

Research Article

A Combined Prediction Model for Subgrade Settlement Based on Improved Set Pair Analysis

Yafeng Li,¹ Wuming Leng,^{1,2} Rusong Nie ,^{1,2} Huihao Mei,^{1,3} Siwei Zhou,¹ and Junli Dong¹

¹School of Civil Engineering, Central South University, Changsha, Hunan 410075, China

²National Engineering Laboratory for High Speed Railway Construction, Central South University, Changsha 410075, China

³Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50010, USA

Correspondence should be addressed to Rusong Nie; nierusong97@csu.edu.cn

Received 4 January 2019; Accepted 1 April 2019; Published 18 April 2019

Academic Editor: Giuseppe D'Aniello

Copyright © 2019 Yafeng Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prediction of subgrade settlement is a complex problem involving various uncertainty factors. To overcome the defects and limitations of the single prediction model, a combined prediction model based on the improved set pair analysis was proposed to take into account the uncertainty and certainty of the single prediction model and make the combined prediction based on the certainty degree, and the criterion of set pair relationship was optimized. In the model, the set pair was first constructed to express the relationship between predicted and measured values. Then the risk of the set pair relationship identification was expressed based on the Bayesian decision theory, and the optimal criterion of set pair relationship was obtained by the adaptive search algorithm. Next, the relationship between the prediction model and the measured data was analyzed to get the certainty degree, and the weight coefficient was obtained according to the certainty degree. Finally the combination of single prediction models was carried out. A case study and comparison with other methods were conducted to confirm the reliability and validity of the proposed model. The result shows that this model fully considers the uncertainty and certainty of the single prediction model and also extends the method of determining the criterion of set pair relationship, which provides ideas for other combined prediction and evaluation problems.

1. Introduction

China's railway is entering a new stage, whether it is a high-speed railway that aims to improve the speed of operation or a heavy-haul railway that aims to improve the axle load of locomotives and vehicles; it imposes stringent requirements on the stability of subgrade. The postconstruction settlement of railway subgrade directly affects the safety, stability, and operating life of the train; therefore, accurate settlement prediction is important for the safety evaluation, reinforcement, and maintenance of the subgrade. Consequently, accurate prediction of subgrade settlement has been a major concern in engineering research.

Current prediction models for subgrade settlement are mainly divided into single prediction models and combined prediction models. The single prediction models are mainly divided into three categories: (1) curve fitting method, such as exponential curve method, hyperbolic method, Asaoka method, and Poisson method [1–5]; (2) backanalysis method,

such as the iterative backanalysis layer by layer method, the modified layer-summation based on elastoplastic theory, and the viscous-elastic BIOT consolidation FEM based on merchant model [6–9]; (3) system theory method, such as time series analysis, gray theory, and neural network method [10–13]. However, single prediction methods are often unsatisfactory: the parameters in the curve fitting method and the system theory method do not have definite physical meaning, and the two methods analyze the settlement law of subgrade from a specific aspect; therefore, it is inevitable to ignore or lose part of the information contained in the data [14]. For the backanalysis method, the solution of constitutive model parameters requires a large amount of experimental and measured data, and the validity of the data has a significant influence on the accuracy of the method, which limits the application of this method. In summary, there are certain defects and limitations in the single prediction models.

To comprehensively analyze the subgrade settlement and combine the advantages of the single prediction model, the

combined prediction model was introduced [15–17]. The combined prediction model has the advantage of reducing the predictive risk and improving the accuracy, which have been widely applied and developed in the settlement prediction. Leng [18] and Wu [19] proposed a combined prediction model based on the correlation coefficient, reciprocal of variance and fitting error, and introduced *MSE*, *MAE*, and *MAP* indicators to evaluate the prediction effect. Zheng [20] and Zhao [21] proposed a combined prediction model based on the residuals and discussed the criteria for selecting basic prediction models. Different from the methods above, Li [22, 23] thought that the combined prediction based on the squares of the errors or the absolute values of the deviations was not necessarily the optimal combination and proposed the superior combination prediction model with the maximum efficiency.

The weight coefficient is the key of the combined prediction model [16]. In the existing literature, the weights are mostly based on the squares of residuals and dispersion, the validity, and the fitting error. The methods of determining the weights are not uniform. Moreover, those combined prediction models often neglect the relationship between the information contained in each single prediction models in the process of portfolio analysis, which restricts the in-depth mining of data information and fails to make full use of existing information for combination prediction.

Subgrade settlement and its prediction is a complex and uncertain system problem affected by many factors. Set pair analysis is a novel analytical method for systematic problems of uncertainty [24]. It dialectically treats uncertainty and certainty by identity-discrepancy-contrary analysis as a whole, which can comprehensively describe the complexity and uncertainty problem. This theory has been applied in the rating, risk assessment, and correlation analysis in civil engineering [25–27], and also provides new methods and ideas for studying the uncertain system problem of subgrade settlement and its prediction.

The set pair relationship is the key of set pair analysis. In the existing method of determining the identity-discrepancy-contrary relationship, the subjective experience method lacks scientificity and rationality, and the association membership method is difficult to apply. The mean standard deviation method and the trisection method are commonly used to determine the criterion of set pair relationship [28–30], but these methods only segment the sample numerically and do not adequately reflect the information and internal relationship contained in the data; thus, the information contained in single models has not been fully utilized and optimally combined.

The main objective of this paper is to introduce a combined prediction model based on set pair analysis and decision theory for subgrade settlement. In the proposed model, the relationship of information from the each prediction model was fully considered based on set pair analysis, and the criterion of set pair relationship was reasonably determined according to Bayesian decision theory and obtained using self-adaption algorithm. Based on the certainty degree, the weight coefficient was obtained and the information from each prediction model was optimally combined. Moreover,

the feasibility and validity of the proposed method were further discussed by case study taking the subgrade settlement of Shuohuang heavy-haul railway as an example, followed by a comparison with single prediction models and a combined prediction model. The proposed model also provides new ideas for other combination prediction and evaluation problems.

2. Theory

2.1. Set Pair Analysis. Set pair analysis is a theory proposed by scholar Zhao for systematic problems of uncertainty [24, 27, 31]. The theory uses the certainty degree μ to quantitatively describe and analyze the relationship between certainty and uncertainty of systematic problems. The mathematical model corresponding to certainty degree μ is written as follows:

$$\mu = a + bi + cj \quad (1)$$

where μ denotes the quantitative expression for set pair relationship, a , b , and c are identity degree, discrepancy degree, and contrary degree, respectively, and $a + b + c = 1$. Parameter i is the discrepancy coefficient within $[-1, 1]$, and parameter j is the contrary coefficient and generally specified as -1 .

2.2. Bayesian Decision Theory. Bayesian decision theory is a powerful tool that enables decision-making even with limited data, succeeding where traditional statistical methods have failed to capture inherent dynamics. It is to make decisions with the least risk or cost under probabilistic uncertainty. This method has been applied in many research fields, such as image processing and machine learning [32, 33].

For a given object x , let $W = \{w_1, w_2, \dots, w_c\}$ be a finite set of c states, and let $A = \{a_1, a_2, \dots, a_t\}$ be a finite set of t possible actions. Let $p(w_j | x)$ be the conditional probability of an object x being in state w_j given that the object is described by x . Let $\lambda(a_i | w_j)$ denote the loss, or cost, for taking action a_i when the state is w_j . For an object x , suppose action a_i is taken. Since $p(w_j | x)$ is the probability that the true state is w_j given x , the expected loss associated with taking action a_i can be expressed as follows:

$$R(a_i | x) = \sum_{j=1}^c \lambda(a_i | w_j) p(w_j | x) \quad (2)$$

3. Combined Prediction Model Based on Set Pair Analysis

The combination prediction of subgrade settlement is essentially the analysis of relationship between each prediction model. The proposed model uses the certainty degree μ to express the certainty and uncertainty of the prediction information of each model, and the certainty degree is obtained based on the set pairs and the criterion of set pair relationship. Next, set pairs are constructed using the absolute error Δ_j , as shown below.

$S = \{s_1, s_2, \dots, s_D\}$ is the measured settlement of monitoring points, and $S_k = \{s_{k1}, s_{k2}, \dots, s_{kD}\}$ is the predicted settlement of the k th prediction model. Set pairs are constructed using the absolute error Δ_i of the predicted and measured values, and the absolute errors are normalized for subsequent analysis.

$$\Delta_i = |s_{ki} - s_i| \quad (3)$$

$$p_i = \frac{\Delta_i - \Delta_{\min}}{\Delta_{\max} - \Delta_{\min}} \quad (4)$$

The absolute error Δ_i and the normalized value p_i express the accuracy and validity of the predicted value, as well as the set pair relationship between the predicted and measured value. If the error value is less than a limit value T_1 , it is considered that the prediction model has a good prediction at this point, and the predicted value is in the identity relationship with the measured value. And if the error value is greater than a limit value T_2 , then the prediction effect is bad, and the predicted value is in the contrary relationship with the measured value. Otherwise, the prediction effect is in general, and the predicted value and the measured value are in a discrepancy relationship. The total number of samples is H . Based on the set pair relationship, the number of samples in which the relationship between the predicted and measured value is identity, discrepancy, and contrary is ε , ϕ , and φ , respectively, and the set pair relationship between the measured values and the predicted values of the model k can be expressed by the certainty degree μ_k as follows:

$$\mu_k = \frac{\varepsilon}{H} + \frac{\phi}{H}i + \frac{\varphi}{H}j \quad (5)$$

The certainty degree μ_k comprehensively reflects the accuracy and effectiveness of prediction model k . The larger the certainty degree μ_k , the better the prediction effect and the larger the weight coefficient of this prediction model. Therefore the certainty degree μ_k can be the basis for determining the weight. Normalize the certainty degree μ_k to obtain the weights and the combination of single prediction models is given as

$$\omega_k = \frac{\mu_k}{\sum_{k=1}^m \mu_k} \quad (6)$$

$$S_i = \sum_{k=1}^m \omega_k S_{ki} \quad (7)$$

4. Criterion of Set Pair Relationship Based on Bayesian Decision Theory

From the set pair analysis above, the set pair relationship is the key. But for the settlement problem in this paper, the criterion of set pair relationship is unknown. The mean standard deviation method and the trisection method are commonly used for systematic problems of uncertainty where the criterion of set pair relationship is unknown, but the two methods determine the criterion numerically and subjectively and do not consider the distribution and

intrinsic relationship of the sample date. Therefore, this paper introduces a new method to determine the optimal criterion of set pair relationship based on the minimum-risk Bayesian decision theory.

In order to apply Bayesian decision theory, the meaning of normalized absolute error p_i is explained here. The normalized absolute error p_i satisfies the form and meaning of probability in Bayesian theory. The smaller the normalized absolute error, the better the prediction effect and the greater the probability that the predictive value is in the identity relationship with the measured value. Conversely, the greater the error, the worse the prediction and the greater the probability that the predictive value is in the contrary relationship with the measured value. Otherwise, the relationship between the predicted and measured value is discrepancy. Therefore, the normalized absolute error p_i can be used as the subjective probability in Bayesian decision theory for risk assessment and parameter decision-making and finally to determine the criterion of set pair relationship.

For a given object x , let $W = \{w, -w\}$ be the set of two states. Let $A = \{a_I, a_D, a_C\}$ be the set of three possible actions, where a_I , a_D , and a_C , respectively, represent actions that classify the object x into identity, discrepancy, and contrary relations. Let p be the conditional probability of an object x being in state w , and the conditional probability of an object x being in state $-w$ is $1-p$. The expected loss R associated with taking the individual actions and the expected loss R_{tot} of taking actions on all samples can be described as follows:

$$R_I = \lambda_{II} \cdot p + \lambda_{IC} \cdot (1 - p)$$

$$R_D = \lambda_{DI} \cdot p + \lambda_{DC} \cdot (1 - p) \quad (8)$$

$$R_C = \lambda_{CI} \cdot p + \lambda_{CC} \cdot (1 - p)$$

$$R_{tot} = \sum R_I + \sum R_D + \sum R_C \quad (9)$$

where R_I , R_D , and R_C are expected loss by determining object x into identity, discrepancy, and contrary relations. λ_{II} , λ_{DI} , and λ_{CI} denote the loss for taking actions of a_I , a_D , and a_C , respectively, when the object x in fact belongs to w . λ_{IC} , λ_{DC} , and λ_{CC} denote the loss for taking the same actions, respectively, when the object x actually does not belong to w .

When the action taken on the object x matches the actual state, the expected loss is minimal, and the criterion of set pair relationship is the most reasonable. Therefore, the principle of determining the set pair relationship can be expressed as the following minimum-risk decision rules:

$$R_I < R_D$$

$$\text{and } R_I < R_C$$

identity

$$R_D < R_I$$

$$\text{and } R_D < R_C$$

discrepancy

$$\begin{aligned}
& R_C < R_I \\
& \text{and } R_C < R_D \\
& \qquad \qquad \qquad \text{contrary} \\
& \qquad \qquad \qquad (10)
\end{aligned}$$

To simplify the decision rules, we make some discussions about the relationship between loss functions. Consider a special kind of loss functions with $\lambda_{II} < \lambda_{DI} < \lambda_{CI}$ and $\lambda_{CC} < \lambda_{DC} < \lambda_{IC}$. That is, the loss of classifying an object x belonging to w into identity relation is less than to the loss of classifying x into the discrepancy relation, and both of the two losses are strictly less than the loss of classifying x into the contrary relation. And the reverse order of losses is used for classifying an object x that does not belong to w . Then we can obtain the principle of determining the set pair relationship expressed in p , substituting formula (8) into formula (10):

$$\begin{aligned}
& p < \frac{\lambda_{CC} - \lambda_{DC}}{(\lambda_{DI} + \lambda_{CC}) - (\lambda_{DC} + \lambda_{CI})} \\
& \text{and } p < \frac{\lambda_{CC} - \lambda_{IC}}{(\lambda_{CC} + \lambda_{II}) - (\lambda_{IC} + \lambda_{CI})} \\
& \qquad \qquad \qquad \text{identity} \\
& p < \frac{\lambda_{DC} - \lambda_{IC}}{(\lambda_{DC} + \lambda_{II}) - (\lambda_{IC} + \lambda_{DI})} \\
& \text{and } p > \frac{\lambda_{CC} - \lambda_{DC}}{(\lambda_{DI} + \lambda_{CC}) - (\lambda_{DC} + \lambda_{CI})} \\
& \qquad \qquad \qquad \text{discrepancy} \\
& p > \frac{\lambda_{DC} - \lambda_{IC}}{(\lambda_{DC} + \lambda_{II}) - (\lambda_{IC} + \lambda_{DI})} \\
& \text{and } p > \frac{\lambda_{CC} - \lambda_{IC}}{(\lambda_{CC} + \lambda_{II}) - (\lambda_{IC} + \lambda_{CI})} \\
& \qquad \qquad \qquad \text{contrary}
\end{aligned} \tag{11}$$

This principle (11) can be written in simple form as follows:

$$\begin{aligned}
& p < \beta \\
& \text{and } p < \gamma \\
& \qquad \qquad \qquad \text{identity} \\
& p < \alpha \\
& \text{and } p > \beta \\
& \qquad \qquad \qquad \text{discrepancy} \\
& p > \alpha \\
& \text{and } p > \gamma \\
& \qquad \qquad \qquad \text{contrary}
\end{aligned} \tag{12}$$

where

$$\begin{aligned}
\alpha &= \frac{\lambda_{DC} - \lambda_{IC}}{(\lambda_{DC} + \lambda_{II}) - (\lambda_{IC} + \lambda_{DI})} \\
\beta &= \frac{\lambda_{CC} - \lambda_{DC}}{(\lambda_{DI} + \lambda_{CC}) - (\lambda_{DC} + \lambda_{CI})} \\
\gamma &= \frac{\lambda_{CC} - \lambda_{IC}}{(\lambda_{CC} + \lambda_{II}) - (\lambda_{IC} + \lambda_{CI})}
\end{aligned} \tag{13}$$

Furthermore, the cost of classifying an object x into the discrepancy relation is closer to the cost of a correct relation than to the cost of an incorrect relation. $(\lambda_{ID} - \lambda_{DC})(\lambda_{CI} - \lambda_{DI})$ is the product of the differences between the cost of making an incorrect action and cost of classifying an object into the discrepancy relation, and $(\lambda_{DI} - \lambda_{II})(\lambda_{DC} - \lambda_{CC})$ is the product of the differences between the cost of making a correct action and cost of classifying an object into the discrepancy relation. So we can get the following inference:

$$(\lambda_{IC} - \lambda_{DC})(\lambda_{CI} - \lambda_{DI}) > (\lambda_{DI} - \lambda_{II})(\lambda_{DC} - \lambda_{CC}) \tag{14}$$

Based on the relationship between the loss functions above, we can get the relationship between α , β , and γ : $0 \leq \beta < \gamma < \alpha \leq 1$. Furthermore, we can simplify the principle (12) to the principle below. Here, we classify the sample at the critical point as identity or discrepancy relation, not discrepancy relation.

$$\begin{aligned}
p &\leq \beta && \text{identity} \\
\beta &< p < \alpha && \text{discrepancy} \\
p &\geq \alpha && \text{contrary}
\end{aligned} \tag{15}$$

The solution of the minimum value of expected loss R_{tol} involves the six loss functions and parameters α , β , and γ , and the six loss functions can be expressed in terms of the parameters α , β , and γ . So we make the following assumptions and simplifications. In six loss functions, the loss λ_{II} generated by determining the object x belonging to w as the identity relation should be 0, and the loss λ_{CC} generated by determining the object x not belonging to w as the discrepancy relation should also be 0. Let us set λ_{IC} equal to 1, so the six loss functions can be simplified to three as follows:

$$\begin{aligned}
\lambda_{CI} &= \frac{1 - \gamma}{\gamma} \\
\lambda_{DC} &= \frac{\beta(\alpha - \gamma)}{\gamma(\alpha - \beta)} \\
\lambda_{DI} &= \frac{(1 - \alpha)(\gamma - \beta)}{\gamma(\alpha - \beta)}
\end{aligned} \tag{16}$$

Assume that the number of samples which are determined to be identity, discrepancy, and contrary relations is ε , ϕ , and φ , respectively. Substitute formula (16) into formulas

(8) and (9), and the expected loss R_{tol} caused by the actions of all objects can be expressed as follows:

$$\begin{aligned}
 R_{tol} &= \sum_{k=1}^{\varepsilon} R_{Ik} + \sum_{k=1}^{\phi} R_{Dk} + \sum_{k=1}^{\varphi} R_{Ck} \\
 &= \sum_{p \geq \alpha} (1-p) + \sum_{p \leq \beta} \left(\frac{1-\gamma}{\gamma} p \right) \\
 &\quad + \sum_{\beta < p < \alpha} \left[\frac{(1-\alpha)(\gamma-\beta)}{\gamma(\alpha-\beta)} p + \frac{\beta(\alpha-\gamma)}{\gamma(\alpha-\beta)} (1-p) \right]
 \end{aligned} \tag{17}$$

When the expected loss R_{tol} is the minimum, the actions on the set pair relationship determination for all samples are the most appropriate, and the parameter (α, β, γ) is the optimal threshold for determining the criterion of set pair relationship using formula (15).

5. Solution of Optimal Threshold Based on Self-Adaption Algorithm

The minimum of the expected loss function R_{tol} involves three parameters: α , β , and γ . The traditional direct solution method has a large solution space and a heavy calculation. The self-adaption algorithm has the advantages of automatically adjusting the processing parameters and constraints according to the characteristics of the data [34–36]. It is widely used for the solution of optimal values and parameter determination and provides a solution for the criterion of set pair relationship. Since the parameter γ is not involved in the criterion, it only appears in the expected loss function. To simplify the algorithm, the parameter γ is discussed in advance. Without involving professional knowledge, the loss generated by determining the object x that actually belongs to w as not belonging to w is equal to the loss generated by determining the object x that does not actually belong to w as belonging to w , that is, $\lambda_{IC} = \lambda_{CI}$. Then $\gamma = 0.5$ can be obtained from formula (16).

The algorithm is briefly described as follows.

Input. $P = \{p_1, p_2, \dots, p_m\}$

Output. α, β

Step 1. Set the initial value: $\alpha = 0.9$, $\beta = 0.1$, and $\gamma = 0.5$. $I = \{\}$, $D = \{\}$, $C = \{\}$, and $T = I \cup D \cup C$.

Step 2. For $p_1 \in P$, let $T = p_1 \cup T$, and calculate the corresponding expected loss R_{tol} based on formula (17).

Step 3. On the premise of $\alpha \geq \beta$, replace the parameters α, β with p_1 , respectively, and calculate the corresponding expected loss R_{tol} . Compare the expected loss values under different (α, β) conditions, and take the corresponding threshold value (α, β) as the optimal parameter value when the expected loss R_{tol} is the minimum.

Step 4. Continue to perform Step 2 and Step 3 on p_2 until all elements in set P have completed the above steps, and

select the final (α, β) to determine the criterion of set pair relationship.

6. Case Study

In order to verify the validity and feasibility of the proposed model, the settlement data of Shuohuang heavy-hual railway subgrade were used to conduct the analysis. The subgrade is in the transition section of the railway and bridge. The thickness of the ballast on the bridge and the transition section is about 1.45m and 0.5m, respectively. The filler of the subgrade is a silty clay with low liquid limit. The profile diagram of the transition section is shown in Figure 1.

The settlement data obtained is along the longitudinal direction of the transition section. In order to avoid the monitoring data affected by the vibration caused by train passing, only the data of the skylight was selected for analysis. At the same time, the information of the train passing through the monitoring point was recorded.

In this paper, the settlement at different monitoring points shows similar rules, so one of the monitoring points can be selected as a demonstration for model calculation and verification (EL2 was selected in the paper). The settlement of the subgrade and the cumulative load of the passing train are shown in Table 1.

For the existing heavy-haul railway with certain years of operation, the instantaneous settlement and the primary consolidation settlement have been basically completed, and the current settlement is composed of the secondary consolidation settlement and the permanent deformation of subgrade. In this case, the subgrade settlement is very small; therefore, the applicable prediction model for this situation is limited. In this paper, hyperbolic model, exponential curve model, GM (1, 1) model, Poisson model, and neural network model with better prediction effect were selected as basic single prediction models to conduct the combined prediction of subgrade settlement under cumulative train load. Since the neural network method was carried out by numerical calculation software, the detailed formula is not given in Table 2. The basic prediction models are shown in Table 2.

Based on the basic prediction models, the predicted results were obtained as shown in Table 4. Then the absolute errors of the predicted results were calculated and normalized. Then we obtained the set P , which contains the connotation of the set pair relationship and the probability. The expected loss function and the criterion of set pair relationship were constructed using the method in section 4, and the parameters (α, β) were solved using the algorithm in Section 5. The solution of parameters is shown in Figure 2, and the criterion of set pair relationship is determined as formula (18).

$$\begin{aligned}
 p &\leq 0.385 && \textit{identity} \\
 0.385 &< p < 0.633 && \textit{discrepancy} \\
 p &\geq 0.633 && \textit{contrary}
 \end{aligned} \tag{18}$$

It can be seen from the criterion of set pair relationship that the prediction results of the single prediction models

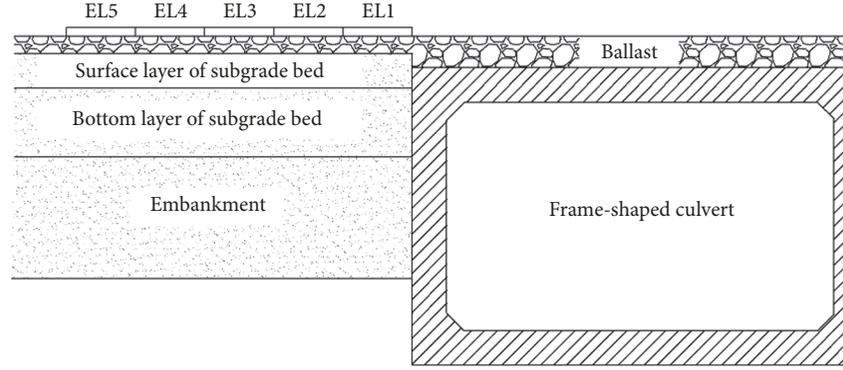


FIGURE 1: Profile diagram of subgrade-bridge transition section.

TABLE 1: Settlement and cumulative load of EL2.

| | | | | | | | |
|--|-------|-------|-------|-------|-------|-------|-------|
| Cumulative load of passing train $Q/\times 10^7 t$ | 7 | 14 | 21 | 28 | 35 | 42 | 49 |
| Measured settlement S/mm | 0.080 | 0.120 | 0.189 | 0.238 | 0.255 | 0.271 | 0.274 |
| Cumulative load of passing train $Q/\times 10^7 t$ | 56 | 63 | 70 | 77 | 84 | 91 | 98 |
| Measured settlement S/mm | 0.297 | 0.304 | 0.307 | 0.327 | 0.331 | 0.332 | 0.335 |

TABLE 2: Single prediction model.

| Prediction model | Regression formula |
|-------------------------|--|
| Hyperbolic model | $S_Q = \frac{Q}{64.32 + 2.27Q}$ |
| Exponential curve model | $S_Q = 0.34(1 - e^{-Q/26.95})$ |
| GM (1, 1) model | $S_k = -0.27e^{-0.24(k-1)} + 0.35$ |
| Poisson model | $S_k = \frac{0.323}{1 + 4.4e^{-0.079k}}$ |
| Neural network method | -- |

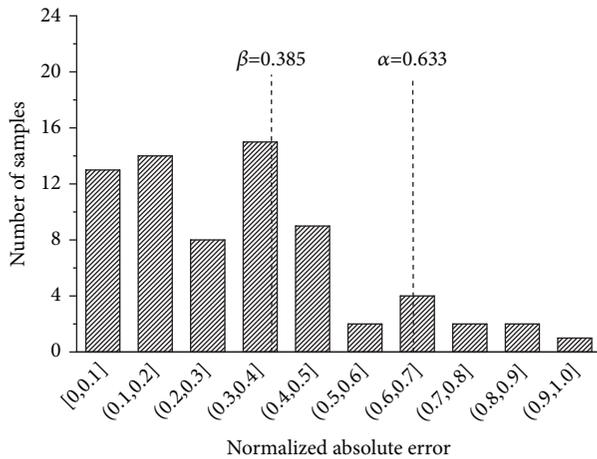


FIGURE 2: Sample information and threshold results.

are mostly in the identity relationship with the measured values, and the prediction effect is good. While a small number of prediction results are in a discrepancy or contrary relationship with the measured values, and the prediction effect is not satisfactory, which lead to the occasional risk and volatility of the single prediction model.

Further, the certainty degree is determined to quantitatively describe the prediction effect and the relationship between the prediction model and the measured data. Count the number of samples in which the set relationship between the predictive and measured values is the identity, discrepancy, and contrary, and the certainty degrees μ_k of three single models are as follows:

$$\begin{aligned}
 \mu_{S-S_1} &= \frac{11}{14} + \frac{1}{14}i + \frac{2}{14}j \\
 \mu_{S-S_2} &= \frac{11}{14} + \frac{1}{14}i + \frac{2}{14}j \\
 \mu_{S-S_3} &= \frac{10}{14} + \frac{2}{14}i + \frac{2}{14}j \\
 \mu_{S-S_4} &= \frac{3}{14} + \frac{9}{14}i + \frac{2}{14}j \\
 \mu_{S-S_5} &= \frac{7}{14} + \frac{6}{14}i + \frac{1}{14}j
 \end{aligned} \tag{19}$$

Under the general conditions of existing research [24–30], the discrepancy coefficient i and contrary coefficient j are usually specified as 0.5 and -1, respectively. And the certainty degree can be further calculated as $\mu_{S-S_1} = 0.679$, $\mu_{S-S_2} = 0.679$, $\mu_{S-S_3} = 0.643$, $\mu_{S-S_4} = 0.393$, and $\mu_{S-S_5} = 0.643$. Normalize the certainty degree, and then obtain the weight coefficient and the combined prediction of subgrade settlement, using formulas (6) and (7). The weight coefficients and combined prediction results are shown in Tables 3 and 4.

In order to compare and analyze the prediction effects of the proposed model, the prediction error was analyzed. The standard error (SE) was selected as the accuracy evaluation index of the prediction model, and the relative standard error (RSE) was selected as the stability evaluation index. The

TABLE 3: Weight coefficient.

| | | | | |
|------------------|-------------------------|-----------------|---------------|----------------------|
| Hyperbolic model | Exponential curve model | GM (1, 1) model | Poisson model | Neural network model |
| 0.224 | 0.224 | 0.212 | 0.129 | 0.212 |

TABLE 4: Prediction results of the proposed model and comparison with those of other models (mm).

| Measured settlement S | Hyperbolic model | Exponential curve model | GM (1, 1) model | Poisson model | Neural network model | Combined prediction model in literature [18] | Proposed model |
|-----------------------|------------------|-------------------------|-----------------|---------------|----------------------|--|----------------|
| 0.080 | 0.087 | 0.078 | 0.080 | 0.091 | 0.090 | 0.081 | 0.082 |
| 0.120 | 0.146 | 0.138 | 0.137 | 0.131 | 0.116 | 0.140 | 0.134 |
| 0.189 | 0.188 | 0.185 | 0.182 | 0.175 | 0.173 | 0.185 | 0.187 |
| 0.238 | 0.219 | 0.221 | 0.217 | 0.217 | 0.234 | 0.219 | 0.226 |
| 0.255 | 0.244 | 0.249 | 0.245 | 0.252 | 0.247 | 0.246 | 0.249 |
| 0.271 | 0.263 | 0.270 | 0.267 | 0.278 | 0.258 | 0.267 | 0.268 |
| 0.274 | 0.279 | 0.286 | 0.284 | 0.296 | 0.262 | 0.283 | 0.279 |
| 0.297 | 0.293 | 0.299 | 0.297 | 0.307 | 0.308 | 0.297 | 0.296 |
| 0.304 | 0.303 | 0.309 | 0.308 | 0.313 | 0.314 | 0.307 | 0.305 |
| 0.307 | 0.314 | 0.316 | 0.316 | 0.317 | 0.316 | 0.315 | 0.312 |
| 0.327 | 0.322 | 0.322 | 0.323 | 0.320 | 0.321 | 0.322 | 0.324 |
| 0.331 | 0.330 | 0.327 | 0.328 | 0.321 | 0.323 | 0.328 | 0.329 |
| 0.332 | 0.336 | 0.330 | 0.332 | 0.322 | 0.323 | 0.332 | 0.333 |
| 0.335 | 0.342 | 0.333 | 0.335 | 0.322 | 0.324 | 0.336 | 0.337 |

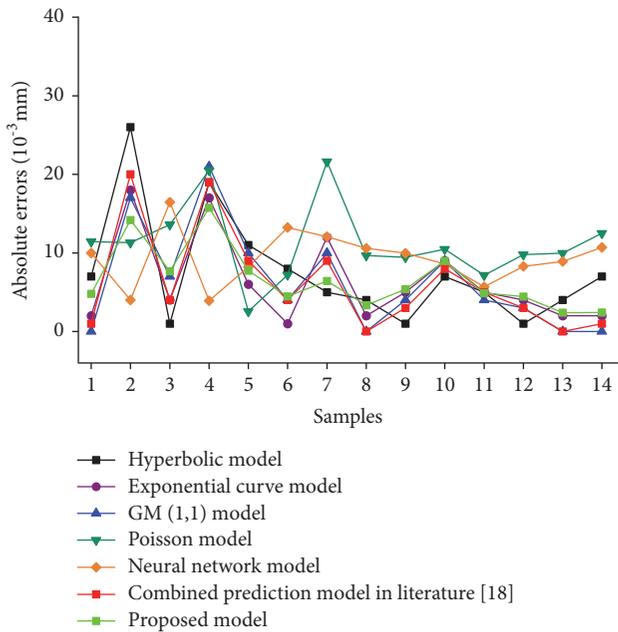


FIGURE 3: Errors of prediction models.

proposed model and the comparison with other prediction models are shown in Figure 3 and Table 5.

The single prediction models have good prediction effect for most samples, but there are some volatility overall. By contrast, the two combined prediction models show better

prediction effect; they have significant advantages in the comprehensive utilization of information and the improvement of prediction stability and accuracy. Furthermore, compared with the combined prediction model in literature [18], the stability and accuracy of the proposed model in this paper are higher, and the prediction effect is better.

7. Conclusions

The current combination prediction models for subgrade settlement neglect the relationship between the information contained in single prediction models and fail to make full use of existing information for combined prediction. Subgrade settlement and its prediction are a complex and uncertain system problem affected by many factors, and set pair theory is a novel analytical method for systematic problems of uncertainty. Therefore, a combination prediction model based on set pair analysis was proposed in this paper. Furthermore, Bayesian decision theory is integrated into the set pair analysis to intelligently determine the criterion of set pair relationship. Finally, the information contained in single prediction models is optimally utilized and combined. With the case study and the analysis of prediction effects, the following conclusions are obtained.

Based on Bayesian decision theory, the criterion of set pair relationship is determined, which overcomes the subjective shortcoming of current methods. Furthermore, the self-adaptation algorithm is used to solve the optimal threshold, which improves the effective recognition of the boundary samples and the accuracy of the set pair relationship.

TABLE 5: Effects of prediction models.

| Prediction model | SE | RSE |
|--|--------|--------|
| Hyperbolic model | 0.0102 | 0.0601 |
| Exponential curve model | 0.0083 | 0.0447 |
| GM (1, 1) model | 0.0089 | 0.0467 |
| Poisson model | 0.0122 | 0.0603 |
| Neural network model | 0.0099 | 0.0497 |
| Combined prediction model in literature [18] | 0.0087 | 0.0480 |
| Proposed model | 0.0077 | 0.0420 |

The certainty degree expresses the certainty and uncertainty relationships between predicted and measured values and describes the accuracy and reliability of the predictive model. Therefore, the certainty degree was used as the basis for determining the weight coefficient, which provides new standards and ideas for the determination of weights. The case study and comparative analysis with other models show that the combined prediction model based on set pair analysis is effective and feasible for subgrade settlement. In addition, the method of determining the criterion of set pair relationship based on decision-making method provides a new idea for the intelligent identification of set pair relations and also provides theoretical support for the application of set pair analysis.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors' research is supported by the National Natural Science Foundation of China (nos. 51878666 and 51678572) and the Innovation Project of Central South University (Grant no. 2018zzts192).

References

- [1] K. Prakash, N. S. Murthy, and A. Sridharan, "Rectangular hyperbola method of consolidation analysis," *Géotechnique*, vol. 37, no. 3, pp. 355–368, 1987.
- [2] A. Asaoka, "Observational procedure of settlement prediction," *Soils and Foundations*, vol. 18, no. 4, pp. 87–101, 1978.
- [3] P. F. Verhulst, "Notice sur la loi que la population suit dans son accroissement," *Quetelet*, vol. 10, no. 10, pp. 113–121, 1837.
- [4] S.-X. Chen, X.-Y. Wang, X.-C. Xu, F. Yu, and S.-L. Qin, "Three-point modified exponential curve method for predicting subgrade settlements," *Rock and Soil Mechanics*, vol. 32, no. 11, pp. 3355–3360, 2011.
- [5] C. F. Huang, Q. Li, S. C. Wu, J. Y. Li, and X. L. Xu, "Application of the Richards model for settlement prediction based on a bidirectional difference-weighted least-squares method," *Arabian Journal for Science & Engineering*, vol. 43, no. 2, pp. 1–9, 2017.
- [6] S. N. M. Tafreshi, T. Shaghghi, G. T. Mehrjardi, A. R. Dawson, and M. Ghadrdan, "A simplified method for predicting the settlement of circular footings on multi-layered geocell-reinforced non-cohesive soils," *Geotextiles and Geomembranes*, vol. 43, no. 4, pp. 332–344, 2015.
- [7] M. Ge, N. Li, W. Zhang, J. Zheng, and C. Zhu, "Settlement behavior and inverse prediction of post-construction settlement of high filled loess embankment," *Yanshilixue Yu Gongcheng Xuebao/Chinese Journal of Rock Mechanics and Engineering*, vol. 36, no. 3, pp. 745–753, 2017.
- [8] C. M. Tan and S. B. Huang, "Back-analysis and prediction for settlement of two-dimensional nonlinear viscoelastic soft clay ground," *Rock and Soil Mechanics*, vol. 23, no. 1, pp. 67–71, 2002.
- [9] S. M. He, "Modified layer-summation based on elastoplastic theory," *Rock and Soil Mechanics*, vol. 24, no. 1, pp. 88–92, 2003.
- [10] H. B. Liu, Y. M. Xiang, and Y. X. Ruan, "A multivariable grey model based on optimized background value and its application to subgrade settlement prediction," *Applied Mechanics and Materials*, vol. 34, no. 1, pp. 173–181, 2013.
- [11] P.-Y. Chen and H.-M. Yu, "Foundation settlement prediction based on a novel NGM model," *Mathematical Problems in Engineering*, vol. 2014, Article ID 242809, 8 pages, 2014.
- [12] X.-W. Ren, Y.-Q. Tang, J. Li, and Q. Yang, "A prediction method using grey model for cumulative plastic deformation under cyclic loads," *Natural Hazards*, vol. 64, no. 1, pp. 441–457, 2012.
- [13] H. Gao, Q.-C. Song, and J. Huang, "Subgrade settlement prediction based on least square support vector regression and real-coded quantum evolutionary algorithm," *International Journal of Grid and Distributed Computing*, vol. 9, no. 7, pp. 83–90, 2016.
- [14] L.-Y. Pan and X.-Y. Xie, "Observational settlement prediction by curve fitting methods," *Yantu Lixue/Rock and Soil Mechanics*, vol. 25, no. 7, pp. 1053–1058, 2004.
- [15] J. N. Bates and C. W. J. Granger, "Combined forecasting," *Journal of Operational Research*, vol. 20, no. 4, pp. 451–468, 1969.
- [16] M. P. Clements and D. I. Harvey, "Combining probability forecasts," *International Journal of Forecasting*, vol. 27, no. 2, pp. 208–223, 2011.
- [17] K. F. Lam, H. W. Mui, and H. K. Yuen, "A note on minimizing absolute percentage error in combined forecasts," *Computers & Operations Research*, vol. 28, no. 11, pp. 1141–1147, 2001.
- [18] W.-M. Leng, Q. Yang, R.-S. Nie, and J. Yue, "Study of post-construction settlement combination forecast method of high-speed railway bridge pile foundation," *Yantu Lixue/Rock and Soil Mechanics*, vol. 32, no. 11, pp. 3341–3348, 2011.
- [19] Q. H. Wu and H. F. Li, "Application of the changeable weight combination model in building settlement predicting," *Journal*

- of Geomatics Science and Technology*, vol. 26, no. 2, pp. 118–120, 2009.
- [20] J.-G. Zheng, T. Wang, and J.-W. Zhang, “Study of settlement prediction methods of loess subgrade,” *Yantu Lixue/Rock and Soil Mechanics*, vol. 31, no. 1, pp. 321–326, 2010.
- [21] M.-H. Zhao, Y. Liu, and W.-G. Cao, “Study on variable-weight combination forecasting method of S-type curves for soft clay embankment settlement,” *Yantu Lixue/Rock and Soil Mechanics*, vol. 26, no. 9, pp. 1443–1447, 2005.
- [22] T. Li, Y.-P. Zhang, and T.-Q. Zhang, “Superior combination forecasting settlement of soft clay embankments,” *Yanshilixue Yu Gongcheng Xuebao/Chinese Journal of Rock Mechanics and Engineering*, vol. 24, no. 18, pp. 3282–3286, 2005.
- [23] H. Y. Chen and D. P. Hou, “Research on superior combination forecasting model based on forecasting effective measure,” *Journal of University of Science and Technology of China*, vol. 32, no. 2, pp. 172–180, 2002.
- [24] K. Q. Zhao, *Set Pair Analysis and Its Preliminary Application*, Zhejiang Science and Technology Press, HangZhou, China, 2000.
- [25] M. Wang and G. Chen, “A novel coupling model for risk analysis of swell and shrinkage of expansive soils,” *Computers & Mathematics with Applications*, vol. 62, no. 7, pp. 2854–2861, 2011.
- [26] M.-W. Wang, P. Xu, J. Li, and K.-Y. Zhao, “A novel set pair analysis method based on variable weights for liquefaction evaluation,” *Natural Hazards*, vol. 70, no. 2, pp. 1527–1534, 2014.
- [27] M. W. Wang, J. L. Jin, and Y. L. Zhou, *Set Pair Analysis Based Coupling Methods and Applications*, Science Press, Beijing, China, 2014.
- [28] W. S. Wang, H. L. Xiang, Y. Q. Li, and S. J. Wang, “A new approach to annual runoff classification based on set pair analysis,” *Journal of Sichuan University*, vol. 40, no. 5, pp. 1–6, 2008.
- [29] T. Yin, X. R. Huang, and W. Wen, “Correlation analysis of encountering probability of runoff in the water-exporting and water-importing area of western route of south-to-north water transfer project by SPA,” *Advanced Engineering Sciences*, vol. 41, no. 2, pp. 48–52, 2009.
- [30] M. W. Wang, D. F. Wei, X. W. Zhou, and P. C. Wang, “Connectional matrix-based combination evaluation method for surrounding rock stability,” *Applied Mathematics & Mechanics*, vol. 36, no. 3, pp. 294–302, 2015.
- [31] H. Garg and K. Kumar, “An advanced study on operations of connection number based on set pair analysis,” *National Academy of Science Letters*, pp. 1–4, 2019.
- [32] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer Science & Business Media, New York, NY, USA, 2013.
- [33] J. V. Stone, *Bayes’ Rule: A Tutorial Introduction to Bayesian Analysis*, Sebtel Press, 2013.
- [34] F. Liang and Y. H. Yan, “Research of dynamic self-adaption intrusion detection model based of clustering,” *Computer Engineering & Design*, vol. 34, no. 3, pp. 814–820, 2013.
- [35] J. He, “An experimental study on the self-adaption mechanism used by evolutionary programming,” *Progress in Natural Science: Materials International*, vol. 18, no. 4, pp. 481–486, 2008.
- [36] G. L. Chen, X. F. Wang, Z. Q. Zhuang, and D. S. Wang, *Genetic Algorithm and Its Application*, Posts & Telecom Press, 1999.

