



## DECISION-MAKING SUPPORT PLATFORM AND SECURITY DESIGN FOR RURAL LEISURE TOUR INDUSTRY BASED ON SEA METHOD AND SVR MODEL

YUNZI GU\*

**Abstract.** An accurate short-term passenger flow forecast of rural tourism can avoid accidents as much as possible. However, the short-term passenger flow of rural leisure tourism shows nonlinear, seasonal, random, and other complex characteristics. Meanwhile, the traditional forecasting methods are often difficult to achieve accurate forecasting. Therefore, this study now used the nonlinear mapping function in the support vector regression to convert the passenger traffic training sample into a high-dimensional feature space and established a linear decision function. Then the influence of periodicity on the prediction effect through seasonal index adjustment was reduced. Finally, event triggering combined with core embedding technology for tamper-proof detection was adopted to improve the security of the platform. The results showed that the minimum absolute error of prediction with improved SVM model was 0.27% compared with traditional model and autoregressive integrated moving average model. After the introduction of the Internet search factor, the traffic prediction result was more accurate, which was 0.0425 smaller than that without the introduction of the Internet search factor. When the concurrency was less than 100 times/s, the average response time difference before and after adding the core embedded program was small, indicating that the security of page tampering technology was high. This research method can effectively predict the passenger flow of rural leisure tourism industry and ensure the safety of the platform.

**Key words:** SVR model; SEA method; Leisure tour industry; Page anti-tampering technology; Countryside

**1. Introduction.** In response to the national call, the rural leisure tour industry has gradually developed in various places, operating numerous farmhouses and attracting urban residents to spend leisure and entertainment in rural farmhouses [11]. According to statistics, the rural leisure tourism industry receives nearly 800 million visitors annually, creating considerable revenue for the countryside. However, the rural leisure tourism industry has the characteristics of loose geographical distribution, uneven development across the country, and low information level. These may lead to greater difficulties in the management mode transformation [4]. The forecast of passenger flow plays an important role in the leisure tourism industry. On the one hand, the forecast value can assist managers to make reasonable scheduling of resources and improve industrial benefits [13]. On the other hand, it can assist the early warning work in the peak period of tourism. The security of the platform is the key to the normal operation of the decision support platform of rural leisure tourism industry [10]. The content of the platform has a wide audience. Meanwhile, the decision is based on the final display to the user in the form of a web page [12]. The content of the page will affect the interests of all parties. Meanwhile, the immutable content of the web page will be the most basic requirement for the security of the platform [5]. This study combines the current situation and demand of the industry, integrates local real-time passenger flow information, network public opinion information, and other data, and builds a relevant decision support platform. Users' preferences and needs can be understood through in-depth analysis of users' behavior path, stay time, conversion rate and other data on websites or apps. The user experience and page design are optimized, the budget and financial plan of the scenic spot are scientifically formulated, and the tourist resources of the scenic spot are optimized. In view of the advantages of Support Vector Regression (SVR) in time-series data processing, a short-term passenger flow prediction model is constructed to improve the management level of scenic spots, optimize the allocation of tourism resources by understanding tourists' needs, and ensure tourists' safety. The innovation of this research is as follows: (1) SVR is used as the basic model to build a short-term passenger flow prediction model. Different from the traditional model, the network related search factor is

---

\*Department of Tourism and Management, Wuhan College of Foreign Language and Foreign Affairs, Wuhan, 430000, China (yunzi\_gu@outlook.com)

introduced. Meanwhile, the compensation effect of the prediction model is improved by optimizing the Seasonal Exponential Adjustment (SEA) weight. (2) The short-term passenger flow prediction model is used as the core module of the decision support platform. Meanwhile, data analysis technology is used to model the aggregated data. (3) The event-triggered combination and core embedding technology are used to achieve the tamper-proof performance of the platform to ensure the safe use of the platform. The contributions of this study are as follows: (1) According to the nonlinear characteristics of passenger flow data, the SVR model is introduced. Meanwhile, the influencing factors of SVR are selected through Grey correlation analysis, which overcomes the shortcomings of the classical time-series prediction model. (2) In view of the obvious seasonal characteristics of annual holiday passenger flow, an adaptive SEA-SVR model based on seasonal adjustment is proposed to directly process the seasonality of the original time-series data, shortening the forecasting time and improving the forecasting accuracy. (3) A decision-making support platform is established for the rural leisure tourism industry, the allocation of tourism resources is optimized, and the management level of scenic spots is improved.

**2. Related Works.** A new type of tourism is the countryside leisure tour. It has become a choice for many people to relax on weekends due to its location in the suburbs and proximity. While the rural leisure tour industry is developing rapidly, its rough management mode has hindered its development. Therefore, optimizing the management mode and effectively using computer technology becomes the direction for optimizing the rural leisure tourism industry. Zheng et al. conducted a relevant data survey to understand the focal actor characteristics and identify the core actors among them in the rural tourism actor network translation through interviews and other forms. The results showed that leisure services, etc. affected the actors [15]. Liu conducted a specific study through practical case studies to grasp the development of leisure agriculture. From the analysis results, the utilization of idle home base was not high and there were more deficiencies in rural tourism, which limited the spatial expansion of leisure tourism [9]. Ding et al. conducted a sampling survey to understand the characteristics of Sichuan farmhouse sewage and analyze its physical and chemical characteristics. The results showed that their effluent discharge was increasing year by year, which negatively affected water resources and should be noticed. Chase L et al. conducted a study on the problems faced by agritourism and proposed a conceptual framework that contains different stratification. The framework made agritourism easier to understand and facilitated the development of agritourism [2].

Hasanipanah et al. chose the improved SVR model to predict air overpressure due to it in my blasting. The improved SVR model was found to have better prediction through comparative analysis [1]. The SVR model was used by Zhang et al. to predict the relevant reaction time of the driver based on Electroencephalogram (EEG) to understand the driver's reaction under unexpected events [14]. The results showed that the correlation between different EEG characteristic parameters. Meanwhile, simple reaction time were different, and the prediction accuracy of SVR model was higher compared with other models [3]. Hu et al. constructed a human face grain yield prediction problem based on Grey correlation analysis using SVR model and optimized the model using adaptive boosting algorithm. The results showed that the model could obtain better prediction results [6]. Lin et al. faced the short-term PV power prediction problem and proposed a hybrid model based on the SVR model. The model could make more accurate predictions under ideal weather conditions through experimental analysis [8].

In summary, the existing methods can well solve the single-factor forecasting. However, these methods have high requirements for the accuracy of historical data and problems such as high time and cost, difficult to determine impact factors, etc. It is difficult to achieve complex and nonlinear short-term passenger flow without considering various interference factors affecting tourism demand. Meanwhile, more references and suggestions cannot be brought to tourism decision-makers. Therefore, the paper starts from the computer technology-assisted aspect and constructs a relevant decision support platform to forecast the short-term passenger flow. Since the SVR model has certain advantages in time-series data processing, the model is used as a prediction model for short-term passenger flow analysis.

### 3. Rural leisure and tourism industry decision-making support platform and safety design.

**3.1. Construction of industry short-term passenger flow forecast model.** In recent years, the leisure tour industry in China's countryside develops rapidly, and the form of the industry is undergoing

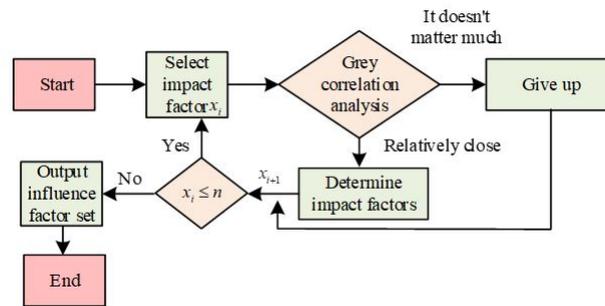


Fig. 3.1: Grey correlation analysis process

major changes. Due to factors such as loose geographical distribution and low Informationization, it has added difficulties to the management of the rural leisure tour industry. Meanwhile, it is difficult to have a comprehensive understanding of the leisure tour development in each region. The forecast of tourist flow mainly includes the traditional time-series method, econometric model, and the current artificial intelligence forecasting methods. However, it is difficult for time-series method and econometric model to realize complex nonlinear short-term passenger flow prediction. Meanwhile, the learning process of artificial neural network is usually slow, and the adaptability to emergencies is poor. SVR provides a new theoretical guidance for nonlinear time-series prediction because of its good generalization ability to deal with small samples and nonlinear data.

To address this situation, the study integrates information such as real-time passenger flow information of the rural leisure tour industry through data analysis technology. Meanwhile, a short-term passenger flow prediction model is constructed through the SVR model to make the passenger flow prediction values more targeted [7]. The prediction model influencing factors are first analyzed before constructing the model. According to the actual situation, the study initially determines three influencing factors, which are the amount of web-related keyword searches, the passenger flow data of yesterday, and various indicators of the weather on the day of travel, and put them as the to-be-entered items of the prediction model.

The passenger flow data of yesterday can provide managers with information such as passenger flow and passenger flow trend. These data help understand the needs of tourists and improve service levels. Meanwhile, the waste of vehicles and human resources can be minimized through reasonable scheduling and operational arrangements to achieve cost savings. Nowadays, people can easily obtain a lot of travel-related information through the Internet to customize their itinerary. Therefore, it is necessary to include the search volume of related keywords in the model. Weather indicators on the day of travel affect whether people decide to travel.

Grey correlation is a method to determine which factor has the greatest impact on the final results. The correlation is a measure of how much they change over time or from object to object for the factors between two systems. The basic idea is to judge whether the connection is close according to the similarity of the shape of the set of sequence curves. The closer the curves are, the greater the correlation between the corresponding sequences, and the smaller the correlation is. The Grey correlation analysis method is selected to make secondary selection of these influencing factors to determine the final influencing factors of the model. The related process is shown in Figure 3.1.

In Figure 3.11, the impact factor to be considered is first determined. Meanwhile, a correlation degree calculation is performed to calculate the correlation between the impact factor and the output sequence. Based on the results of this calculation, the influence factor is selected twice to obtain the final influence factor of the model. There are certain differences in magnitudes among the data, which can affect the analysis results. Therefore, these data need to be pre-processed. The Z-score function can compare and analyze the values between different data sets. The principle is to subtract the mean from the original data and divide by the standard deviation. The result is the Z-score value. The Z-score function is used to standardize the data. Therefore, there is comparability between different data sets to better data analysis and mining. The impact

factor is normalized and de-scaled by the function, and the relevant calculation formula is shown in Eq. (3.1).

$$x^* = \frac{x - v}{\sigma} \quad (3.1)$$

In Eq. (3.1),  $x$  and  $x^*$  represent the original series and the de-quantized series, respectively. The mean of  $x$  is expressed as  $v$ . The standard deviation of  $x$  is expressed as  $\sigma$ . The human comfort index is calculated by quantitative analysis to take into account all weather indicators. The formula for calculating the human comfort index is given in Eq. (3.2).

$$ssd = (1.818 * t + 18.18)(0.88 + 0.0002 * f') + \frac{t' - 32}{45 - t'} - 3.2 * v + 18.2 \quad (3.2)$$

In Eq. (3.2), the human comfort index is set as  $ssd$ , the average temperature is expressed as  $t'$ , the relative humidity is set as  $f'$ , and the wind speed is expressed as  $v$ , respectively. The nonlinear mapping function is set as  $\phi(x)$ . The overhead feature space is set as  $F$ .  $x_i$  represent the model influence factors.  $i$  denotes the ordinal number. The sample data are defined as  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$ .  $y_i$  denotes the target output value. The SVR model uses  $\phi(x)$ , which can map  $x_i$  into  $F$ . The relevant regression problem in this space is solved to overcome the nonlinear regression problem in the original space. In this case, the SVR model function is shown in Eq. (3.3).

$$f(x) = w^T \phi(x) + b, \phi : R^n \rightarrow F, w \in F \quad (3.3)$$

In Eq. (3),  $f(\cdot)$  denotes the function,  $w$  and  $b$  denote the variables, and  $R$  denotes the function. Because SVR adopts the principle of minimizing structural risk, it is necessary to find the corresponding function to solve the related function regression problem, and the related calculation formula is shown in Eq. (3.4).

$$R_{SVR}(C) = R_{emp} + \frac{1}{2} \|\omega^2\| = \frac{1}{n} \sum_{i=1}^n |y_i - f(x_i)| + \frac{1}{2} \|\omega^2\| \quad (3.4)$$

In Eq. (3.4), the structural risk function is expressed as  $R_{SVR}(\cdot)$ . The penalty function is set as  $C$ , which is a constant. The larger the value of  $C$  is, the smaller the allowable correlation error will be, and the corresponding generalization ability will be worse, which means the complexity of the model is higher. The empirical risk function is expressed as  $R_{emp}(\cdot)$ .  $\frac{1}{2} \omega^2$  denotes the Euclidean norm.  $C \frac{1}{n} \sum_{i=1}^n |y_i - f(x_i)|$  denotes the training set error. The relevant expression in  $|y_i - f(x_i)|$  is shown in Eq. (3.5).

$$|y_i - f(x_i)| = \begin{cases} 0, & |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| - \varepsilon, & \text{else} \end{cases} \quad (3.5)$$

In Eq. (3.5),  $\varepsilon$  denotes the insensitive loss function, and more stable estimation results can be obtained by introducing  $\varepsilon$ . The function regression problem is represented by minimizing the cost generalization function, and its related expression is shown in Eq. (3.6).

$$\min \frac{1}{2} \|\omega^2\| + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (3.6)$$

In Eq. (3.6),  $\xi_i$  and  $\xi_i^*$  denote the relaxation variables, which are introduced to make the solution of Eq. (3.7) exist.

$$\begin{cases} y_i - w^T \phi(x) - b \leq \varepsilon + \xi_i^*, & i = 1, 2, \dots, n \\ -y_i + w^T \phi(x) + b \leq \varepsilon + \xi_i, & i = 1, 2, \dots, n \\ \xi_i, \xi_i^* \geq 0, & i = 1, 2, \dots, n \end{cases} \quad (3.7)$$

In solving the functional regression optimization problem, the Lagrange multiplier is introduced and the optimality condition (Karush-Kuhn-Tucher, KKT) is used to obtain the regression function as shown in Eq. (3.8).

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (3.8)$$

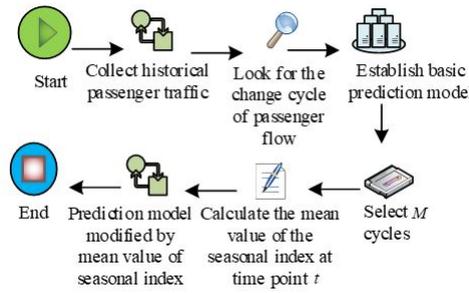


Fig. 3.2: Process of the SEA method

In Eq. (3.8),  $\alpha_i$  and  $\alpha_i^*$  denote Lagrange multipliers,  $\alpha_i \in [0, C]$ . The kernel function  $K(x_i, x) = \phi(x_i)\phi(x)$  can satisfy the Mercer condition. The Gaussian radial basis kernel function is chosen for use in the prediction model, and the relevant mathematical expression is shown in Eq. (3.9).

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \exp\left(\frac{-|x_i - x|^2}{\sigma^2}\right) + b \tag{3.9}$$

In Eq. (3.9), exp represents the multiplicative power operation function of e.  $K(x_i, x) = \exp\left(\frac{-|x_i - x|^2}{\sigma^2}\right)$  and  $x$  are the optimal support vectors obtained from the model search. In the passenger flow forecasting, the cyclical variation affects the forecasting results. Therefore, the study chooses SEA to reduce this effect, and its related process is shown in Figure 3.2.

In Figure 3.2, firstly, the historical statistics of passenger flow are collected. The period of passenger flow change is set as  $T$ , which is searched for. The basic prediction model is constructed, and  $M$  cycles are selected to calculate the seasonal index, and the relevant calculation formula is shown in Eq. (3.10).

$$C_{s,t} = \frac{a_{s,t}}{a'_{s,t}} \tag{3.10}$$

In Eq. (3.10), the selected historical periods are denoted as  $s, s = 1, 2, \dots, M$ . The specific time points of the prediction are denoted as  $t, t = 1, 2, \dots, T$ . The sample true value is set as  $a_{s,t}$ . The sample prediction result is  $a'_{s,t}$  after prediction by the original model. The mean value of the seasonal index at the time points of  $t$  is calculated. The prediction model is modified by this mean value. The relevant calculation formula is shown in Eq. (3.11).

$$\begin{cases} AJ_t = \frac{1}{M} \sum_{s=1}^M C_{s,t} \\ F_t(x_t) = \varphi(x_t) * AJ_t \end{cases} \tag{3.11}$$

In Eq. (3.11), the forecast model adjusted by setting the SEA method is  $F_t(x_t)$ .  $\varphi(x_t)$  denotes the original model. The average seasonal adjustment index at the time point of  $t$  is  $AJ_t$ , which is also the compensation coefficient. The forecasting model after introducing the seasonal adjustment index is shown in Eq. (3.12).

$$f(x_{n+t}) = \left(\sum_{s=1}^M (\alpha_i - \alpha_i^*) K(x_i, x_{n+t}) + b\right) * AJ_t \tag{3.12}$$

The formula at  $AJ_t$  can be interpreted as meaning that the seasonal index for each of the past  $M$  years contributes to the coefficient for the current compensation coefficient. The problem can be transformed so that each of the given  $M$  models affect the solution of  $AJ$ . These models form an integrated model. The optimal collaborative model, the best  $AJ$ , is obtained by assigning to each model the corresponding weight  $w'_i$ . In

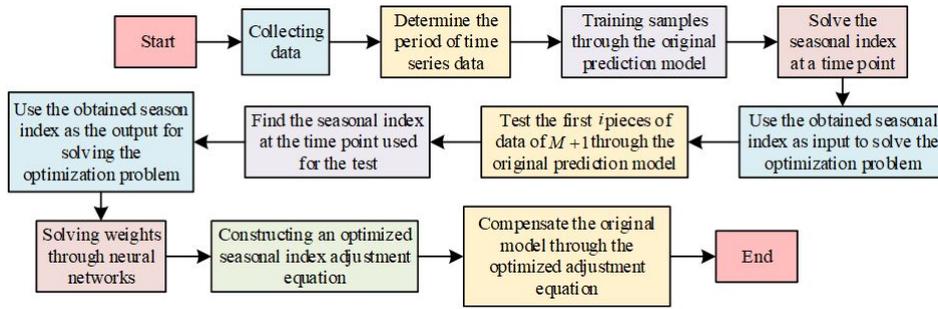


Fig. 3.3: Specific process for improving the SEA method

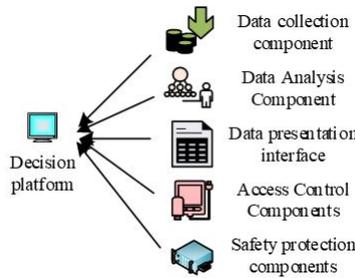


Fig. 3.4: Main functional components of the platform

solving the weights of each model, the solution can be performed by a neural network. Therefore, the specific flow of the improved SEA method is shown in Figure 3.3.

In Figure 3.3, data collection is performed to determine the time-series data for T, and the data need to have M period history data. A training sample is taken from  $\varphi(x_t)$  and the results obtained are used as input to solve the optimization problem by solving AJ for each time point according to the formula of  $AJ_t$ . The mathematical expression of the output vector  $\lambda$  is shown in Eq. (3.13).

$$\lambda = \frac{a_{m+1,t}}{a'_{m+1,t}} \tag{3.13}$$

In Eq. (3.13),  $t = 1, 2, \dots, m$ . m denoted the number of historical data periods. The weights are solved by using neural network. The number of hidden nodes is set to 1 and the weights from hidden nodes to output nodes are automatically adjusted to 1. The optimized seasonal index is set to  $AJ_t^*$  and the correlation equation of this index is constructed as shown in Eq. (3.14).  $\varphi(x_t)$  is compensated by this equation.

$$AJ_t^* = \sum_{s=1}^M w'_s C_{s,t} \tag{3.14}$$

**3.2. Page anti-tampering technology and construction of related decision support platform.**

After completing the establishment of the short-term passenger flow prediction model, a rural leisure tour industry decision support platform is constructed. Visitors can see personalized leisure tour content recommendations, which are conducive to their choice of places to visit. Operators can receive comprehensive marketing analysis guidance data given by the platform and keep abreast of other operators' dynamics in this platform. Among them, the platform's main functional components are shown in Figure 3.4.

In Figure 3.4, the platform mainly includes data collection component, data analysis component, data display interface, access control component, and security protection component. The data collection component

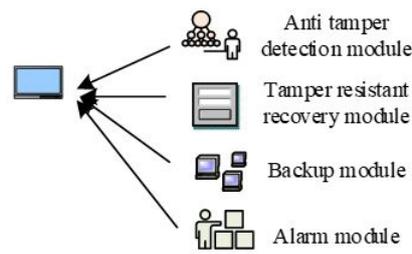


Fig. 3.5: Modules of the tamper-proof system

collects data such as real-time traffic information and the trajectory of visitors browsing the platform. The data analysis component makes use of data mining technology and relevant mathematical modeling to conduct directional analysis on the aggregated data. These data involve prediction of tourist flow, recommendation of popular products, development evaluation of leisure tourism merchants, etc. Then valuable information can be extracted, which can provide data support for the decision-making platform. The security protection component is distributed among the modules of the platform. For the platform pages, the anti-tampering technology of the pages makes them tamper-proof and protects the platform from network tampering attacks. Access control component is the core of the platform. This paper studies how to control system data access by designing role-based access control security access control component. The main method is to divide tourists, merchants, and regulators into three roles. Roles have different permissions based on regions. That is, roles in a region can access only the content of the current region and its subordinate regions. When the user accesses the data, the identity of the role (tourist, merchant or regulator) needs to be verified, and the relevant data can be obtained when the verification is passed. Meanwhile, the anti-tampering system has several modules, as shown in Figure 3.5.

In Figure 3.5, the anti-tampering system has four modules including anti-tampering detection module and backup module. The event triggering mechanism is combined with the core embedding technology in the anti-tamper detection module. The core technology is the notify feature in Linux in the event triggering mechanism through which the server-side files are monitored and illegal tampering is detected. Meanwhile, recovery is performed while tampering is detected with the aid of the resync file synchronization tool. Meanwhile, the core embedding technology will be based on Tomcat's filter technology for specific implementation, intercepting and verifying all Web requests. The location where the tampering occurred can be recovered in time and the backup files can be encrypted in the anti-tampering recovery module. The backup module backs up Web server-side files and has a disaster recovery role. In the alarm module, when the anti-tampering detection module finds an attack event, this alarm module sends an email to the administrator in time, etc. From the functions of different modules, the core of the anti-tampering system is the anti-tampering detection module, in which the core issue is the page tampering detection technology. In the core embedding technology to check the integrity of outgoing pages, the study selects Hashed Message Authentication Code (HMAC) encryption. Meanwhile, time-seeded random numbers are chosen for the selection of "challenge" numbers. HMAC is a message authentication code that uses a password hashing function, combined with an encryption key, to generate a message authentication code after a special calculation. It can be used to ensure data integrity and can be used to authenticate a message. The mathematical expression of the encryption algorithm is shown in Eq. (3.15).

$$HMAC(K', M) = H(K' \oplus opad || H(K' \oplus ipad || M)) \quad (3.15)$$

In Eq. (3.15), HMAC denotes the encryption algorithm, H denotes the chosen Hash algorithm, and the Message-Digest 5 (MD5) algorithm of the Hash algorithm is chosen in the study. The authentication password is denoted as  $K'$ . The block size processed in H is denoted as B. B is the processed block size, not the output Hash value. opad and ipad denote the repetition of 0x5a and 0x36, respectively. B is the repeated several times. The algorithm is applied to the page verification, and the related flow is shown in Figure 3.6.

In Figure 3.6, the system is first initialized, the file Hash value is calculated, and the calculation result is

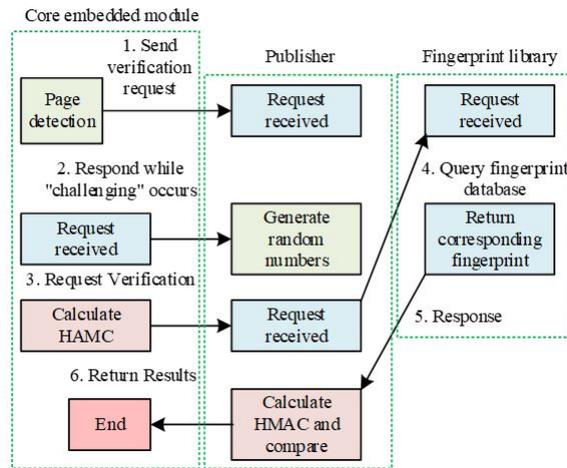


Fig. 3.6: Page verification process

backed up. Before the system is released to the public, the files to be released to the Web server need to be backed up. These files will be operated by MD5. Meanwhile, the corresponding file fingerprint information will be obtained and stored in the database for future verification work. One of the files has a unique location in the database. When the file is changed on the publishing server and the data is updated on the web server, the backup file is updated and the file hash value is recalculated. Meanwhile, the relevant fingerprint database is updated so that the relevant files are synchronized to ensure the normal operation of the system verification process. In the verification request, the filter is used as the client and the publishing server is used as the server based on the HMAC encryption idea. The filter sends a network verification request to the publishing server. The current system time is treated as a random number seed and a random number is generated and returned to the client as the answer upon receiving the request. The random number is the “challenge” number. The random number and the MD5 value of the Web page to be authenticated are encrypted with the HMAC-MD5 algorithm to obtain a message digest.

The page is verified and the message digest is sent to the server side. When the server side receives the message digest, it performs the HMAC-MD5 algorithm encryption on the basis of random numbers and fingerprint library keys. Meanwhile, the results obtained are compared with the message digest sent by the client. When the two are consistent, it directly informs the filter of this client. After this client releases the web page, the external request can get the relevant Web resources. On the contrary, it sends emails and other messages to the administrator to inform him/her and calls the data backup module to recover the tampered contents. For the page verification process, it is transient in nature, and the process is executed according to the authentication rules specified in advance. It is unforgeable in nature, and the random seed is selected according to the current time character of the publishing server. The random number generated from the seed is used as the “challenge number”, which cannot be known by the third party in advance. The random number generated by the seed is used as the “challenge number”, which cannot be known by the third party in advance to ensure the security of the page verification.

#### 4. Application analysis of decision support platform for rural leisure tourism industry.

**4.1. Performance analysis of the improved SVR model.** The ARIMA model mainly focuses on the regression of the data itself and the extrapolation of the time trend, which is a kind of time series. The main purpose of ARIMA is to predict the future value of the deformation series. The improved SVR model used in the study was analyzed. The Autoregressive Integrated Moving Average Model (ARIMA), an SVR model, was used as a comparison model to study the forecast results of the three models in this dataset for 2019, as shown in Figure 4.1.

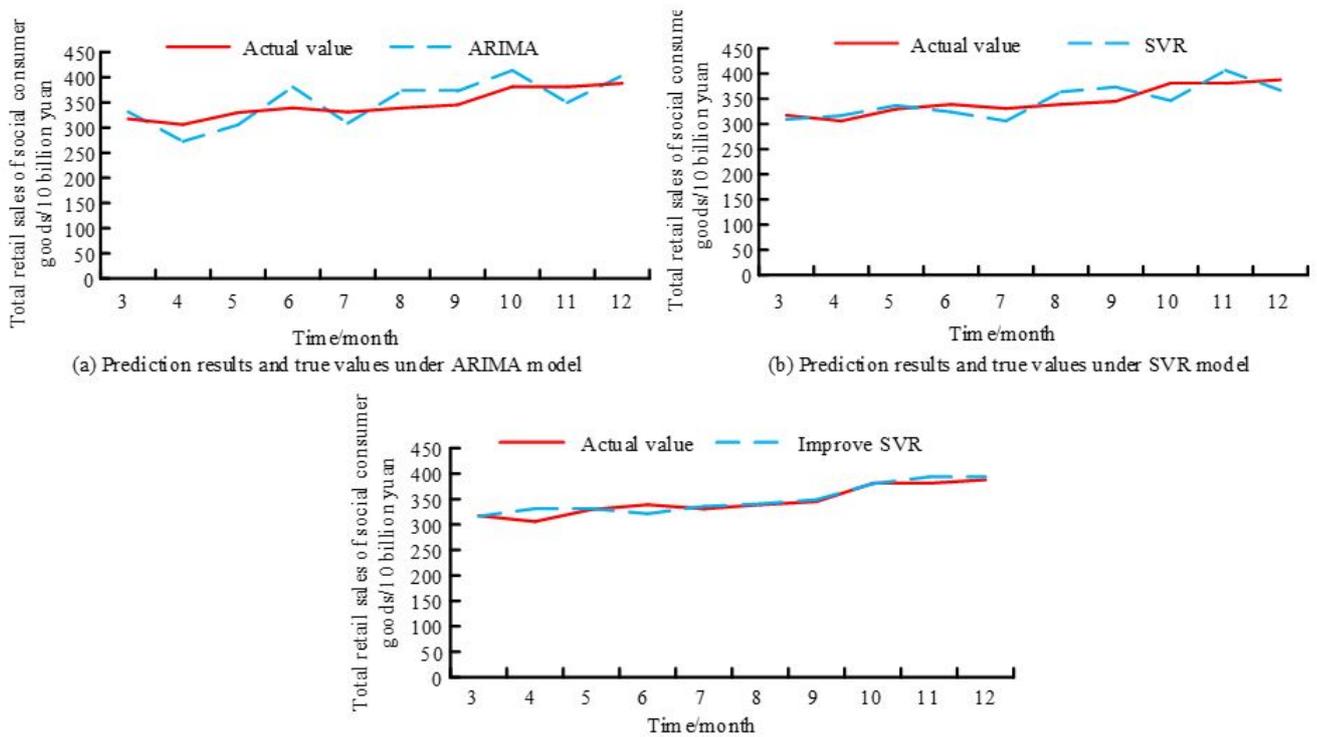


Fig. 4.1: Prediction results of three models

In Figure 4.1(a), the prediction results of different months have different degrees of difference from the actual value, and the predicted value fluctuates above and below the actual value. In Figure 7(b), the change of the prediction results is similar to that in Figure 7(a), and the magnitude of fluctuation decreases. The prediction results in Figure 7(c) also fluctuate above and below the actual value, but the magnitude of fluctuation is relatively minimal. In Figure 4.1(a), in March, the prediction result of ARIMA model is 33.128 billion yuan, which is 14.02 billion yuan more than the actual value, and the actual value is 317.26 billion yuan. In May, the prediction result of the ARIMA model is 305.30 billion yuan, which is 33.31 billion yuan less than the actual value. In Figure 4.1(b), in October, the predicted value of SVR model is \$346.29 billion, which is \$34.75 billion less than the actual value. In August, the predicted value of SVR model is 363.64. In Figure 4.1(c), in March, the predicted value of improved SVR model is \$316.28 billion, which is \$0.98 billion less than the actual value and \$7.05 billion more than the SVR model. The prediction effect of the improved SVR model is known based on the difference between the predicted and actual values of the three models. Absolute error refers to the difference between the measured value and the true value of a measurement. The smaller the value, the more accurate the prediction. The absolute errors of the predictions of the three models were obtained to further analyze the prediction differences of these three models as shown in Figure 8.

In Figure 4.2(a), the absolute error of prediction of the ARIMA model fluctuates up and down within 6.00% to 13.00%. In March, the corresponding absolute error of prediction is 4.42%, which is 6.47% smaller than that of April, while the absolute error of model prediction is 10.89% in mid-April. In May, the corresponding absolute error of prediction is 12.30%, which is the maximum absolute error of prediction of the ARIMA model. In May, the corresponding absolute error of prediction is 12.30%, which is the maximum absolute error of prediction of ARIMA model, and 6.29% larger than the minimum absolute error of prediction of the model. In Figure 4.2(b), the maximum absolute error of prediction of the SVR model is 9.12%, corresponding to the month of October, which is 3.73% larger than that of December. In May, the minimum absolute error of prediction of the model is 2.20%, while the absolute error of prediction corresponding to June is 4.37%. In Figure 4.2(c), the minimum

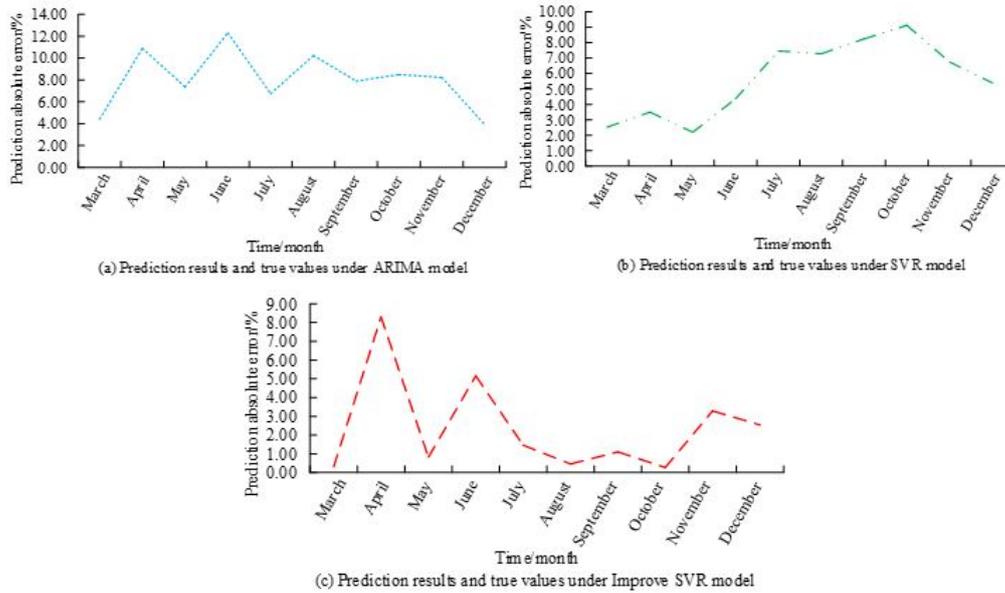


Fig. 4.2: Prediction absolute error of three models

absolute error of prediction of the improved SVR model is 0.27%, which is smaller than the minimum absolute error of the other two models. In June, the corresponding absolute error of prediction is 5.16%, which is 4.06% larger than that of September, while the corresponding absolute error of prediction in November is 3.29%. In April, the corresponding absolute error of prediction is the largest, which is 8.30%. This is smaller than that of the other models. The maximum absolute error of prediction is smaller than that of the other models. This shows that the improved SVR model has better performance among these three models.

**4.2. 2 Analysis of the effect of application in decision support platform based on the improved SVR model.** In this study, the traffic data were provided by the Jiuzhaigou tourism official website (<http://www.jiuzhaigou.com/>). The daily traffic data of 2013-2016 were selected, a total of 9.56 million people. Historical weather indicators were queried through historical meteorological website (<http://mlishi.tianqicom/>), including the highest temperature, lowest temperature, wind speed, air, and relative humidity, etc. The human comfort index was to calculate the index value according to the indexes of the day. Then the index was used to classify and determine the comfort level. The related keyword search volume was provided by Baidu Search Index (<http://index.baidu.com/>). The keywords "Jiuzhaigou Valley", "Jiuzhaigou ticket" and "Jiuzhaigou tourism" were selected in this study. The daily keyword search volume was calculated as the input factor. Then the desensitization of sensitive information, such as partial hiding or replacement of personal identity information such as name, ID number and phone number, was studied. Therefore, it could ensure that the real identity of users was not exposed in the anonymization process. Meanwhile, a strict access control mechanism was established to prevent data from being obtained by unauthorized personnel.

The constructed compensated prediction model was analyzed and simulation experiments were conducted. The OSX EI Capitan system was selected to obtain the 2017-2020 passenger flow data from an official tourism website. Therefore, queries on each index of relevant historical weather were conducted through the historical weather network. Meanwhile, the keyword search volume of Jiuzhaigou was queried through the Baidu search index. The time-series were plotted from January 2016 to October 2019 in Figure 4.3.

Figure 4.3(a) shows the daily passenger flow statistics, Figure 4.3(b) shows the daily human comfort index statistics, and Figure 4.3(c) shows the daily search volume statistics for the keyword "jiuzhaigou". The data in these three subplots show a certain periodicity, especially in Figure 4.3(b), where the periodicity of the data is

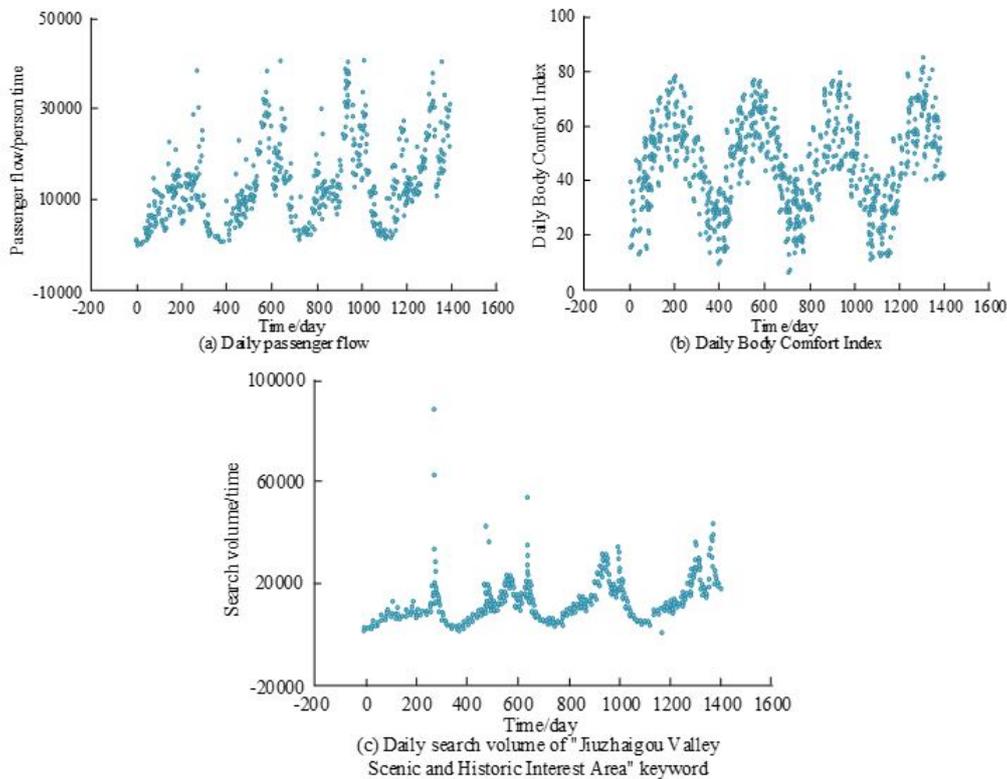


Fig. 4.3: Relevant time-series diagram

more obvious. In Figure 4.3(a), when the time is on the 264th day, visitors are 39,457, which is 39,408 more than that on the 400th day, which is 49. When the time is on the 389th day, visitors are 114. In Figure 4.3(b), the situation of peak human comfort index occurs near the 200th, 600th, 1000th, and 1400th days. The time interval of the peak appears tends to decrease as time goes on. In Figure 4.3(c), except for the keyword “jiuzhaigou”, which is searched for more than 60,000 times, the keyword is searched for less than 40,000 times at other times. The collected data are calculated and processed. The SVR model is selected to predict the daily passenger flow. Meanwhile, the “jiuzhaigou” Baidu search factor is used as an input factor to analyze its impact on the model prediction, as shown in Figure 4.4.

In Fig. 4.4(a), the corresponding traffic values are different at different times. The overall difference between the model prediction and the actual sample value is smaller after adding the web search factor. On day 5, passengers are 6597, which are 49 more than passengers at the same time without the web search factor. On day 7, passengers are 7335, which are 668 more than passengers at the same time without the web search factor, and 357 less than the actual value. In Figure 4.4(b),  $R$  denotes the correlation coefficient and  $MSE$  denotes the mean square error. After adding the network search factor, the  $MSE$  value of the training output is 0.0845, which is 0.0153 smaller than that without the network search factor. The  $R$ -value is 0.9284, which is slightly larger than that without the network search factor. The inclusion of network search factor can improve the prediction effect, which affects the passenger flow. The keyword “jiuzhaigou” was added to the input factor. The SVR model was improved to make the prediction. The SVR, ARIMA, and seasonal SVR models were used for comparison.

In Figure 4.5(a), based on the graph, the improved SVR model used in the study is located in a fold that is closer to the actual sample data. At day 6, the improved SVR model predicts 6668 passengers, which is 1665 passengers less than the SVR model considering the season and 102 passengers less than the actual value.

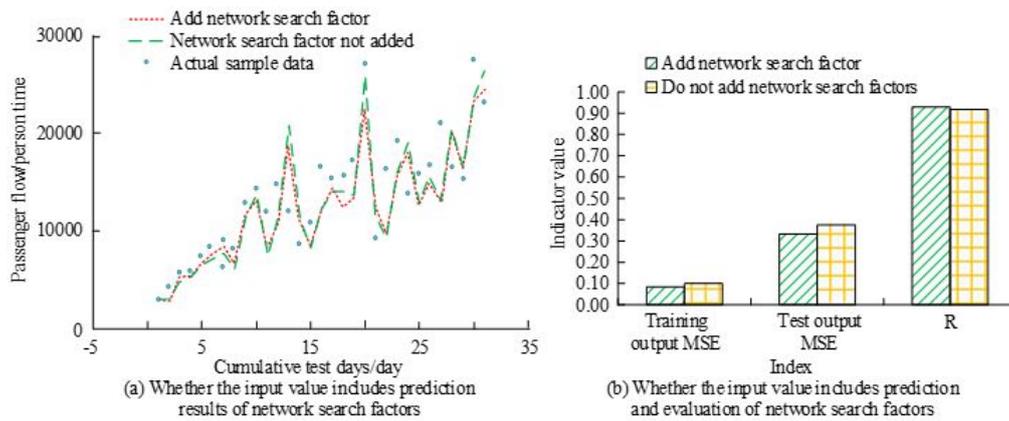


Fig. 4.4: Relevant prediction results

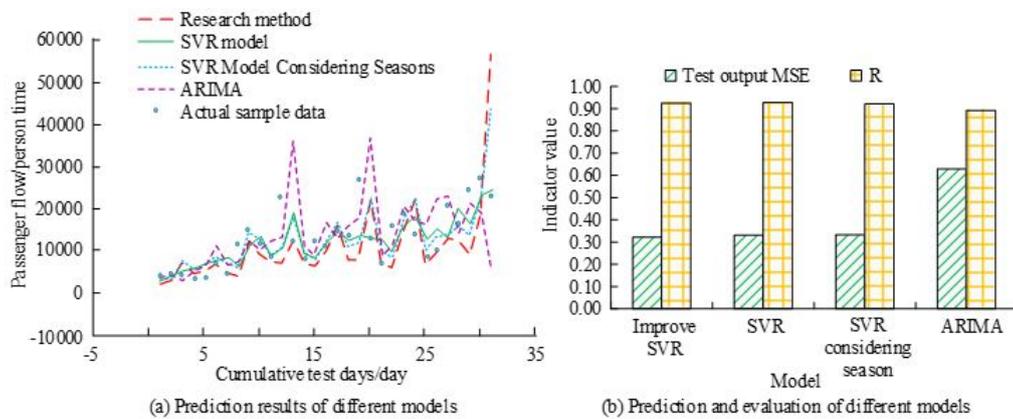


Fig. 4.5: Prediction results of different models

At day 21, its predicted passenger flow is the same as the actual value. In Figure 4.5(b), the tested output MAE value of the ARIMA model is 0.6299 at the maximum, which is 0.3075 larger than that of the improved SVR. This corresponds to a minimum MAE value of 0.3224. The improved SVR model has a higher R-value of 0.9254, which is slightly smaller than the SVR but larger than the other two algorithms. This shows that the improved SVR model has better performance. The performance of the anti-tampering scheme is analyzed, and the related results are shown in Figure 4.6.

From Figure 4.6, adding core embedded programs to the server affects the performance of the server. When the concurrency is less than 100 times/s, the average response time difference before and after the program is added is very little. The concurrency is greater than 100 times/s, and the time difference between these two increases continuously as the concurrency increases. When the concurrency is 150 times/s, the response time under adding the core embedded program is 15.11s, which is 3.83s more than when it is not added.

**5. Conclusion.** A decision support platform was constructed for the leisure tour industry in the countryside to grasp the development of the leisure tour industry in various parts of the countryside, understand the local passenger flow. The improved SEA method was used, the SVR model was improved, and a related passenger flow prediction model was constructed. An anti-tampering system was established to combine the event triggering mechanism with the core embedding technology for anti-tampering detection. The HMAC en-

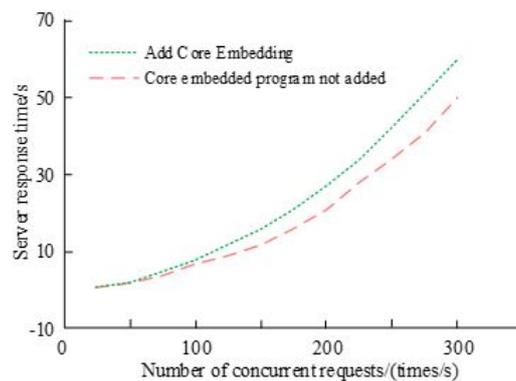


Fig. 4.6: Performance Analysis of Tamper Proof Schemes

encryption algorithm was used for page verification to prevent the page from being tampered. The results showed that the improved SVR model had better performance compared with SVR and ARIMA. The difference of the improved SVR model with the actual value was smaller. In March, the prediction value was \$31.628 billion, which was \$0.98 billion less than the actual value and \$7.05 billion more than the SVR model. The absolute error of prediction was smaller. The minimum absolute error of prediction was 0.27%, which was smaller than the other two models. Adding the network search factor improved the accuracy of passenger flow prediction. After adding the network search factor, the MSE value of the training output was 0.0845, which was 0.0153 smaller than that without the network search factor. The R value was 0.9284, which was slightly larger than that when the network search factor was not added. In the analysis of the application of the improved SVR model, its prediction performance was better than that of the other models. At day 6, the improved SVR model predicted 6668 passengers, which was 1665 passengers less than the SVR model considering the season and 102 passengers less than the actual value. The minimum MAE value of this model's test output was 0.3224, which was 0.3075 smaller than the ARIMA model. The core embedded program affects the performance of the server in the performance analysis of the anti-tampering scheme. The concurrency was greater than 100 times/s, and its response time increased. This shows that the application of the research method is good. Since the time of page verification is not short, the verification technique can be improved later to improve the verification efficiency.

**Funding.** The research is supported by: Empowering Rural Tourism New Model by using Digital Technology - Research on Educational Science Planning in Hubei Province of China (No. B2021549).

#### REFERENCES

- [1] L. CHASE, M. STEWART, B. SCHILLING, B. SMITH, AND M. WALK, *Agritourism: Toward a conceptual framework for industry analysis*, Lyson Center for Civic Agriculture and Food Systems, 8 (2018), pp. 13–19.
- [2] R. DING, Y. LI, X. YU, P. PENG, AND L. WEI, *Characteristics of rural agritainment sewage in sichuan, china*, Water Science & Technology, 79 (2019), pp. 1695–1704.
- [3] M. HASANIPANAH, A. SHAHNAZAR, H. B. AMNIEH, AND D. J. ARMAGHANI, *Prediction of air-overpressure caused by mine blasting using a new hybrid pso-svr model*, Engineering with Computers, 33 (2017), pp. 23–31.
- [4] T. HEIN, *A legacy of outstanding angertainment success*, Fruit & Vegetable Magazine, 73 (2017), pp. 14–16.
- [5] D. HU, C. ZHANG, W. CAO, X. LV, AND S. XIE, *Grain yield predict based on gra-adaboost-svr model*, Journal on Big Data, 3 (2021), pp. 65–76.
- [6] J. HU, X. QIAN, H. CHENG, H. TAN, AND X. LIU, *Remaining useful life prediction for aircraft engines based on phase space reconstruction and hybrid vns-svr model*, Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 41 (2021), pp. 3415–3428.
- [7] J. HU, X. QIAN, C. TAN, AND X. LIU, *Point and interval prediction of aircraft engine maintenance cost by bootstrapped svr and improved rfe*, Journal of Supercomputing, 79 (2023), pp. 7997–8025.
- [8] J. LIN AND H. LI, *A short-term pv power forecasting method using a hybrid kmeans-gra-svr model under ideal weather condition*, Journal of Computer and Communications, 8 (2020), pp. 102–119.

- [9] D. LIU, *Development and utilization of rural idle homesteads in the context of rural revitalization—a case study of leisure agriculture*, Asian Agricultural Research, 11 (2019), pp. 57–60.
- [10] P. SINGH, P. GUPTA, AND K. JYOTI, *Tasm: technocrat arima and svr model for workload prediction of web applications in cloud*, Cluster Computing, 22 (2019), pp. 619–633.
- [11] X. XIAO, *An empirical study on angertainment consumption behavior of urban residents – a case study of huangpi district of wuhan city*, Asian Agricultural Research, 10 (2018), pp. 92–99+104.
- [12] D. XU, Q. ZHANG, Y. DING, AND H. HUANG, *Application of a hybrid arima-svr model based on the spi for the forecast of drought-a case study in henan*, Journal of Applied Meteorology and Climatology, 59 (2020), pp. 1239–1259.
- [13] R. YU, H. CHEN, G. CHEN, AND C. LIU, *Spatial distribution of rural tourism destination and influencing factors in hubei province-a case study of high-star agritainment*, Economic Geography, 38 (2018), pp. 210–217.
- [14] J. ZHANG, Z. M. WU, Y. F. PAN, AND Z. Z. GUO, *Predicting method of simple reaction time of driver based on svr model*, Zhongguo Gonglu Xuebao/China Journal of Highway & Transport, 30 (2017), pp. 127–132.
- [15] L. ZHENG AND H. LIU, *Identification of focal actors in the translation of the rural tourism actor-network: a case in china*, Environmental Engineering and Management Journal, 17 (2018), pp. 1813–1823.

*Edited by:* Zhengyi Chai

*Special issue on:* Data-Driven Optimization Algorithms for Sustainable and Smart City

*Received:* Nov 21, 2023

*Accepted:* Jun 3, 2024