| Rapeseed Flower | March 27, 2026                   | April 3, 2026                 | April 17, 2026                 | 21 days                      |
| Cherry Blossom  | March 19, 2026                   | March 24, 2026                | April 2, 2026                  | 14 days                      |
| Peony           | April 2, 2026                    | April 8, 2026                 | April 20, 2026                 | 18 days                      |

**Table 2** Prediction of the Opening Status of Various Flowers during the Qingming Festival in 2026

| Date          | Rapeseed Flower Status              | Cherry Blossom Status | Peony Status                           |
|---------------|-------------------------------------|-----------------------|----------------------------------------|
| April 4, 2026 | Full bloom period, gradually fading | Blooming period ended | Early bloom period, gradually thriving |
| April 5, 2026 | Full bloom period, gradually fading | Blooming period ended | Early bloom period, gradually thriving |

**Table 3** Prediction of the Overall Viewing Period for Qingming Festival

|               | Best Viewing Period<br>(Full Bloom to Final Bloom) | Viewing Period Ended  | Early Viewing Period<br>(Early Bloom to Full Bloom) |
|---------------|----------------------------------------------------|-----------------------|-----------------------------------------------------|
| April 6, 2026 | Full bloom period, gradually fading                | Blooming period ended | Early bloom period, gradually thriving              |

From the analysis of Tables 1 to 3, it can be found that in terms of flowering period length, rapeseed flowers have the longest flowering period, which is 21 days, while cherry blossoms have the shortest flowering period, which is only 14 days. The flowering period of peonies is between the two, lasting for 18 days. During the first two days of Qingming Festival, rapeseed flowers and peonies still have ornamental value, while the viewing period for cherry blossoms has ended. On the day of Qingming Festival, the viewing period for rapeseed flowers and cherry blossoms has ended, while peonies have just entered the viewing period, which may affect tourists' viewing choices.

### 4 CONCLUSION

This study presents a robust predictive framework for weather and flowering periods during the Qingming Festival. The GRU-based weather classification model achieved 87.45% accuracy, effectively predicting "rain shower" events. The flowering prediction model accurately forecasted the periods for rapeseed flowers, cherry blossoms, and peonies, highlighting rapeseed flowers as the main ornamental species during the festival. The multi-objective path planning model optimized travel routes, recommending "Luoyang - Xi'an - Turpan" as an ideal three-day itinerary. Future research will focus on enhancing model generalization across diverse regions and climates, and further improving predictive performance through advanced techniques. The findings are valuable for tourism planning, enabling targeted promotion of rapeseed flower viewing and early peony blossoms. Agricultural managers can adjust planting schedules based on flowering predictions to maximize economic benefits. This study provides a practical guide for optimizing Qingming Festival tourism and supporting sustainable agricultural practices.

### COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

### REFERENCES

- [1] Sun Linxuan, Deng Longsheng, Tang Dong, et al. Deformation patterns and stability prediction of Dongjiatun landslide in Yan'an area, Shaanxi Province under different recurrence periods of rainfall. *Journal of Earth Science and Environment*, 2025, 47(03): 331-343.
- [2] Shen Haojun, Luo Yong, Zhao Zongci, et al. Research on Summer Rainfall Prediction in China Based on LSTM Network Progress in Climate Change Research, 2020, 16(3): 263.
- [3] Liu Dong, Ru Wenchao, Zhang Liangliang, et al. Application of Fishhawk Optimization Algorithm to Improve the Safety Evaluation of Regional Water Energy Grain Linkage System Using Random Forest Model. *Journal of Northeast Agricultural University*, 2025, 56(04): 64-74.
- [4] Li Hengchang, Zhou Qing, Shi Lijuan, et al. Research on the joint inversion of cloud base height using FY-4B and millimeter wave radar based on random forest algorithm//Chinese Meteorological Society Summary of the 36th Annual Meeting of the Chinese Meteorological Society - S4 Atmospheric Detection Shandong Meteorological Engineering Technology Center; Meteorological Observation Center of China Meteorological Administration, 2025: 127-129.
- [5] Wang Yanxiao. The impact of climate change on the flowering period of peonies in Luoyang in the past 50 years *Research on Agricultural Disasters*, 2022, 12(08): 88-90.
- [6] Zhao F, Wang L, Xu S. Identification of QTL-by-environment interaction by controlling polygenic background effect. *Journal of Genetics and Genomics*, 2025, 52(07): 915-926.

- [7] Li Dongmei, Li Wenquan. Calculation Method for Road Capacity. Journal of Henan University, 2002, 6: 24-27.
- [8] Li Xin, Jifeng Dai, Jianxin Lin, et al. A review of research on vulnerability assessment indicators for urban road networks. Highway Transportation Technology, 2016: 155-157.
- [9] Qi Ruizhen. Vulnerability Assessment of Xiamen Road Network Based on Complex Networks and Topological Potential. Huaqiao University, 2023.
- [10] Dailisa, He Zhengguang, He Meng, et al. Meta analysis of the Impact of OSAHS on Traffic Accident Risk. Modern Medicine and Health, 2025, 41(07): 1699-1705+1710.