Research Article

A Hybrid Model for Prediction in Asphalt Pavement Performance Based on Support Vector Machine and Grey Relation Analysis

Xuancang Wang [ID],[1] **Jing Zhao** [ID],[1] **Qiqi Li** [ID],[1] **Naren Fang** [ID],[1] **Peicheng Wang** [ID],[2] **Longting Ding** [ID],[1] and **Shanqiang Li** [ID][1]

[1]*School of Highway, Chang'an University, Xi'an 710064, China*
[2]*School of Information Engineering, Chang'an University, Xi'an 710064, China*

Correspondence should be addressed to Jing Zhao; zhaojingzi0203@163.com

Pavement performance prediction is a crucial issue in big data maintenance. This paper develops a hybrid grey relation analysis (GRA) and support vector machine regression (SVR) technique to predict pavement performance. The prediction model can solve the shortcomings of the traditional model including a single consideration factor, a short prediction period, and easy overfitting. GAR is employed in selecting the main factors affecting the performance of asphalt pavement. The SVR is performed to predict the performance. Finally, the data collected from the weather station installed on Guangyun Expressway were adopted to verify the validity of the GRA-SVR model. Meanwhile, the contrast with the grey model (GM (1, 1)), genetic algorithm optimization BP [[parms resize(1),pos(50,50),size(200,200),bgcol(156)]]081%, −0.823%, 1.270%, and −4.569%, respectively. The study concluded that the nonlinear and multivariate prediction model established by GRA-SVR has higher precision and operability, which can be used in long-period pavement performance prediction.

1. Introduction

Big data maintenance is a central issue in highway management. Highway maintenance mileage accounted for 97.7% of the mileage of traffic in China by the end of 2018. Notably, the expressway has been transferred from the construction to the maintenance period. With the popularity of big data technology, roads have entered the era of big data maintenance. However, the reason why the performance of asphalt pavement is a vital component of maintenance management and operation is that the rational allocation of maintenance decision-making and maintenance funds are determined by an accurate prediction model in the later period. Therefore, the scientific establishment of the pavement performance prediction model is significant for asphalt pavement maintenance and can provide a model for big data maintenance.

The pavement management system (PMS) is applied for road life cycle management. However, it generally uses analytical tools and statistical methods to predict pavement performance [1]. Predicting of pavement performance is critical, but it is very complex, because the performance of asphalt pavement is affected by the combination of structural design, material properties, construction quality traffic load natural factors, and maintenance [2]. The pavement performance prediction model is a relationship that characterizes the variation of pavement performance with time, material, and traffic load [3].

There are different methods available for the determination of pavement performance; many scholars have attempted to develop a scientifically derived accurate model. There are four types of prediction models: uncertainty model, certainty model, dynamic model, and bionic model [4].

(i) Uncertainty model: the commonly used model is the grey theoretical model that has the characteristics of a small amount of data, high prediction accuracy, and a simple calculation method. Therefore, it is widely used in pavement performance prediction. For example, Zhang et al. [5], Shen and

Du [6], Wang and Li [7], and Zhang and Ji [8] used this model to predict pavement smoothness and rutting. Peng et al. [9] applied Weibull distribution to pavement performance prediction and obtained ideal results.

(ii) Certainty model: it is an empirical method; it takes advantage of using traditional regression as a tool to fit the data that come from experiments and finite element mechanics to get the form and parameters of the model. For example, Sun and Liu [10] proposed the decay equation of asphalt pavement performance which was obtained through engineering experiment. Abed et al. [11] investigated the variability effect of thickness and stiffness of pavement layers; they used the Monte Carlo method to obtain the probability distribution function of pavement performance by using the parameters. Gong et al. [12] proposed a regularized regression method to estimate the asphalt concrete moduli with data available from the long-term pavement performance (LTPP) database.

(iii) Bionic model: this model has high prediction accuracy. Yang et al. [13] used the genetic neural network model to estimate rutting and driving quality. Bianchini and Bandini [1] proposed the neuro-fuzzy hybrid model to predict the present serviceability index (PSI). Ferreira and Lima Cavalcante [14] and Beltran and Romo [15] presented the application of artificial neural networks (ANN) in pavement performance.

(iv) Dynamic model: it is based on the traditional model. For instance, Shen et al. [16] improved the traditional grey model and proposed a dynamic grey model. Chen et al. [17] combined the US PME model with grey prediction theory and mechanical experience method, proposed a dynamic grey prediction model, and established a DGM-PME combination model to forecast the rutting. Chu and Durango-Cohen [18] used the autoregressive moving average time series state space method to predict the structural strength of the pavement. El-Badawy et al. [19] developed a comprehensive bottom-up fatigue cracking distress dynamic prediction model integrating the Mechanistic-Empirical Pavement Design Guide (MEPDG) and performance test methodology.

To date, various pavement performance prediction models have been proposed by scholars, but the models still have defects. For example, the grey model just adopts the time factor and does not take into account other factors such as natural environment and traffic load which may have maximum impact. And, as the forecast period increases, the stability and accuracy of the prediction are decreased. Weibull distribution model is only suitable for small sample data prediction. The certainty model is mainly determined by factors like the initial performance index of asphalt pavement and the road age. It is simple and convenient to use for it does not consider the reasonable dynamic data. It only can predict short period performance. The genetic neural network and ANN model are prone to overfitting when data are insufficient. The dynamic prediction model can make full use of the later data to predict longer periods. Simultaneously, the reason why the model can only consider the impact of time on pavement performance is that the model is based on time series. Hence, a new model is needed to be devised to be applied to pavement performance prediction.

Recently, support vector machines have been applied in various fields. Zhao et al. [20] proposed a k-means and SVM hybrid model for the development of an electric vehicle urban driving cycle. Hoang et al. [21] used it to recognize the pavement crack. Wang et al. [22] proposed a support vector machine online model for predicting metro ridership. Karballaeezadeh et al. [23] applied this model to the prediction of road residual life and compared the model with an artificial neural network (ANN) and multilayer perceptron (MLP) models. The results show that the support vector machine model has the highest accuracy.

The factors affecting the performance of asphalt pavement were processed firstly by GRA. The SVR with advantages of minimizing structural risk and strong generalization performance was then used to establish a hyperplane as a decision surface. Finally, the asphalt pavement performance prediction model was established to provide a model that can be applied to maintenance decision-making, maintenance fund investment, and big data of pavement maintenance.

The structure of this paper is as follows. Section 2 mainly introduces the main principles of GRA-SVR. Section 3 contains the modeling process of the whole model. Section 4 mainly uses the model to verify the example. Finally, the results are analyzed.

2. Methodology

2.1. Basic Principles of GRA-SVR

2.1.1. Basic Principle of Grey Relation Analysis. The grey system theory holds that the complex objective systems which are all ordered and discrete data must contain inherent laws [24]. There are many factors affecting the performance of asphalt pavement, but the effects of various factors are not very clear, so that we can call the factors grey. Therefore, GRA is used to quantitatively reflect the correlation between asphalt pavement performance and various factors. This method can find the main factors from many factors that affect pavement performance. The corresponding statistical data of the influencing factors in the system are converted into geometric curves by the method, and the closer the curve geometry is to the dependent variable, the greater the degree of association is [25].

2.1.2. Basic Principle of Support Vector Machine Regression. SVR is a model derived from the support vector machine (SVM) proposed by VAPNIK [26]. The SVM model is a machine learning method that mainly solves the

classification problems of small samples, nonlinearities, and high-dimensional data [27, 28]. Its principle is based on the VC theory of statistical principle and structural risk minimization, and the optimal solution in data mining is sought by establishing an optimal hyperplane [29]. Usually, we reduce the dimension of the sample to simplify the problem, while the SVM method is the opposite. It uses the kernel function to map the sample points to high-dimensional and even infinite-dimensional space to deal with linear problems as shown in Figure 1.

Regression is essentially similar to classification. The SVM classification model is to manage a plane so that the support vectors of the two classification sets or all the data are farthest from the classification plane, and the SVR model is to find a regression plane so that all data of a collection could be closest to the plane, as shown in Figure 2. The SVR can predict the prediction vector of the test data by establishing a nonlinear relationship between the data tested in the training data and the support vector. Most of the various influencing factors of asphalt pavement performance are nonlinear. The specific method is as follows.

Assume the sample set $(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)$, $x \in R^n, y \in R, x \in R^n, y \in R$, Then, y and x in the sample set can be expressed as follows [2]:

$$f(x) = w \cdot x + b, \tag{1}$$

where w and b are the coefficients of the hyperplane.

If the original data fit well with the support vector machine regression, then $\min 1/2\|w^2\|$ is as follows [2]:

$$\text{s.t.} \quad \begin{cases} w \cdot x_i + b - y_i \leq \varepsilon, \\ y_i - w \cdot x_i - b \leq \varepsilon, \\ \qquad\qquad i = 1, 2, \ldots, l, \end{cases} \tag{2}$$

where ε is a positive number.

Equation (1) is transformed into (3) by introducing the Lagrangian logarithm [2]:

$$f(x) = w \cdot x + b = \sum_{i=1}^{1} (a_i - a_i^*)(x_i \cdot x) + b, \tag{3}$$

where a_I and a_i^* are the sample support vectors, which take a value of zero in most cases.

The above process is the linear regression principle of SVR, but the effects of the factors including rainfall, traffic volume, maximum temperature, and minimum temperature for the pavement performance are nonlinear. When dealing with the nonlinear problem of the SVR, the sample x_i is mapped to a high-dimensional space by $\psi: x \longrightarrow H$. An optimal hyperplane should be constructed to solve the "dimensionality disaster"; the inner space operation is implemented using the original spatial parameters when ψ is unknown. The internal kernel function $K(x_i, x_j) = \psi(x_i) \times \psi(x_j)$ can be obtained when the kernel function satisfies the condition of Mercer [30]. At the same time, Lagrange changes are introduced to get equation (4) [31]:
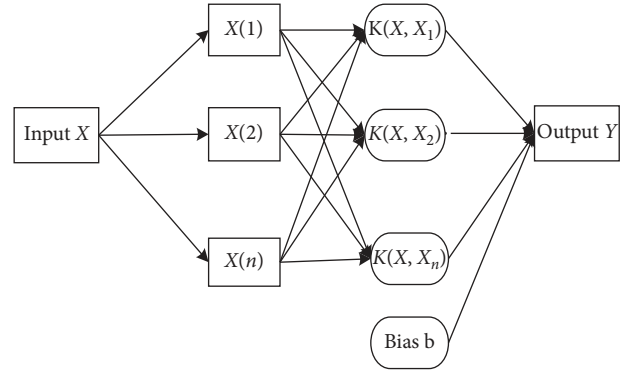

FIGURE 1: Architecture diagram of support vector machine.

$$L(w, \xi, b, a, \beta) = \sum_{i=1}^{1} a_i - \frac{1}{2} \sum_{i=1}^{1} y_i y_j a_i a_j K(x_i x_j). \tag{4}$$

Finally, the transformed regression function [31] is as follows:

$$f(x) = w \cdot x + b = \sum_{i=1}^{1} (a_i - a_i^*) K(x_i \cdot x) + b. \tag{5}$$

This method can avoid overfitting caused by traditional methods. SVR nonlinear regression fitting could control the fitting process by increasing the dimension. The high generalization performance that is closely related to the choice of kernel function is a big advantage of SVR.

Commonly used kernel functions are listed as follows [32]:

(1) Linear kernel function: $K(\mathbf{x}, x_i) = \mathbf{x}^T x_i$;

(2) Polynomial kernel function: $K(\mathbf{x}, x_i) = (\mu \mathbf{x}^T x_i + r)^p, \mu > 0$;

(3) RBF kernel function: $K(x, x_i) = \exp(-g\|x - x_i\|^2), g > 0$;

(4) Sigmoid kernel function: $K(\mathbf{x}, x_i) = \tanh(\mu \mathbf{x}^T x_i + r), \mu > 0$.

μ, r, and p are parameters of the kernel function.

However, each type of kernel function has different advantages and disadvantages:

Linear kernel functions are used to generalize linear samples.

Polynomial kernel functions are mostly used to process text data.

Although Sigmoid kernel function has higher accuracy, it is complicated, which increases the complexity of the whole model.

Therefore, in this paper, the RBF kernel function is used for support vector machine regression prediction.

2.2. Construction of GRA-SVR Asphalt Pavement Performance Model

2.2.1. Selection of the Best Parameters.
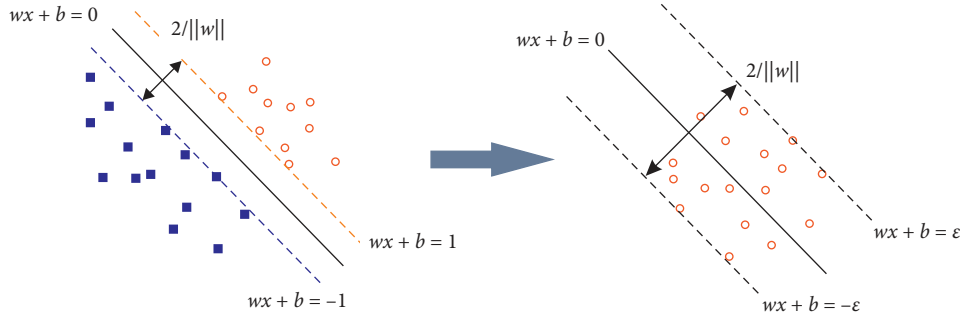It is important to select the appropriate penalty parameter c and kernel

FIGURE 2: Difference of SVM and SVR diagram.

function parameter g to ensure the accuracy of the entire model when using SVR for prediction. Therefore, the CV method is generally adopted to solve this problem, which is a statistical analysis method for verifying the performance of the model. The principle is to group the original data and divide them into verification and training sets. In this way, it is possible to effectively avoid the states of underlearning and overlearning and ultimately obtain the accuracy. Common CV methods are as follows:

(1) Hold-Out Method: the method randomly divides the data into two categories: one is the training set used to train the model, and the other is the verification set used to verify the model [20]. The final accuracy is the performance metric of the model.

(2) LOO-CV: assuming there are N samples in the original data, that is why the model is called N-CV, so each sample is an independent verification set, and the remaining N-1 samples are training sets; thus, N models were obtained. The average accuracy of the final validation set is used as a performance indicator for the model. However, due to the high computational cost, the model has difficulties in practical operation.

(3) K-CV: the original data are equally divided into K groups. The data of each group are used as verification set once, and the remaining data of other K-1 groups are used as a training set; therefore, K models are obtained. Then, the average of the classification accuracy calculated from the final verification set of those K models is used as the performance index of this model [33]. This method is more accurate due to the fact that it can effectively avoid the states of underlearning and overlearning.

According to the comparable selection of the three methods, the K-CV model is finally adopted to cross-validate and select the best penalty parameter c and function parameter g. The specific method is as follows. Firstly, the parameters c and g are limited to a specific range, and then the K-CV model is used for the training set in the range to obtain the accuracy. Finally, the parameters c and g which make the training set with the highest accuracy are selected as the optimal parameters. The concrete implementation can be implemented using the libsvm3.20 tool.

2.2.2. Construction of Asphalt Pavement Performance Model.
The pavement performance is affected by many factors. The factors, acting on performance, are uncertain and non-linear. Hence, the performance and factors integrate a grey system. Therefore, the grey correlation analysis can be used as an attribute processor to select several important influencing factors, and then the SVR is used to perform the regression prediction. Through the establishment of the comprehensive model GRA-SVR to predict the trend of pavement performance under the influence of various factors, the specific modeling process is shown in Figure 3.

Specific steps are as follows:

(1) Select dependent and independent variables.

(2) Establish a raw data matrix: $Xi = \{x_i(k), k = 1, 2, \ldots, n\}$, $i = 1, 2, 3, 4$. $x_i(k)$ represents a certain level of the first influencing factor.

(3) Data normalization.

(4) Calculating the difference sequence [34] is as follows:

$$\Delta_i(k) = |x_0'(k) - x_i'(k)|,$$
$$\Delta_i = (\Delta_i(1), \Delta_i(2), \ldots, \Delta_i(n)), \quad (6)$$
$$i = 1, 2, \ldots, m.$$

(5) Achieving the largest and smallest difference of the sequence [34] is as equation (7). Write the maximum value as M and the minimum value as N:

$$M = \max_i \max_k \Delta_i(k),$$
$$N = \min_i \min_k \Delta_i(k). \quad (7)$$

(6) Calculating the correlation coefficient of each sample [35] is as follows:

$$\gamma_{0i}(k) = \frac{m + \xi M}{\Delta_i(k) + \xi M}, \quad \xi \in (0, 1), k = 1, 2, \ldots, n;$$
$$(8)$$
$$i = 1, 2, \ldots, m.$$

ξ is called the resolution coefficient. When $\xi \leq 0.5463$, the resolution is the best. Usually, the value of ξ is 0.5, which is also taken in this paper.
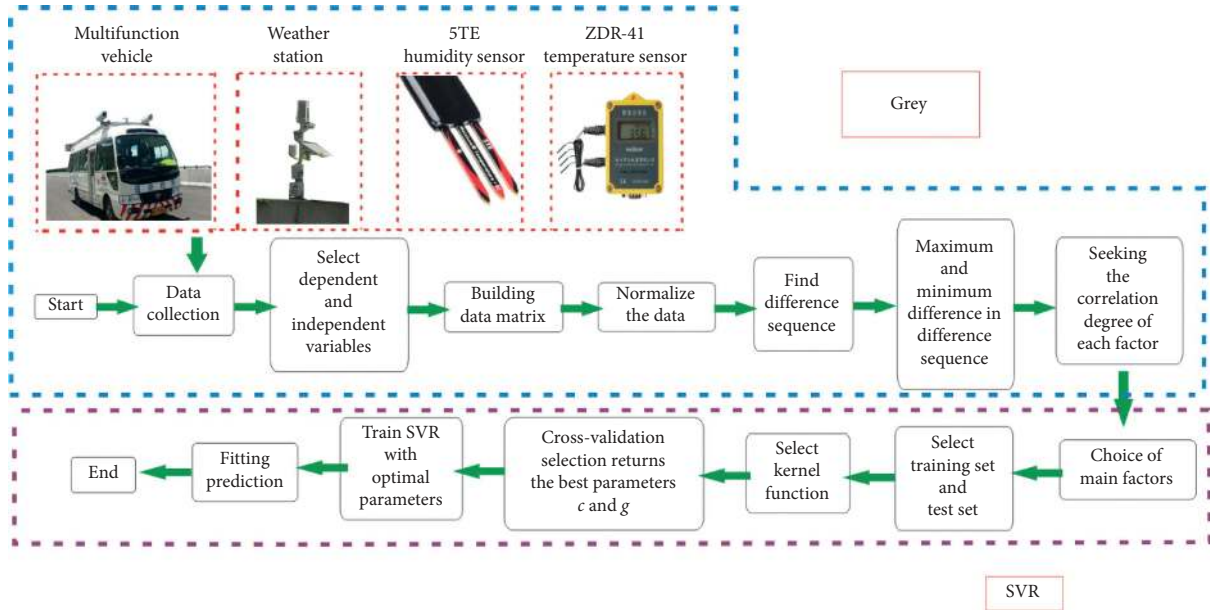
FIGURE 3: Flowchart of the GRA-SVR modeling process.

(7) Calculating the correlation between each influencing factor and the system [35] is as follows:

$$\gamma_{0i} = \frac{1}{n} \sum_{i=1}^{n} \gamma_{0i}(k); \quad i = 1, 2, \ldots, m. \qquad (9)$$

(8) Choose the factors that have a greater influence on pavement performance.

(9) To improve the accuracy and training speed of the model and prevent big numbers of consuming decimals during the calculation process, the data should be normalized and processed to the interval [0, 1].

(10) RBF which is researched has a high precision [36, 37], and this paper selects the RBF kernel function to predict the performance.

(11) K-CV model is used to cross-validate and select the best penalty parameter c and function parameter g.

(12) Using the optimal parameters for SVR fitting, the prediction data are obtained.

3. Case Verification

3.1. Data Acquisition. This paper is based on the highway from Guangzhou to Yunfu (Guangyun highway) and the installed weather station in 2010, and it can collect the climate data including road temperature, humidity, wind speed, and solar radiation. The installation details and pavement structure are shown in Figures 4 and 5. Among them, the pavement temperature detection uses the ZDR-41 temperature sensor, subgrade temperature, and humidity testing to use a 5TE sensor (see Figure 6). The climate of Guangdong province is humid and the temperature is extremely high, rising to 41°C. Under the influence of large traffic volume, the rutting is serious as shown in Figure 7. The RDI prediction models GRA-SVR, PPI, GA-BP, and GM

(1, 1) were established to analyze the accuracy of each model, which were based on the RDI, maintenance funds, traffic volume, and data collected by the weather station from 2011 to 2018 (see Table 1 for the survey results).

The factors, pavement structure, and materials should be considered in performance prediction. Usually, the pavement structure needs to be calculated as a numerical value. To address this issue, the structures number [12, 38–40] is usually adopted. However, it needs to be calculated in two cases as follows:

(i) Different structures: in this case, the thickness and material of each layer of the road are different. The structural number [41] (SN) is adopted according to the AASHTO guide for design of pavement structures. The road network level performance prediction can apply this case. The specific calculation method is as follows:

$$\text{SN} = a_1 D_1 + a_2 D_2 m_2 + a_3 D_3 m_3, \qquad (10)$$

where a_i is i^{th} layer coefficient; this parameter needs to be obtained through experiments, D_i is i^{th} layer thickness, and m_i is the ith layer drainage coefficient.

(ii) Same structure: the performance of the pavement material can be affected by the environment, and the structural bearing capacity is changed. The pavement structural bearing capacity can be expressed by the pavement structure strength ratio (SSR) [42]. The specific calculation method is

$$\text{SSR} = \frac{l_0}{l}, \qquad (11)$$

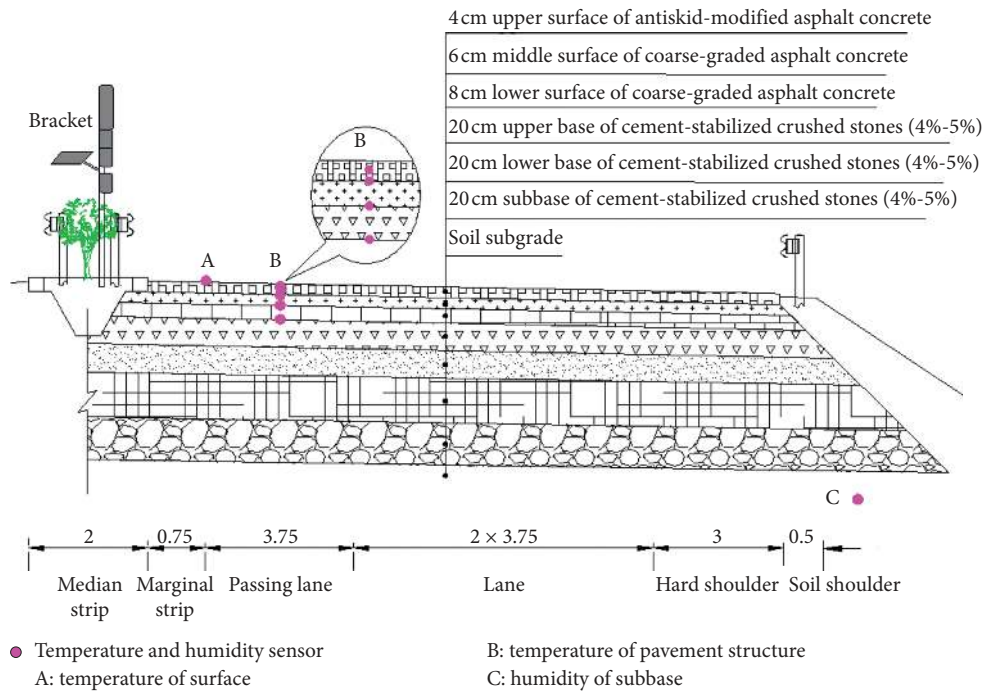where l_0 is pavement deflection standard value (0.01 mm), where l is pavement measurement

Bracket

4 cm upper surface of antiskid-modified asphalt concrete
6 cm middle surface of coarse-graded asphalt concrete
8 cm lower surface of coarse-graded asphalt concrete
20 cm upper base of cement-stabilized crushed stones (4%-5%)
20 cm lower base of cement-stabilized crushed stones (4%-5%)
20 cm subbase of cement-stabilized crushed stones (4%-5%)
Soil subgrade

B

A   B

C ●

2	0.75	3.75	2 × 3.75	3	0.5
Median strip	Marginal strip	Passing lane	Lane	Hard shoulder	Soil shoulder

● Temperature and humidity sensor   B: temperature of pavement structure
A: temperature of surface   C: humidity of subbase
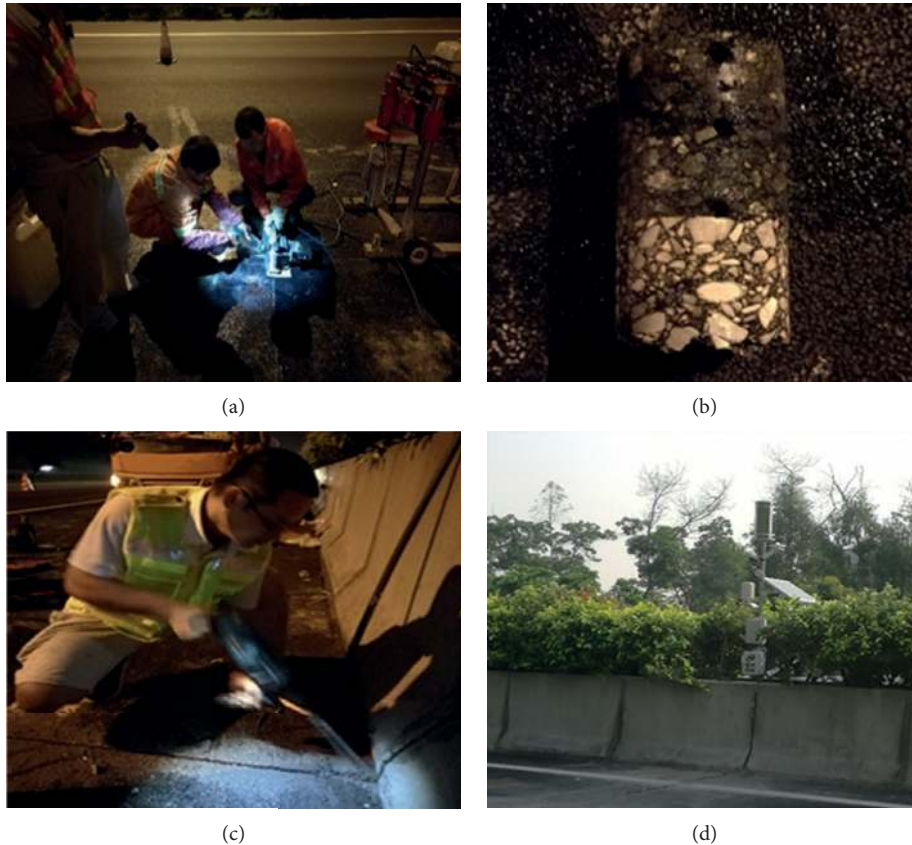
Figure 4: Sensor layout.

(a)

(b)

(c)

(d)

Figure 5: Weather station layout. (a) Drilling cores of asphalt pavement, (b) installation of temperature sensor in a pavement structure, (c) installation of temperature sensor on road surface, and (d) bracket mounting.
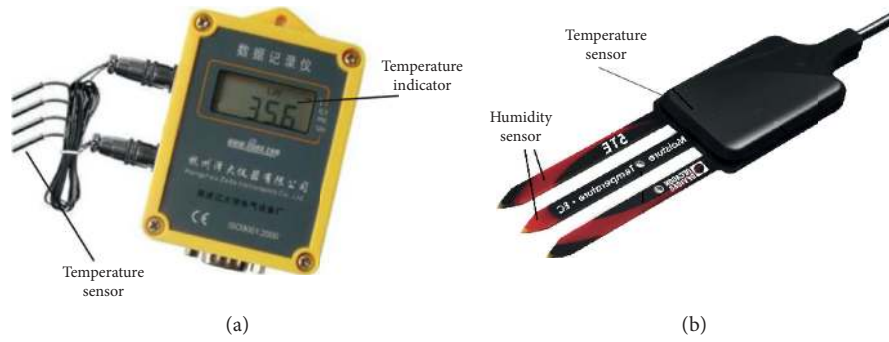
FIGURE 6: (a) Pavement sensor and (b) subgrade sensor.

FIGURE 7: Rutting of Guangyun Expressway.

representing deflection (0.01 mm); this parameter needs to be obtained through multifunction vehicle.

This paper relies on engineering only one pavement structure, so the calculation of SSR represents the influence of pavement structure on pavement performance.

3.2. Grey Relation Analysis. The correlation of the data can be analyzed in Table 2; the correlation degree of each influencing factor can be obtained, as shown in Table 2.

The effects of various factors on rutting are sorted as follows:

$$\gamma_2 < \gamma_9 < \gamma_{18} < \gamma_7 < \gamma_{12} < \gamma_{10} < \gamma_8 < \gamma_{17} < \gamma_{14} < \gamma_{13} < \gamma_{16}$$
$$< \gamma_5 < \gamma_{19} < \gamma_{11} < \gamma_{15} < \gamma_6 < \gamma_1 < \gamma_4 < \gamma_3.$$

(12)

Generally, the greater the degree of relevance, the better the correlation of factors to the main direction of system development, that is, the greater the influence of this factor on the evaluation index. When $\gamma > 0.8$ is well correlated, when $\gamma = 0.6 \sim 0.8$, the correlation is good. We can see that γ of these 18 factors is greater than 0.6, indicating that these factors have an impact on the rutting. Among the 19 factors, γ of 12 factors is greater than 0.8, indicating that these 12 factors have a strong influence on the formation of rutting.

So, the better relevant factors that have the greatest impact were selected to establish the model, and the other factors were removed. The selected results are as follows:

Equivalent single axle loads > maintenance funds > pavement structure strength ratio > mean value of soil moisture > highest temperature in the middle surface > highest temperature in the road surface > annual cumulative total radiation > annual average rainfall > lowest temperature in middle surface > highest temperature in the upper surface > lowest temperature of upper surface > highest temperature in lower surface.

The following can be observed from the above analysis:

(1) The primary factor, the formation of rutting, is the equivalent single axle loads. The greater equivalent single axle loads are, the more serious the rutting is. The reason is that, under the action of traffic load, large shear stress will be generated in the asphalt pavement, which will cause irreversible cumulative deformation in the surface layer.

(2) The maintenance funds have a significant repairing effect on the rutting. For example, in this section of the highway, the maintenance funds were RMB 81,500 in 2013. The traffic volume and rainfall increased, but the rutting disease was significantly improved in 2014.

TABLE 1: Datasheet of RDI and various influencing factors of Guangyun Expressway (2011–2018).

Year	2011	2012	2013	2014	2015	2016	2017	2018
RDI	94	90.2	90.1	91.4	89.8	86.4	84.6	85.5
PCI	99.7	97.9	97.1	95.1	92.5	91.4	88.7	87.5
SRI	98	95.5	89.5	80.8	86.9	84.2	84.4	85.6
SSR	2.61	1.63	1.58	1.92	1.36	1.27	1.35	1.15
Service life	2	3	4	5	6	7	8	9
Equivalent single axle loads (10^3)	1214	1504	1600	1860.15	1984.46	2173.91	2386.93	2023.45
Maintenance funds (/million yuan)	3.23	0.875	8.15	6.34	7.64	8.54	8.0	6.57
Annual average rainfall (/mm)	1667.7	1490.5	1647.6	2224.5	1752.5	1645.6	2321	2013.1
Mean value of soil moisture	17.4	17.8	19.5	22.1	18.6	17.5	22.4	19.1
Mean value of environment humidity (%RH)	75.3	74.5	83.8	77.3	76	73.7	72.1	88.8
Annual maximum wind speed (m/s)	7.4	6.2	6.1	5.8	5.3	6.3	5.6	7.9
Highest temperature of environment (/°C)	37.6	37.8	38.9	39.2	39.6	39	40.4	38.2
Lowest temperature of environment (/°C)	3.4	2.5	2.5	2.7	0.1	3.8	3.7	4.2
Highest temperature of road surface (/°C)	65.1	61.8	60.3	62.4	68.5	62.5	65.1	65.2
Lowest temperature of road surface (/°C)	4.2	6.3	5.9	4.3	5.9	6	6.4	5.8
Highest temperature of upper surface (/°C)	55.1	56.2	57.1	60.2	58.6	58.8	57.9	63.1
Lowest temperature of upper surface (/°C)	7.5	7.9	6.9	7.8	6.5	7.5	8	8.1
Highest temperature in middle surface (/°C)	56.7	60.3	59.4	58.5	60.5	59.4	61.2	42.8
Lowest temperature in middle surface (/°C)	6.3	5.4	5.8	6.5	6.8	6.9	6	6.1
Highest temperature in lower surface (/°C)	46.7	47.6	46.9	45.5	44.2	46.7	43.2	44.9
Lowest temperature in lower surface (/°C)	9.1	8.5	8.9	9.3	9.5	10.2	9.8	10.4
Annual cumulative total radiation	1014	1085	1045	1054	1240	1093	1105	1166

(3) The degree of relevance SSN is 0.9301. It shows that SSN has a greater impact on the rutting. The specific reason is that water, solar radiation, and temperature have an impact on the pavement material, and the structural bearing capacity is insufficient, resulting in the occurrence of rutting.

(4) The annual cumulative radiation ages the asphalt and accelerates the formation of the rutting. After the aging of the asphalt, the overall shear resistance of the asphalt surface layer is reduced, resulting in a decrease in the rutting resistance. For example, the annual cumulative radiation was the largest in 2015, and the rutting in 2016 was more serious.

(5) The maximum shear stress generally occurs in the midsurface, and the rainfall and wind speed accelerate the heat dissipation of the highest temperature of the environment and road surface. Based on the above factors, the influence of the highest temperature of the middle layer on the formation of the rutting is greater than the highest temperature of the road surface and the upper layer.

(6) Under the action of traffic load, the water infiltrated into the asphalt surface layer by soil and rainfall will become high-pressure water, which will reduce the bond behavior between asphalt and aggregate, resulting in lower pavement strength and lower resistance to rutting.

(7) The lowest temperature of the road surface would cause other diseases on the asphalt pavement, which indirectly lead to the occurrence of rutting.

The dimensionally reduced data are normalized by software, and the processing results are shown in Table 3.

3.3. Penalty Parameter Selection. In this paper, the optimal penalty parameter c and function parameter g are solved by K-CV cross-validation model to select the best penalty parameter c and function parameter g (see Figure 8). The axis of abscissa indicates the value of c after taking the base 2 logarithm. The ordinate axis represents the value of g after taking the base 2 logarithm. Contour lines indicate errors in the range of c and g. When the error is the smallest, the corresponding c and g are the best. First, c and g are initially selected. The range of c is within $2^{\wedge}(-6) \sim 2^{\wedge}(6)$ and that of g is within $2^{\wedge}(-8) \sim 2^{\wedge}(8)$. When the error is 0.0572, the optimal penalty parameter is $c = 64.0$ and $g = 0.0039$.

By primary election, the range of values for c can be reduced to $2^{\wedge}(-3) \sim 2^{\wedge}(2)$ and g can be reduced to $2^{\wedge}(-4) \sim 2^{\wedge}(4)$(see Figure 9). At the same time, reduce the interval between the contour and the three-dimensional view. When the error is 0.0605, the optimal penalty parameter is $c = 4.0$ and $g = 0.0884$.

4. Results and Discussion

The GRA-SVR, GM (1, 1) [43], GA-BP [44], and PPI model were applied and compared to predict the RDI of 2018 which was based on the training set consisting of various factors and RDI from 2011 to 2017. The PPI [10]model is as follows:

$$PPI = PPI_0 \left\{ 1 - \exp\left[-\left(\frac{a}{y} \right)^{\beta} \right] \right\}, \tag{13}$$

where PPI is the performance index; PPI_0 is the initial performance index; y is the road age; α and β are mode parameters. In this paper $PPI_0 = 94$; $y = 8$; $\alpha = 13.2$; $\beta = 1.409$.

TABLE 2: Relevance of each influencing factor.

| Influencing factor	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}	γ_{11}	γ_{12}	γ_{13}	γ_{14}	γ_{15}	γ_{16}	γ_{17}	γ_{18}	γ_{19}
γ	0.9301	0.6794	1.0397	0.9698	0.8539	0.9241	0.7685	0.7999	0.7052	0.7998	0.8806	0.7866	0.8409	0.8326	0.9049	0.8472	0.8077	0.7593	0.8622

γ_1 is pavement structure strength ratio; γ_2 is the service life; γ_3 is the equivalent single axle loads; γ_4 is the maintenance funds; γ_5 is the average annual rainfall; γ_6 is the mean value of soil moisture; γ_7 is the mean value of environment humidity; γ_8 is the annual maximum wind speed; γ_9 is the highest temperature of environment; γ_{10} is the lowest temperature of the environment; γ_{11} is the highest temperature of road surface; γ_{12} is the lowest temperature of road surface; γ_{13} is the highest temperature of upper surface; γ_{14} is the lowest temperature of upper surface; γ_{15} is the highest temperature in middle surface; γ_{16} is the lowest temperature in middle surface; γ_{17} is the highest temperature in lower surface; γ_{18} is the lowest temperature in lower surface; γ_{19} is the annual cumulative total radiation.

TABLE 3: Standardized data after normalization.

| Time	2011	2012	2013	2014	2015	2016	2017
RDI	1	0.596	0.585	0.723	0.553	0.191	0
Equivalent single axle loads	0	0.247	0.329	0.551	0.657	0.818	1
Maintenance funds	0.307	0	0.949	0.713	0.883	1	0.930
Pavement structure strength ratio	1	0.269	0.231	0.485	0.067	0	0.060
Mean value of soil moisture	0	0.080	0.420	0.940	0.240	0.020	1
Highest temperature in middle surface	0	0.800	0.600	0.400	0.844	0.600	1
Highest temperature of road surface	0.585	0.183	0	0.256	1	0.268	0.585
Annual cumulative total radiation	0	0.314	0.137	0.177	1	0.350	0.403
Average annual rainfall	0.213	0	0.189	0.884	0.315	0.187	1
Lowest temperature in middle surface	0.600	0	0.267	0.733	0.933	1	0.400
Highest temperature of upper surface	0	0.216	0.392	1	0.686	0.725	0.549
Lowest temperature of upper surface	0.667	0.933	0.267	0.867	0	0.667	1.000
Highest temperature in lower surface	0.795	1	0.841	0.523	0.227	0.795	0
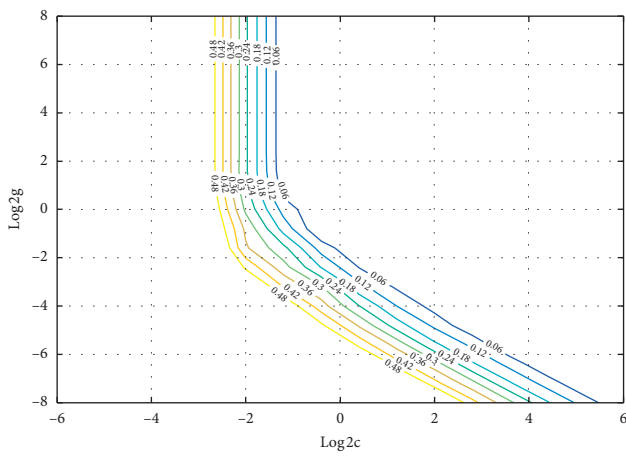
(a)
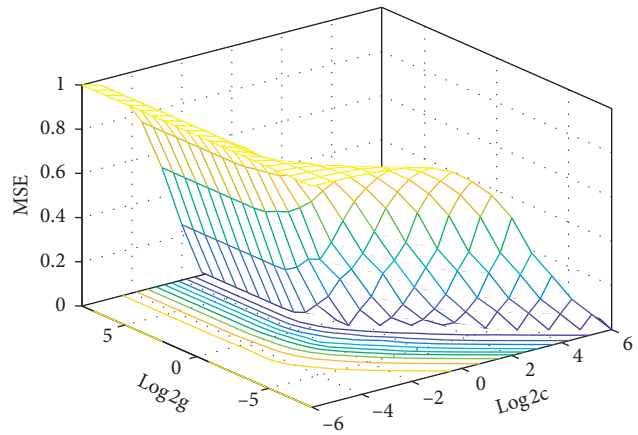
(b)

FIGURE 8: Best primary selection of penalty parameters. (a) Parameters c and g versus the accuracy rate in two dimensions; (b) parameters c and g versus the accuracy rate in three dimensions.
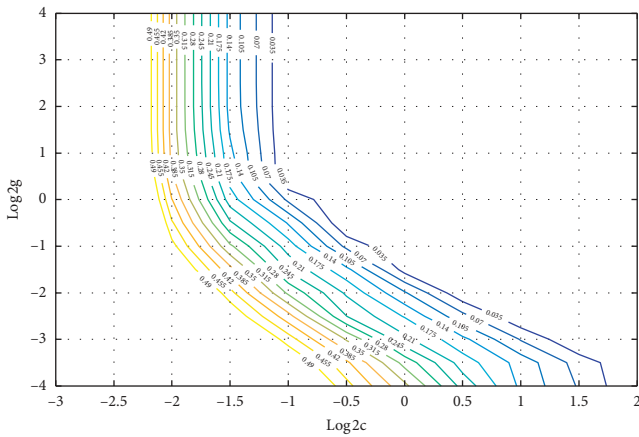
(a)
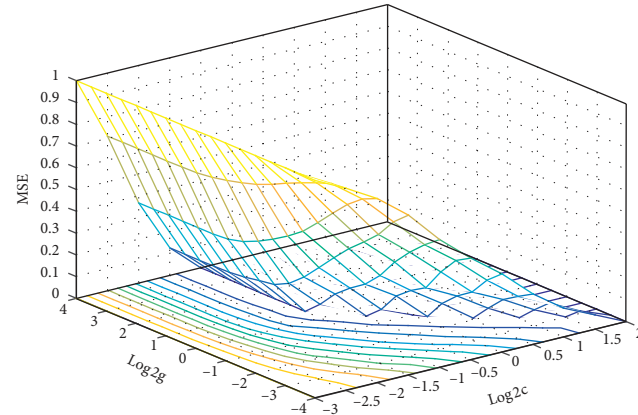
(b)

FIGURE 9: Best final selection of penalty parameters. (a) Parameters c and g versus the accuracy rate in two dimensions and (b) parameters c and g versus the accuracy rate in three dimensions.

TABLE 4: Comparison of predicted and actual values of RDI.

Time	Original value	GRA-SVR		GM (1, 1)		GA-BP		PPI	
		Predictive value	Absolute error	Predictive value	Absolute error	Predictive value	Absolute error	Predictive value	Absolute error
2011	94.0	94.00	—	94.00	—	94.00	—	94.00	−0
2012	90.2	90.27	−0.070	91.65	1.447	90.19	0	93.97	3.801
2013	90.1	90.03	0.068	90.47	0.368	90.10	−0.002	93.57	3.87
2014	91.4	90.63	0.724	89.30	−2.096	91.41	0.003	92.15	2.166
2015	89.8	89.73	0.070	88.16	−1.645	89.80	0	89.49	2.348
2016	86.4	86.47	−0.070	87.02	0.621	86.05	−0.352	85.84	3.091
2017	84.6	84.67	−0.066	85.90	1.301	84.26	−0.341	81.59	1.242
2018	85.5	85.56	−0.069	84.80	−0.704	86.58	1.082	79.10	3.907

TABLE 5: Precision comparison of forecast results for the three models.

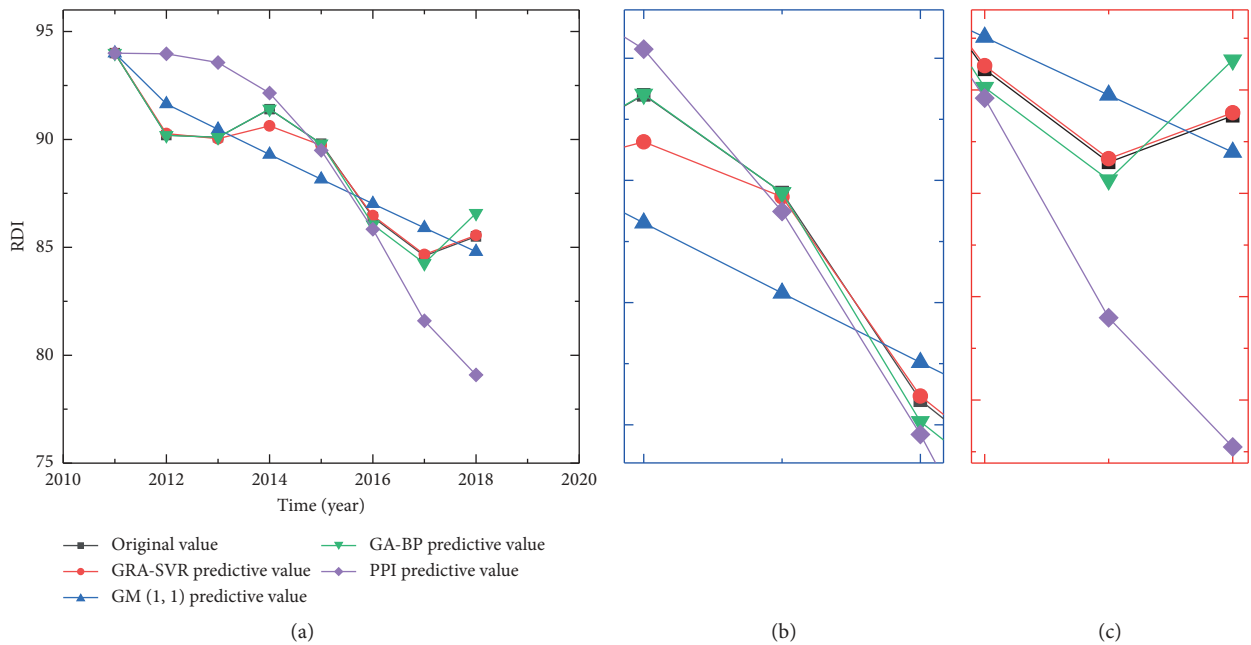| Model	Correlation coefficient	RMSE	Relative error (%)
GRA-SVR	0.992	0.298	−0.081
GM (1, 1)	0.856	1.304	−0.823
GA-BP	0.984	0.448	1.270
PPI	0.879	3.270	−4.569

FIGURE 10: Trend charts of RDI predicted value of different models.

The comparative analysis of the predicted and actual values of different models is shown in Table 4, the accuracy comparison was shown in Table 5, sand the corresponding variation trend and actual value of different models were shown in Figures 10 and 11.

The evaluation parameters of the four models obtained from Table 5 in predicting RDI are as follows:

Correlation coefficient: GM (1, 1) (0.856) < PPI (0.879) <GA-BP (0.984) <GRA-SVR (0.992)

RMSE: GA-BP (0.298) < GRA-SVR (0.499) <GM (1, 1) (1.304) <PPI (3.270)

Relative error: GRA-SVR (0.081) <GM (1, 1) (0.823) <GA-BP (1.270) <PPI (4.569)

The GRA-SVR and GA-BP models all showed good performance in terms of the overall correlation and deviation of the predicted value from the true value. However, with respect to relative error in 2018, GRA-SVR is the best, followed by GM (1, 1). Figure 11 shows the relative errors of the predicted and true values for the four models from 2011 to 2018. It can be observed that the relative error of the GA-BP model is the smallest, higher than GRA-SVR in 2016, and higher than GM (1, 1) in 2018 from 2011 to 2015. This is
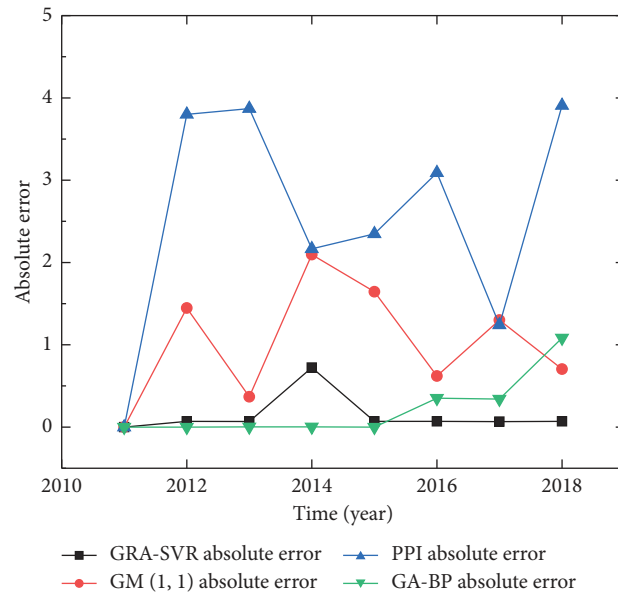
FIGURE 11: Trend charts of the actual value of different models.

TABLE 6: Performance comparison of four models.

Model	Operability	Prediction period	Accuracy	Consideration of factors
GRA-SVR	★	★★★	★★★	★★★
PPI	★★★	★	★	★★
GM (1, 1)	★★	★★	★★	★
GA-BP	★	★★	★	★★★

★ means performance in general, ★★ means better performance, and ★★★ means the best performance.

because the model is prone to overfitting for samples with small data, resulting in reduced prediction accuracy.

The trends of the predicted and actual values from different model RDIs were depicted in Figure 10(a). It can be seen that the GRA-SVR and GA-BP models display nonlinear trends, which are close to the actual value. The other two models show a linear relationship, which is different from the actual value.

All four models have good accuracy in short period prediction (see Figure 10(b)), but the accuracy would change with the prediction period increasing (see Figure 10(c)); the GRA-SVR model has the highest prediction accuracy because the old data were replaced by the new prediction data as the new training set. The GA-BP takes second place. Thirdly, the GM (1, 1) model just used the data of 7 years, and the accuracy reduced as the new data are not replenished in time with the time increases. The PPI model has the worst prediction accuracy, which was due to the fact that the model only uses the first-year data for prediction. As the prediction period increases, the controllability of the model decreases. In order to verify the accuracy of the model, the pavement surface condition index (PCI) and pavement skidding resistance index (SRI) prediction applied this model. The relative error was −0.115% and 0.111%, respectively.

For the GRA-SVR and GA-BP model modeling process, more important factors that affect the production of rutting should be considered, so the modeling process is more complex than the other two models, but the prediction results are stable. The PPI model just considers the age and regional conditions, and the main factors affecting the pavement performance were unutilized; therefore, the prediction accuracy is lower. In the GM (1, 1) model, the time factor was only considered, whose prediction accuracy depends greatly on the accuracy of the annual data. If the data of a certain year are deviated, the whole system trend will have a large error, and the ease of operation of the model is between the other models. Therefore, the GRA-SVR model is suitable for multivariate, long-period, and nonlinear prediction of pavement performance.

The accuracy, prediction period, and operability of the three models are compared and analyzed. The results are shown in Table 6.

Overall, our study establishes the model that has offered better performance than other models. However, there are also limitations. In the future study, we want to choose the best parameters with better methods including genetic algorithm and particle swarm optimization. These algorithms are also widely used in other fields. If we find a better optimization method, we can make the prediction accuracy higher. We will build the database with more road information. Then, the GRA-SVR model at the computing terminal is used to predict the performance. Some decision model is applied to maintenance decision. Finally, the results
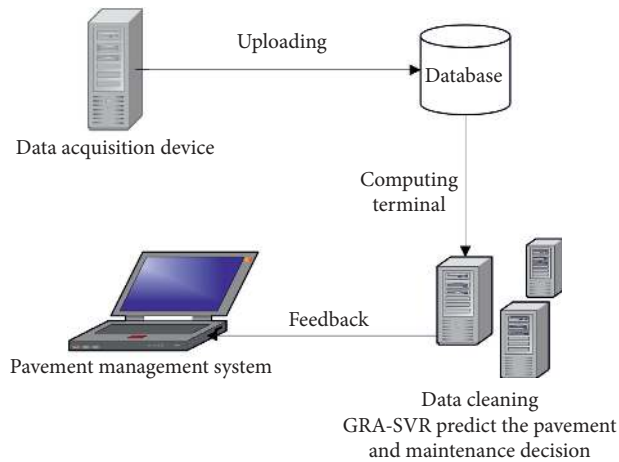
Figure 12: Conception of use of the model.

are uploading the pavement management system (see Figure 12). We firmly believe that this will have far-reaching implications for road maintenance projects.

5. Conclusion

In this study, a GRA-SVR predictive hybrid model, combining the grey correlation analysis with support vector machine regression, was proposed for the first time to be applied to predict the performance of asphalt pavement. The main conclusions are drawn as follows:

(1) The main factors including equivalent single axle loads, maintenance funds, highest temperature in the middle surface, pavement structure strength ratio, average value of soil moisture, highest temperature in the road surface, lowest temperature in the road surface, highest temperature in the upper surface, annual average rainfall, annual cumulative total radiation, highest temperature in the upper surface, annual average rainfall, lowest temperature of upper surface, highest temperature in lower surface, lowest temperature in lower surface, and annual maximum wind speed are well correlated in pavement performance.

(2) Compared with other models, the GRA-SVR model is highly accurate and time-independent, which makes it suitable for short and long period predictions.

In conclusion, the GRA-SVR model is applicable for a multivariate, long period, and nonlinear performance of pavement prediction and is restricted by the amount of data. It is reliable for asphalt pavement maintenance decision-making. At the same time, this model can also be applied to big data road maintenance prediction.

Data Availability

This paper is from the Guangdong Provincial Department of Transportation (2015-02-011), and the data come from the project team experiment.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

References

[1] A. Bianchini and P. Bandini, "Prediction of pavement performance through neuro-fuzzy reasoning," *Computer-Aided Civil And Infrastructure Engineering*, vol. 25, no. 1, pp. 39–54, 2010.

[2] Q. R. Li, Z. Y. Guo, and Y. J. Wang, "Evaluation of the performance of expressway asphalt pavement based on PCA-SVM," *Journal of Beijing University of Technology*, vol. 44, no. 2, pp. 283–288, 2018.

[3] Z. Lan, "Performance evaluation and prediction of expressway asphalt pavement," Southeast University, Nanjing, China, Doctor degree, 2015.

[4] C. Jin and J. X. Zhang, "Summary of research on performance prediction of asphalt," *Journal of China & Foreign Highway*, vol. 37, no. 5, pp. 31–35, 2017.

[5] D. Zhang, X. Li, Y. Zhang, and H. Zhang, "Prediction method of asphalt pavement performance and corrosion based on grey system theory," *International Journal of Corrosion*, vol. 2019, Article ID 2534794, 9 pages, 2019.

[6] D. Shen and J. Du, "Grey model for asphalt pavement performance prediction," in *Proceedings of the Intelligent Transportation Systems Conference*, pp. 668–672, Washington, WA, USA, October 2004.

[7] K. Wang and Q. Li, "Gray clustering-based pavement performance evaluation," *Journal of Transportation Engineering-ASCE - J TRANSP ENG-ASCE*, vol. 136, no. 1, pp. 38–44, 2010.

[8] X. Zhang and C. Ji, "Asphalt pavement roughness prediction based on gray GM (1, 1 | sin) model," *International Journal of Computational Intelligence Systems*, vol. 12, no. 2, pp. 897–902, 2019.

[9] T. Peng, X. L. Wang, and S. F. Chen, "Pavement performance prediction model based on Weibull distribution," *Applied Mechanics and Materials*, vol. 378, pp. 61–64, 2013.

[10] L. J. Sun and X. P. Liu, "Standard decay equation for pavement performance," *Journal of Tongji University (Natural Science)*, vol. 23, no. 5, pp. 512–518, 1995.

[11] A. Abed, N. Thom, and L. Neves, "Probabilistic prediction of asphalt pavement performance," *Road Materials and Pavement Design*, vol. 20, pp. 247–264, 2019.

[12] H. Gong, Y. R. Sun, and B. S. Huang, "Estimating asphalt concrete modulus of existing flexible pavements for mechanistic-empirical rehabilitation analyses," *Journal of Materials in Civil Engineering*, vol. 31, no. 11, Article ID 04019252, 2019.

[13] J. Yang, J. J. Lu, and M. Gunaratne, "Application of neural models for forecasting or pavement crack index and pavement condition rating," in *Gain access to New Resources through the TRB Global Affiliate Program*, Vol. 152, Department of Civil and Environmental Engineering, University of South Florida, Tampa, FL, USA, 2003.

[14] A. Ferreira and R. Lima Cavalcante, "Application of an artificial neural network based tool for prediction of pavement

performance," in *Proceedings of the ISAP Conference on Asphalt Pavements*, Fortaleza, Brazil, 2018.

[15] G. I. Beltran and M. P. Romo, "Assessing artificial neural network performance in estimating the layer properties of pavements," *Ingeniería e Investigación*, vol. 34, no. 2, pp. 11–16, 2014.

[16] J. M. Shen, Y. G. Dong, W. J. Zhou, and X. Wang, "A grey dynamic multi-attribute association decision model based on exponential function," *Control and Decision*, vol. 31, no. 8, pp. 1441–1445, 2016.

[17] X. W. Chen, H. N. Wang, Z. Chen, and Y. Zhan-ping, "Correction of MEPDG rutting prediction model based on mathematical statistics method," *Journal of Chang'an University (Natural Science Edition)*, vol. 33, no. 6, 2013.

[18] C.-Y. Chu and P. L. Durango-Cohen, "Estimation of infrastructure performance models using state-space specifications of time series models," *Transportation Research Part C: Emerging Technologies*, vol. 15, no. 1, pp. 17–32, 2007.

[19] S. M. El-Badawy, M. G. Jeong, and M. El-Basyouny, "Methodology to Predict Alligator Fatigue Cracking Distress Based on Asphalt Concrete Dynamic Modulus," *Transportation Research Record*, vol. 2095, pp. 115–124, 2009.

[20] X. Zhao, Q. Yu, J. Ma, Y. Wu, M. Yu, and Y. Ye, "Development of a representative EV urban driving cycle based on a k-means and SVM hybrid clustering algorithm," *Journal of Advanced Transportation*, vol. 2018, Article ID 1890753, 18 pages, 2018.

[21] N.-D. Hoang, Q. Nguyen, and D. T. Bui, "Image processing-based classification of asphalt pavement cracks using support vector machine optimized by artificial bee colony," *Journal of Computing in Civil Engineering*, vol. 32, no. 5, pp. 1–14, 2018.

[22] X. Wang, N. Zhang, Y. Zhang, and Z. Shi, "Forecasting of short-term metro ridership with support vector machine online model," *Journal of Advanced Transportation*, vol. 2018, Article ID 3189238, 13 pages, 2018.

[23] N. Karballaeezadeh, S. Danial Mohammadzadeh, S. Shamshirband, P. Hajikhodaverdikhan, A. Mosavi, and K.-w. Chau, "Prediction of remaining service life of pavement using an optimized support vector machine (case study of Semnan-Firuzkuh road)," *Engineering Applications of Computational Fluid Mechanics*, vol. 13, no. 1, pp. 188–198, 2019.

[24] M. Dong, "A grey relational analysis between some selected affective factors and English test performance," *Canadian Social Science*, vol. 10, no. 6, pp. 195–200, 2014.

[25] K. J. Chen, X. N. Li, and Y. Y. Qiu, "Gray correlation analysis on influencing factors of engineering material price in Fujian province," *Journal of Highway and Transportation Research and Development*, vol. 35, no. 4, pp. 137–145, 2018.

[26] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.

[27] M. J. Abdi and D. Giveki, "Automatic detection of erythemato-squamous diseases using PSO-SVM based on association rules," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 603–608, 2013.

[28] Z. Liu, H. Cao, X. Chen, Z. He, and Z. Shen, "Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings," *Neurocomputing*, vol. 99, pp. 399–410, 2013.

[29] Q. H. Liu, Z. X. Zhang, H. F. Lin, and Y. Zhu, "Study on prediction of asphalt pavement performance based on support vector machine," *Highway Engineering*, vol. 43, no. 2, pp. 201–205, 2018.

[30] J. P. Yin, "Research on model selection and parameter selection of SVM," Harbin Institute of Technology, Harbin, China, Doctor degree, 2016.

[31] X. Xue and M. Xiao, "Application of genetic algorithm-based support vector machines for prediction of soil liquefaction," *Environmental Earth Sciences*, vol. 75, no. 10, 2016.

[32] S. Abdollahi, H. R. Pourghasemi, G. A. Ghanbarian, and R. Safaeian, "Prioritization of effective factors in the occurrence of land subsidence and its susceptibility mapping using an SVM model and their different kernel functions," *Bulletin of Engineering Geology and the Environment*, vol. 78, no. 6, pp. 4017–4034, 2019.

[33] X. W. Dong, Y. W. Wang, G. S. Zhang, and C. X. Zhou, "The prediction of cross-company software defects based on migration learning," *Computer Engineering and Design*, vol. 37, no. 3, pp. 684–689, 2016.

[34] X. Wang, C. An, Q. Fu et al., "Grey relational analysis and optimization of guide vane for reactor coolant pump in the coasting transient process," *Annals of Nuclear Energy*, vol. 133, pp. 431–440, 2019.

[35] M. Zhang, J. Yi, and D. Feng, "Reasonable thickness design of expressway pavement structures based on gray relation analysis of subgrade soil improvement," *Science Progress*.

[36] I. Aydin, M. Karakose, and E. Akin, "A multi-objective artificial immune algorithm for parameter optimization in support vector machine," *Applied Soft Computing*, vol. 11, no. 12, pp. 204–211, 2011.

[37] X. Wang, Z. Q. Wang, G. Jin, and J. Yang, "Land reserve prediction using different kernel-based support vector regression," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 30, no. 4, pp. 204–211, 2014.

[38] U. Rusmanto, I. Syafi, and D. Handayani, "Structural and functional prediction of pavement condition (A case study on south arterial road, Yogyakarta)," in *Proceeings of the AIP Conference Proceedings*, H. Prasetyo, N. Hidayati, E. Setiawan et al., Eds., American Institute of Physics, Paris, France, June 2018.

[39] C. Jia-Ruey and C. Sao-Jeng, "Development of a rutting prediction model through accelerated pavement testing using group method of data handling (GMDH)," in *Proceedings of the 2009 Fifth International Conference on Natural Computation (ICNC 2009)*, pp. 367–371, Tianjin, China, August 2009.

[40] J. R. Chang, S. H. Chen, D. H. Chen, and Y. B. Liu, "Rutting prediction model developed by genetic programming method through full scale accelerated pavement testing," in *Proceedings of the 2008 Fourth International Conference on Natural Computation*, M. Z. Guo, L. Zhao, and L. P. Wang, Eds., IEEE Computer Society, p. 326, Jinan, China, October 2008.

[41] AASHTO guide for design of pavement structures, *AASHTO Guide for Design of Pavement Structures*, The American Association of State Highway and Transportation Officials, Washington, DC, USA, 1993.

[42] Highway Performance Assessment Standards, *Highway Performance Assessment Standards*, Ministry of Transport of the People's Republic, Beijing, China, 2018.

[43] J. L. Deng, "Introduction to the grey theory," *Grey Systems*, vol. 1, no. 1, pp. 1–24, 1989.

[44] D. Zheng, Z.-D. Qian, Y. Liu, and C.-B. Liu, "Prediction and sensitivity analysis of long-term skid resistance of epoxy asphalt mixture based on GA-BP neural network," *Construction And Building Materials*, vol. 158, no. 15, pp. 614–623, 2018.