

Research Article

An Improved KNN-Based Slope Stability Prediction Model

Shuai Huang ¹, Mingming Huang ², and Yuejun Lyu¹

¹*Institute of Crustal Dynamics, China Earthquake Administration, Beijing 100085, China*

²*Beijing Meteorological Information Center, Beijing 100089, China*

Correspondence should be addressed to Mingming Huang; mmhuang1205@163.com

Received 20 April 2020; Revised 27 May 2020; Accepted 17 June 2020; Published 16 July 2020

Academic Editor: Tayfun Dede

Copyright © 2020 Shuai Huang 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.

An accurate slope prediction model is important for slope reinforcement before the disaster. The k -nearest neighbor (KNN) algorithm, as a simple and effective nonparametric machine learning method, is widely applied in classification recognition. In our study, the k -nearest neighbor (KNN) algorithm is improved to reduce its sample dependence and improve the robustness of the algorithm, and then the prediction model of the slope stability is proposed based on the improved k -nearest neighbor (KNN) algorithm. Extensive experimental results show that our proposed prediction model achieves high prediction performance in this regard. Moreover, a comparison between our proposed prediction model and the finite element method, which is the classical theoretical method of slope stability, was made, which will provide an important approach to predicting the slope stability for slope engineering. Finally, shaking table test of a slope model is conducted to evaluate whether the slope is stable or not, and the experimental results are in good agreement with the prediction results of our proposed prediction model, which further demonstrates its effectiveness.

1. Introduction

Landslide is a complex natural phenomenon of slope instability, and it usually causes huge losses to human life and property. It is widely understood that slope stability depends on different parameters, such as cohesion, internal friction angle, rainfall, and earthquake. At present, numerical analysis is commonly adopted in the slope stability analysis. However, numerical analysis will not help analyze slope stability solely because slope is a complex dynamic system affected by many factors [1]. Consequently, the prediction of the slope stability should be of practical significance. The aim of this work is to propose a prediction approach of the slope stability based on machine learning techniques.

Predicting the slope stability is still a challenge. The factors that influence the slope stability are various and complicated, and the main influence factors can be roughly divided into three categories [2], including physical and mechanical properties of the slope soil (unit weight, cohesion, and the angle of internal friction), natural topography (slope height and slope angle), and external factors (rainfall infiltration, groundwater seepage, and earthquake load). It is

difficult to predict the slope stability due to various and complicated factors [3]. Lin et al. [4] chose six typical slope parameters—unit weight, cohesion, internal friction angle, slope inclination, slope height, and pore water ratio—to establish the evaluation index system and predicted the slope stability using four supervised learning methods. Zhao et al. [5] chose six input variables—density, friction angle, friction coefficient, slope angle, slope height, and pore water pressure—for the prediction of slope stability using the relevance vector machine method and found that the RVM is a robust tool for the prediction of slope stability. Samui and Kothari [6] chose six input variables—unit weight, cohesion, angle of internal friction, slope angle, height, and pore water pressure coefficient—for the prediction of slope stability using the least square support vector machine method and found that the developed LSSVM is a robust model for slope stability analysis. Hu et al. [7] used the support vector machine method to forecast the slope instance and found that the forecasting results are consistent with the actual states of slope stability. Li and Jiang [8] chose six characteristic parameters—unit weight, cohesion, angle of internal friction, slope angle, height, and pore water pressure

coefficient—for the prediction of slope stability using KNN and found that the KNN method was more accurate than the backpropagation neural network algorithm. Xiong and Li [9] applied PNN to rock slope stability forecasting, and the results of the case study show that the analysis results are completely consistent with the actual situation. Consequently, machine learning approaches are being increasingly used for slope stability. However, the main external influencing factors of slope stability have been neglected in previous studies, such as rainfall and earthquake, which are the main inducing factors of landslide.

It is worth noting that the k -nearest neighbor (KNN) algorithm [10–12], which is one of the most well-known algorithms in classification recognition, has been proven to be very effective in prediction. To improve prediction accuracy, scholars have carried out studies adjusting its weights [13, 14]. Dudani [15] proposed a weighted voting method named the distance-weighted k -nearest neighbor (WKNN) rule, which is the first distance-based vote weighting schemes. Gou et al. [16] presented a dual weighted k -nearest neighbor (DWKNN) rule that extended the linear mapping of Dudani. Furthermore, because the KNN algorithm generally requires a preset k value and runs multiple experiments with different k values to obtain the best prediction results, some new ideas are proposed to improve the selection of the k value. Zheng [17] proposed a strategy of dynamically setting k values. Liu and Zhang [18] proposed a scheme reconstructing points of the test dataset by learning the correlation matrix, in which different k values are assigned to different points of test data based on the training data. In addition, Ma et al. [19] proposed a coefficient-weighted KNN classifier and a residual-weighted KNN classifier for making classification decisions on the basis of sparse coefficients in the sparse representation. Gou et al. [20] proposed the two-phase probabilistic collaborative representation-based classification (TPCRC) to enhance the power of pattern discrimination in PCRC. Huang et al. [21, 22] analyzed the factors influencing the rockfall runout distance, predicted the rockfall runout distance based on an improved KNN algorithm, and predicted sand liquefaction using the local mean-based pseudo-nearest neighbor algorithm; however, the accuracy of the prediction still needs to be improved. From previous studies, it is found that the KNN algorithm is widely used in classification prediction, but it depends on the number of training samples. In this study, we improved the KNN algorithm to reduce its sample dependence and improve the robustness of the algorithm and built the prediction model of the slope.

2. Establishment of the Prediction Model

2.1. Our Improved KNN Algorithm. KNN, as a simple, effective, and nonparametric prediction method, was first proposed by Cover and Hart to solve text prediction problems [18]. Its principle is to expand the area from the test sample point x constantly until k training sample points are included. In addition, the test sample point x is classified into the category that most frequently appears in the nearest k training sample points.

- (1) KNN algorithm implementation steps are shown in Figure 1.
- (2) Dudani [15] first introduced a weighted voting method for the KNN, called the distance-weighted k -nearest neighbor rule (WKNN). In the WKNN, the closer neighbors are weighted more heavily than the farther ones, using the distance-weighted function. The weighted function of the WKNN is shown as follows:

$$w_i = \begin{cases} \frac{d_k - d_i}{d_k - d_1}, & d_k \neq d_1, \\ 1, & d_k = d_1, \end{cases} \quad (1)$$

$$d_i = \sqrt{(x - x_i^{NN})^T (x - x_i^{NN})}.$$

Accordingly, the prediction result of the query is made by the majority weighted voting as defined in the following:

$$c' = \arg \max_c \sum_{(x_i^{NN}, c_i^{NN}) \in \bar{T}} \delta(c = c_i^{NN}). \quad (2)$$

- (3) DWKNN [16] is based on the WKNN: different weights are given to k -nearest neighbors according to their distances, with closer neighbors having greater weights. The dual distance-weighted function of the DWKNN is defined as

$$w_i = \begin{cases} \frac{d_k - d_i}{d_k - d_1} \cdot \frac{d_k + d_1}{d_k + d_i}, & d_k \neq d_1, \\ 1, & d_k = d_1. \end{cases} \quad (3)$$

Then, we label the query \bar{x} by the majority weighted vote of k -nearest neighbors, the same as

$$c' = \arg \max_c \sum_{(x_i^{NN}, c_i^{NN}) \in \bar{T}} w_i \delta(c = c_i^{NN}). \quad (4)$$

Through comparative study, we find that the method in equation (5) for improvement has better robustness and less sample dependence. Thus, in this study, we used this method to predict the slope stability:

$$w_i = \begin{cases} \ln\left(\left(\frac{d_k - d_i}{d_k - d_1}\right) \cdot \left(\frac{d_k + d_1}{d_k + d_i}\right)\right), & d_k \neq d_1, \\ 1, & d_k = d_1. \end{cases} \quad (5)$$

Accordingly, we classify the query point x into class c by majority weighted voting of its neighbors as shown in the following:

$$c' = \arg \max_c \sum_{(x_j^{NN}, c_j^{NN}) \in \bar{T}} w_j \delta(c = c_j^{NN}). \quad (6)$$

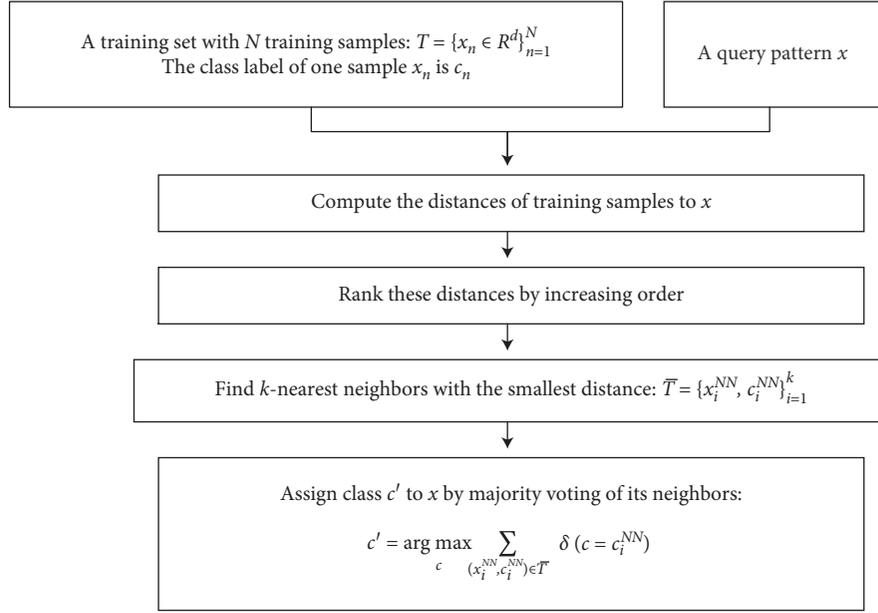


FIGURE 1: KNN algorithm implementation steps.

2.2. *Establishment of the Prediction Model.* The prediction model of the slope stability based on our improved algorithm can be expressed as follows.

Let X denote a set of the slope stability sample, and suppose X is $X = \{x_n \in R^m\}_{n=1}^N$, where x_i represents the feature of the i -th surrounding rock stability sample, N is the total number of features, and m is the feature dimension. In addition, let y_i represent the slope stability levels, and $y_i \in \{0, 1, 2, 3, 4\}$, $i = 1, 2, \dots, N$. Therefore, the sample set of the prediction model is shown as follows:

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ x_N & y_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdot & \cdot & x_{1m} & y_1 \\ x_{21} & x_{22} & \cdot & \cdot & x_{2m} & y_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{N1} & x_{N2} & \cdot & \cdot & x_{Nm} & y_N \end{bmatrix}. \quad (7)$$

Given the unknown sample $x = (x_1, x_2, \dots, x_m)$, our proposed slope stability prediction model based on our improved KNN algorithm can be expressed as

$$\bar{y} = \arg \min_{w_i} f_i(x) = \arg \min_{w_i} d(x, \bar{x}_i^{PNN}), \quad (8)$$

where \bar{x}_i^{PNN} is the nearest neighbors of the unknown sample x in class w_i ($i = 1, 2, 3$). Hence, the unknown sample x is classified into class \bar{y} that has the closest neighbor among all classes.

3. Prediction Model of the Slope Stability Based on Our Improved KNN Algorithm

The prediction model is established using the training samples in [3]. There are 50 cases which are used for training, and 14 cases are used for testing.

3.1. *Data Information and Predictors.* The slope stability prediction is performed to find the nonlinear relationship between the influencing factors and the slope stability. The main influencing factors can be roughly divided into three categories [3], including physical and mechanical properties of the slope soil (unit weight, cohesion, and the angle of internal friction), natural topography of a slope (slope height and slope angle), and external factors (rainfall infiltration, groundwater seepage, and earthquake load). In our study, we chose the representative factors—unit weight, cohesion, internal friction angle, slope height, slope angle, groundwater level, earthquake intensity, and rainfall intensity—as the influencing factors.

By comparison of the slope codes of the earthquake-prone countries (China, Japan, European countries, and the United States), evaluation methods of the slope seismic stability in different specifications were determined at home and abroad, as shown in Table 1.

By comparing slope codes in different countries, we used safety factor and permanent displacement to evaluate the slope stability. According to the safety factor of the slope, Xiong [23] classified the slope stability into five grades which are particular instability, instability, potential instability, basic stability, and stability. The five grades are labeled as I, II, III, IV, and V, respectively, as depicted in Table 2.

3.2. *Normalization.* Since the range of each predictor is significantly different and the test results might rely on the values of a few predictors, they are preprocessed using normalization [24]. We compute the upper and lower bound of each predictor, and the process for the used normalization is represented as

TABLE 1: Slope codes of the earthquake-prone countries.

Nation	Codes	Evaluation methods	Evaluation indicators
China	<Technical code for building slope engineering> (GB50330-2013)	Quasi-static method	Safety factor
	<Design code for engineered slopes in water resources and hydropower projects>(SL386-2007)	Quasi-static method	Safety factor
	<Specifications for design of highway subgrades> (JTG D30-2004)	Quasi-static method	Safety factor
	<Code for seismic design of railway engineering> (GB50111-2006)	Quasi-static method	Safety factor
	<Code for design on subgrade of railway> (TB10001-2005)	Quasi-static method	Safety factor
	<Code for design of high-speed railway> (TB10621-2009)	Quasi-static method	Safety factor
Japan	<Design standards for railway structures>	Quasi-static method and Newmark method	Safety factor and permanent displacement
European countries	Eurocode 7: geotechnical design	Quasi-static method	Safety factor
The United States	California Geological Survey's Guidelines (2008)	Quasi-static method	Safety factor

TABLE 2: Slope stability grade.

Slope stability	Particular instability	Instability	Potential instability	Basic stability	Stability
Grades	I	II	III	IV	V
Grade in samples	0	1	2	3	4

$$\bar{y}_i = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}}, \quad (9)$$

$$y_{\min} = \min\{y\} = \min_{j=1,\dots,n}\{y_j\}, \quad (10)$$

$$y_{\max} = \max\{y\} = \max_{j=1,\dots,n}\{y_j\}, \quad (11)$$

where $y = (y_1, y_2, \dots, y_n)$ is each predictor.

Accordingly, the value of each predictor is normalized to between 0 and 1 based on equations (9)–(11).

3.3. Criteria for Our Prediction Model Performance. The accuracy, computed based on the percentage of all test samples classified correctly, is used to evaluate the prediction performance of the slope stability. Accuracy tells us about the number of samples which are correctly predicted, and it is defined as follows:

$$\text{accuracy} = \frac{\#\text{test samples predicted correctly}}{\#\text{test samples}}, \quad (12)$$

where $\#\text{test samples}$ denotes the total number of test samples and $\#\text{test samples predicted correctly}$ is the number of test samples that are predicted correctly.

3.4. Procedure Algorithm of Our Proposed Prediction Model. In this study, we improve the KNN algorithm to further overcome the influence of neighborhood k . Let $T = \{x_n \in R^d\}_{n=1}^N$ denote a training set with M classes which are w_1, w_2, \dots, w_m . Training samples for each class are $T_{w_i} = \{x_j^i \in R^d\}_{j=1}^{N_i}$, where T_{w_i} is the subset of the training samples w_i , d is the dimensional feature space, and N is the

training samples. In our improved KNN algorithm, the class label of a query point x is computed as shown in the following steps.

For computing the nearest neighbors k from the set T for the unknown query point x , let $\bar{T} = \{x_i^{NN}, c_i^{NN}\}_{i=1}^k$, where \bar{T} is the set of nearest neighbors k for x . And the k -nearest neighbors $x_1^{NN}, x_2^{NN}, \dots, x_k^{NN}$ are sorted in the ascending order according to the distance between their Euclidean distance and x . By assigning different weights to the nearest neighbors k , the weight w_j of the j -th nearest neighbor is defined as

$$w_j = \begin{cases} \exp\left(-\left(\left(\frac{d_k - d_j}{d_k - d_1}\right) \cdot \left(\frac{d_k + d_1}{d_k + d_j}\right)\right)\right), & d_k \neq d_1, \\ 1, & d_k = d_1. \end{cases} \quad (13)$$

Accordingly, we classify the query point x into class \bar{c} by majority weighted voting of its neighbors as shown in the following:

$$\bar{c} = \arg \max_c \sum_{(x_j^{NN}, c_j^{NN}) \in \bar{T}} w_j \cdot \delta(c = c_j^{NN}). \quad (14)$$

3.5. Slope Stability Prediction. In this section, our proposed prediction model is trained by 50 typical slope stability cases and tested by 14 typical slope stability cases. The neighborhood size k ranges from 1 to 7 with an interval of 1, which is inspired by [25]. The 50 typical slope stability cases are shown in Table 3, and the 14 typical slope stability cases are shown in Table 4. This prediction experiment is

TABLE 3: Training samples of the slope stability.

No.	Slope angle (°)	Slope height (m)	Cohesion (MPa)	Internal friction angle (°)	Unit weight (kN/m ³)	Groundwater level (m)	Earthquake intensity	Rainfall intensity (mm)	Grade
1	49	79	0.020	15.0	18.5	80	3	12	I
2	49	79	0.020	20.0	18.0	20	3	52	II
3	48	78	0.045	32.0	17.3	10	3	81	II
4	48	78	0.009	20.4	17.7	50	3	58	IV
5	48	78	0.015	21.0	11.0	10	4	82	IV
6	48	78	0.058	35.0	31.3	20	4	69	V
7	47	77	0.068	37.0	31.3	20	4	72	II
8	46	76	0.063	32.0	25.0	50	4	33	I
9	46	76	0.042	15.0	18.8	10	3	80	II
10	46	75	0.048	40.0	25.0	50	3	31	III
11	45	75	0.017	37.5	22.0	20	4	72	II
12	45	75	0.055	36.0	25.0	80	4	13	V
13	45	75	0.100	45.0	22.4	80	4	29	IV
14	45	75	0.063	32.0	25.0	100	4	5	IV
15	44	74	0.015	24.5	18.4	10	3	89	II
16	44	74	0.001	20.0	18.8	20	3	65	I
17	43	73	0.032	29.5	19.0	10	4	90	V
18	43	73	0.032	33.0	27.0	20	4	50	IV
19	43	73	0.03	20.0	18.8	100	4	2	III
20	42	72	0.012	30.0	19.6	50	3	29	I
21	42	72	0.004	32.0	19.2	80	3	52	IV
22	42	72	0	35.0	17.5	10	3	86	V
23	41	71	0.014	31.0	27.3	10	3	70	II
24	41	71	0.014	31.0	27.3	10	4	70	II
25	41	71	0.032	29.7	27.3	80	4	16	IV
26	40	70	0.010	39.0	18.5	80	4	25	III
27	40	70	0.025	20.0	17.6	20	4	63	I
28	40	70	0.005	10.0	21.8	80	4	13	IV
29	39	69	0.09	32.0	27.0	50	4	31	III
30	38	68	0.038	35.0	18.5	80	3	13	III
31	37	67	0.03	15.0	21.6	20	3	65	IV
32	37	67	0.009	30.0	19.2	10	4	88	II
33	37	67	0.023	35.0	17.4	20	4	43	V
34	37	67	0.005	10.0	14.1	100	3	6	III
35	36	66	0	18.2	19.7	20	3	49	IV
36	36	66	0.030	30.0	19.7	20	3	43	V
37	36	66	0.010	27.0	19.7	10	3	16	III
38	36	66	0.020	20.0	20.0	20	4	72	II
39	35	65	0.010	20.0	17.7	20	4	56	IV
40	35	65	0.003	27.6	19.6	20	4	69	IV
41	35	65	0.120	45.0	25.0	10	4	30	III
42	35	65	0.039	38.0	28.4	100	4	5	V
43	32	62	0.020	20.0	20.0	10	3	88	II
44	32	62	0.017	37.5	22.8	20	3	50	IV
45	31	62	0.069	30.0	21.5	50	4	45	IV
46	31	61	0.040	30.0	20.0	0	3	16	III
47	31	61	0.069	30.0	21.5	50	4	45	I
48	30	60	0.017	37.5	22.8	50	3	37	III
49	30	60	0.014	26	11.9	50	4	44	IV
50	30	60	0.011	0	16.5	10	4	70	I

implemented in Eclipse 3.7.2 by Java language programming, and the hardware environment is Inter Core i7-6700 CPU 3.40 GHz.

As shown in Table 4, our proposed prediction model has high accuracy and reliability, and the prediction results of the proposed prediction model are in good agreement with the actual results. The accuracy of our proposed prediction

model is up to 92.85%. This illustrates that our proposed prediction model is feasible to predict the slope stability, which shows that our proposed prediction model could be used to evaluate the slope stability before the design and construction of slope engineering.

Next, the prediction performance of our proposed prediction model is compared with other prediction models

TABLE 4: Test samples of the slope stability.

No.	Slope angle (°)	Slope height (m)	Cohesion (MPa)	Internal friction angle (°)	Unit weight (kN/m ³)	Groundwater level (m)	Earthquake intensity	Rainfall intensity (mm)	Grade	Prediction results
1	29	60	0	23.5	18.7	50	4	30	III	III
2	29	59	0.023	25.0	12.0	100	4	3	V	V
3	28	59	0.001	32.5	20.4	100	3	6	III	III
4	28	58	0.006	18.0	18.1	10	4	89	IV	IV
5	28	58	0.009	32.0	21.8	10	4	60	IV	IV
6	27	57	0.012	16.3	17.0	20	3	58	I	I
7	27	57	0.008	26.0	17.6	80	4	57	II	II
8	25	56	0.020	15.0	18.5	10	4	57	IV	IV
9	25	55	0.015	30.0	18.8	50	4	56	IV	IV
10	21	52	0.017	37.5	22.8	20	3	55	II	II
11	20	50	0.010	37.0	20.0	20	3	52	I	I
12	20	50	0	20.0	23.0	50	4	50	II	II
13	20	49	0	24.5	20.0	80	4	50	III	II (×)
14	20	47	0.014	25.0	18.8	20	4	49	V	V

based on the KNN algorithm [10], WKNN algorithm [15], and DWKNN algorithm [16]. The following prediction experiments will show whether our proposed prediction model will achieve better prediction performance. The comparison results between different prediction models are shown in Figures 2 and 3.

As can be seen in Figures 2 and 3, the prediction accuracy of our proposed prediction model is somewhat better than the prediction accuracy of the prediction models based on KNN, WKNN, and DWKNN algorithms in almost all of the test cases, which shows that our proposed prediction approach performs better than other approaches with the increasing of the neighborhood size k . It can be found that the accuracy of our proposed prediction model is the highest when the neighborhood size k is 4, and our proposed prediction model achieves an accuracy of 92.85%. This result suggests that our proposed prediction model based on the improved KNN algorithm has the robustness to the sensitivity of different choices of the neighborhood size k with a good prediction performance in predicting the slope stability.

4. Engineering Application of Our Proposed Prediction Model

To further determine the performance of our proposed prediction approach based on the improved KNN algorithm in engineering applications, we also conduct experiments to see the prediction performance for evaluating the slope stability along the Sichuan-Tibet railway in China and compared the prediction results with the finite element method and shaking table test results. The intensities of the historical earthquakes were within a radius of 500 km around.

Our research group drilled lots of boreholes in our survey region along the Sichuan-Tibet railway. On the basis of mass borehole data, the values of the influencing factors—unit weight, cohesion, internal friction angle, and groundwater level—are obtained, and we use our proposed prediction approach based on the improved KNN algorithm

to predict the stability of the slope along the Sichuan-Tibet railway. The seismic activity around the Sichuan-Tibet railway is relatively frequent.

4.1. Slope Stability Prediction of the Sichuan-Tibet Railway.

Cutting slopes along the Sichuan-Tibet railway in China are chosen as the research object. We simplified the slope shape, and the simplified slope models are established with finite element software MIDAS GTS NX. Mohr–Coulomb elastoplastic model is used to model the stress-strain behavior of the soil. And the grid size of the finite element model is 0.5 m, as shown in Figure 4. Furthermore, the bottom is set as the fixed boundary, and the left and right are set as viscoelastic artificial boundaries.

In the numerical simulation model, the quality damping coefficient α and the stiffness damping coefficient β are fixed as 0.2 and 0.0019, respectively. So, the damping coefficient of the numerical simulation model is calculated by the Rayleigh damping formula, as illustrated in the following:

$$[C] = \alpha[M] + \beta[K], \quad (15)$$

where α denotes the quality damping coefficient and β is the stiffness damping coefficient. Moreover, the quality damping coefficient and the stiffness damping coefficient are computed by

$$\alpha = \frac{2\omega_i\omega_j(\xi_i\omega_j - \xi_j\omega_i)}{\omega_j^2 - \omega_i^2}, \quad (16)$$

$$\beta = \frac{2(\xi_i\omega_j - \xi_j\omega_i)}{\omega_j^2 - \omega_i^2}, \quad (17)$$

where ω_i is the natural frequency of the first model, ω_j is the natural frequency of the second model, and ξ_i and ξ_j are conventional damping ratios ranging from 2% to 7%.

The engineering geological conditions of the slope along the Sichuan-Tibet railway were investigated, and the influencing factor values were determined based on the indoor

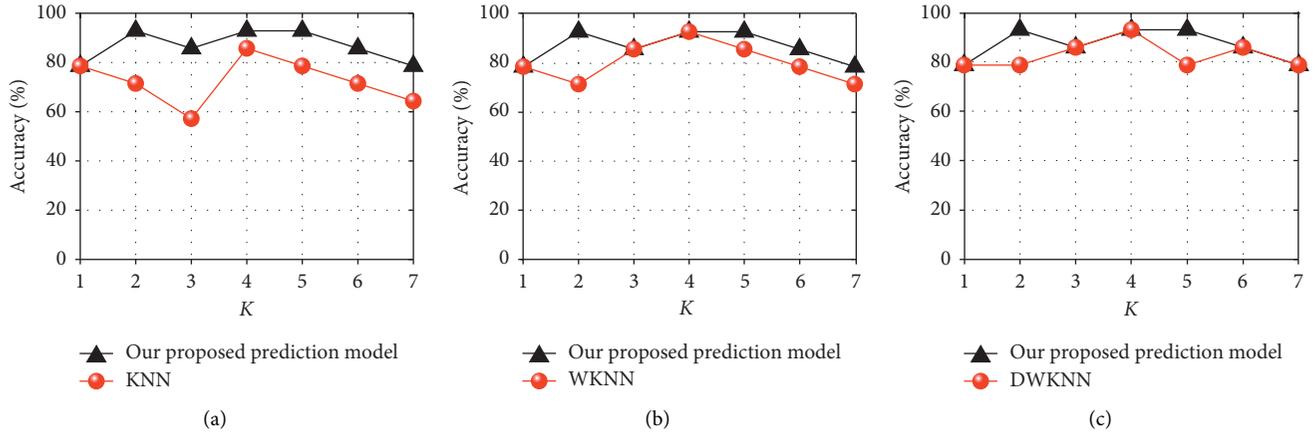


FIGURE 2: Prediction results between our proposed approach with other approaches. (a) KNN algorithm. (b) WKNN algorithm. (c) DWKNN algorithm.

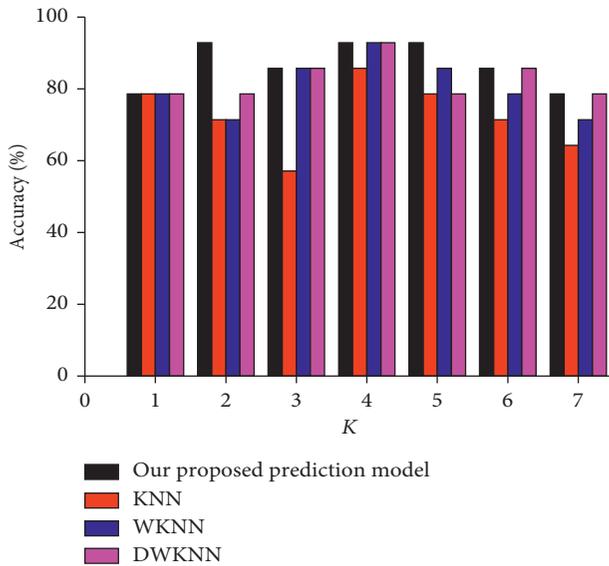


FIGURE 3: Slope stability prediction results with different neighborhood sizes.

experiment. The value ranges of the influencing factors are shown in Table 5.

First, we should know the influence laws of all the factors on slope stability, and we compute the safety factors of the slope under different influencing factors, as shown in Figure 5.

As shown in Figure 5, the safety factors increase with the increasing of the cohesion and internal friction angle, while the safety factors decrease with the increasing of other factors. Also, we found that the slope stability is significantly affected by slope angle, slope height, cohesion, internal friction angle, groundwater level, and peak acceleration.

In order to more directly demonstrate the influence of different influencing factors on slope stability, different influencing factors are normalized. And the safety factors under the normalized influencing factors are shown in Figure 6.

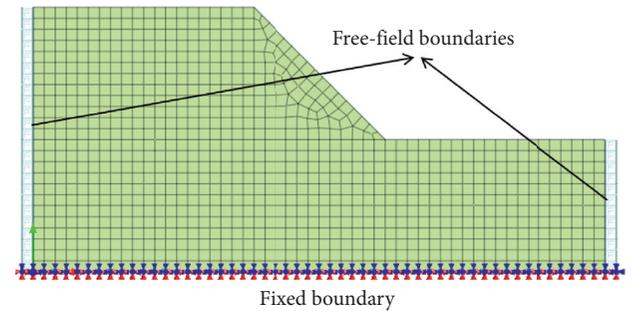


FIGURE 4: Analysis model of the slope.

TABLE 5: Value ranges of the influencing factors.

No.	Influencing factors	Value ranges
1	Slope angle (°)	45~50
2	Slope height (m)	20~32
3	Cohesion (MPa)	0.01~0.2
4	Internal friction angle (°)	15~32
5	Unit weight (kN/m ³)	15~28.5
6	Groundwater level (m)	0~30
7	Earthquake intensity	4~7
8	Rainfall intensity (mm)	28~60

As shown in Figure 7, it can be found that the factors selected in our study are all sensitive to slope stability which shows the accuracy of the influencing factors chosen in our study. Also, we found that the slope is the most unstable under the influence of the peak acceleration, which shows that the impact of potential future earthquakes on the slope cannot be ignored. Consequently, we could use our proposed prediction model to predict the slope stability under the potential future earthquakes, and some reinforcement measures can be taken according to the predicted results, which are important and useful for solving the realistic engineering problems.

Based on the nonlinear finite element method and the strength reduction methods, the slope damage contour can be obtained under different slope stable-states. We could

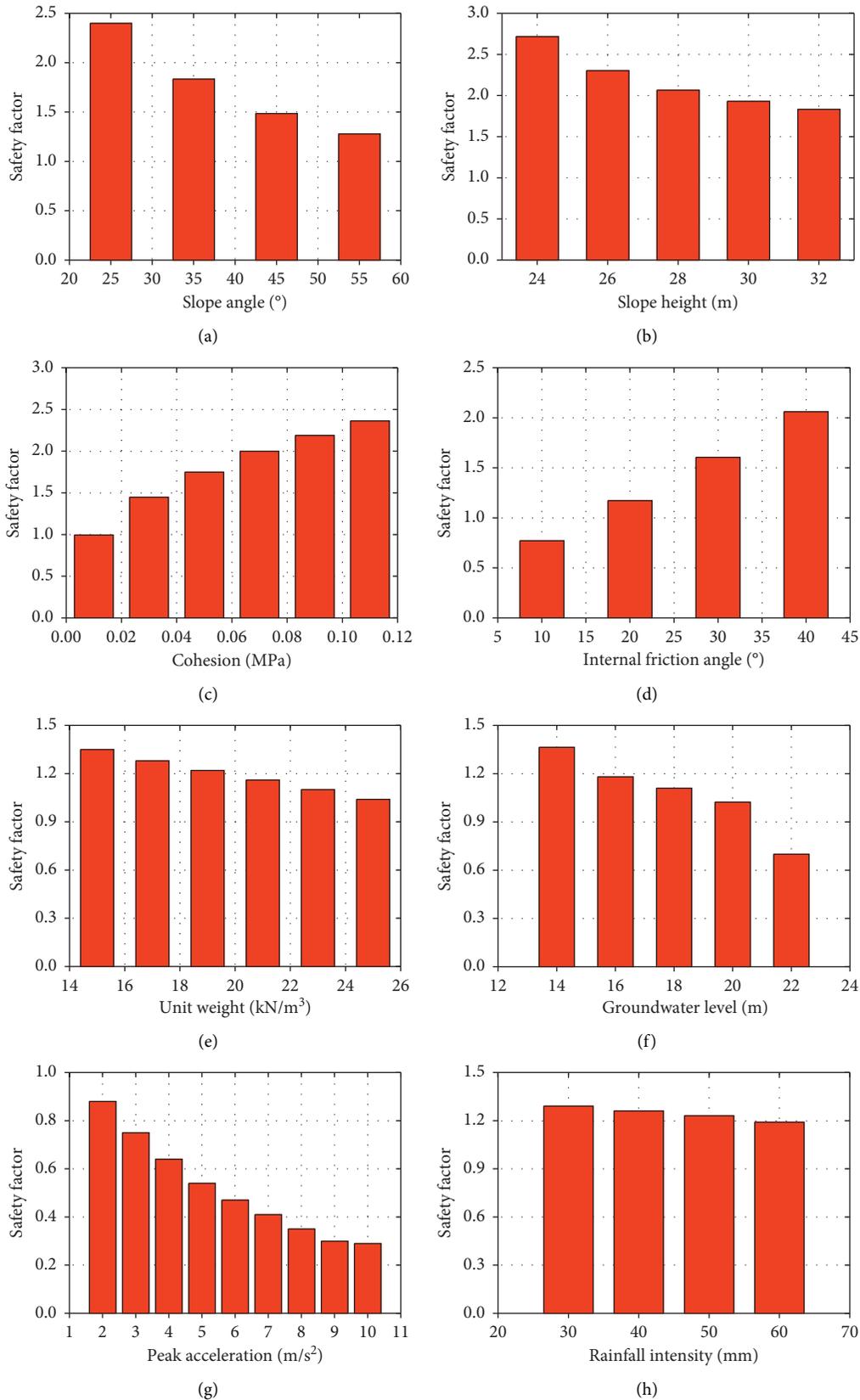


FIGURE 5: Safety factors under different influencing factors. (a) Slope angle. (b) Slope height. (c) Cohesion. (d) Internal friction angle. (e) Unit weight. (f) Groundwater level. (g) Peak acceleration. (h) Rainfall intensity.

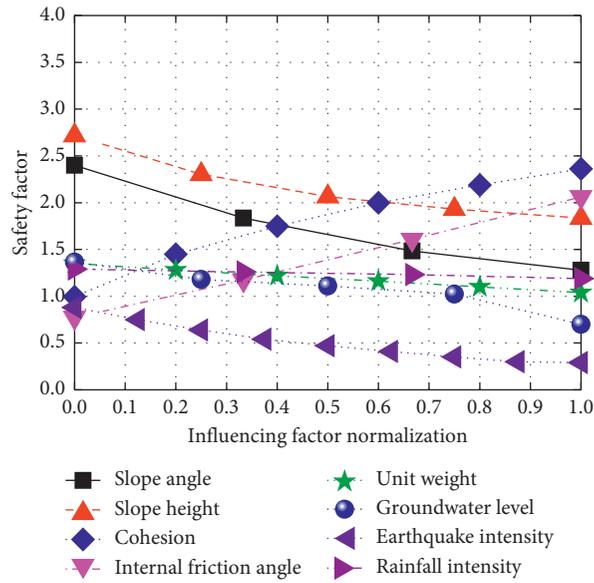


FIGURE 6: Safety factors under different influencing factors after normalization.

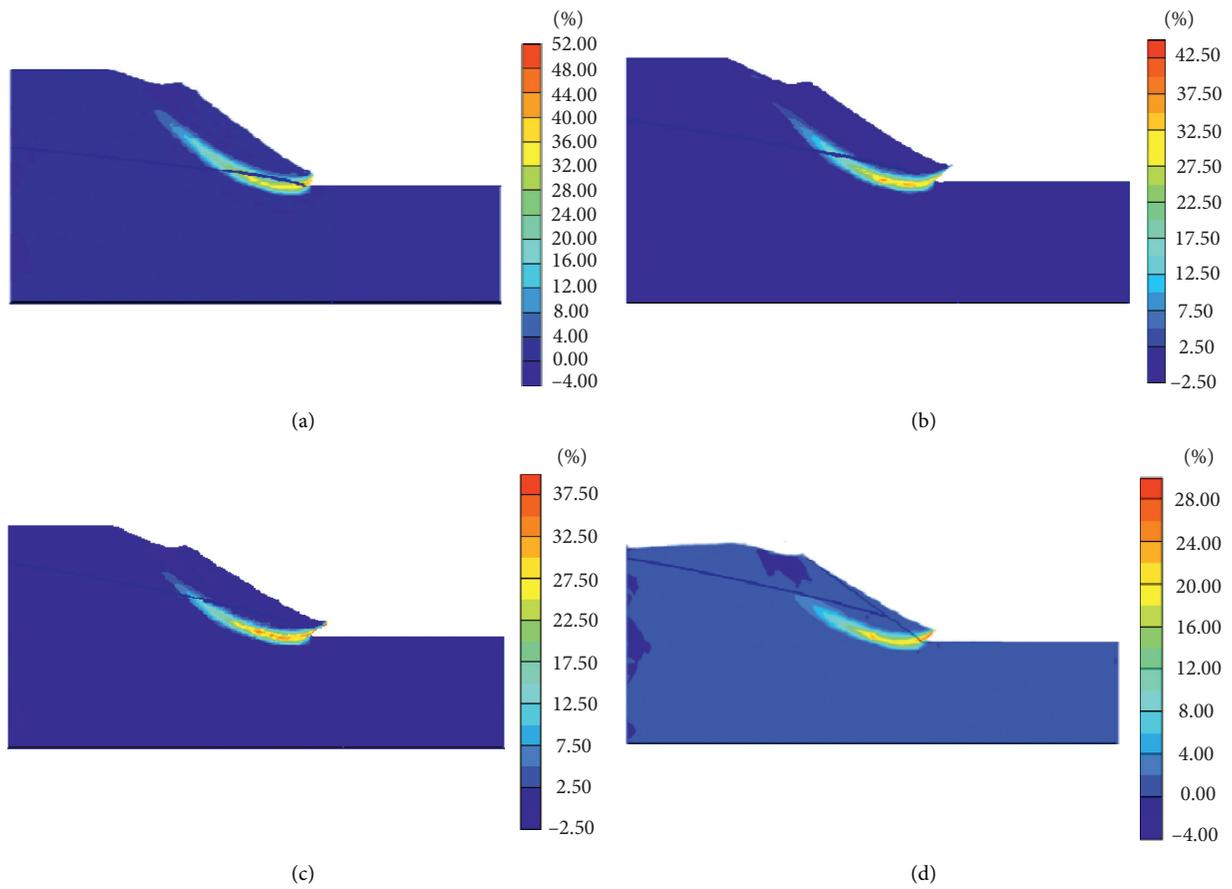


FIGURE 7: Analysis model of the slope. (a) Groundwater level 10 m (grade IV). (b) Groundwater level 15 m (grade III). (c) Groundwater level 20 m (grade II). (d) Groundwater level 25 m (grade I).

TABLE 6: Slope stability prediction along the Sichuan-Tibet railway.

No.	Slope angle (°)	Slope height (m)	Cohesion (MPa)	Internal friction angle (°)	Unit weight (kN/m ³)	Groundwater level (m)	Earthquake intensity	Rainfall intensity (mm)	Our proposed prediction model	Finite element results
1	50	32	0.010	15	28.5	80	6	60	I	I
2	50	31	0.050	20	18.0	20	4	52	III	III
3	49	31	0.145	32	17.3	20	4	31	IV	IV
4	49	31	0.092	20	17.7	50	4	51	III	III
5	49	30	0.065	21	15.0	80	4	32	II	II
6	49	30	0.158	35	21.3	80	4	69	IV	IV
7	48	30	0.068	31	21.3	50	4	52	III	III
8	48	29	0.063	32	22.0	50	6	58	III (×)	I
9	48	29	0.042	25	18.8	20	5	38	II	II
10	48	29	0.011	15	25.0	80	4	50	V	V
11	47	28	0.057	28	21.0	20	5	42	II	II
12	47	27	0.055	26	20.0	80	4	33	II	II
13	47	27	0.152	25	22.4	80	4	29	IV	IV
14	48	28	0.063	22	21.0	50	5	45	II	II
15	47	26	0.200	25	18.4	20	4	49	III	III
16	47	26	0.122	20	18.8	20	4	55	III	III



(a)



(b)

FIGURE 8: Shaking table test equipment. (a) Hydraulic system. (b) Shaking table.



FIGURE 9: Test model of the slope.

more intuitively determine the slope failure degree by the contour. Grade V indicates that the slope is in a stable state; thus, we only plot the slope damage contour for slope stability grades I, II, III, and IV. In this section, the groundwater levels vary, and other factor values remain constant. The slope damage contours for different slope stability grades are shown in Figure 7.

As shown in Figure 7, the slope shows different stability states when the slope is at different groundwater levels. The slope stability degree could be determined through the plastic zone distribution. And we could more intuitively determine the slope stability grades by the finite element method. Thus, we could verify the accuracy of our proposed prediction model by comparing the finite element results.

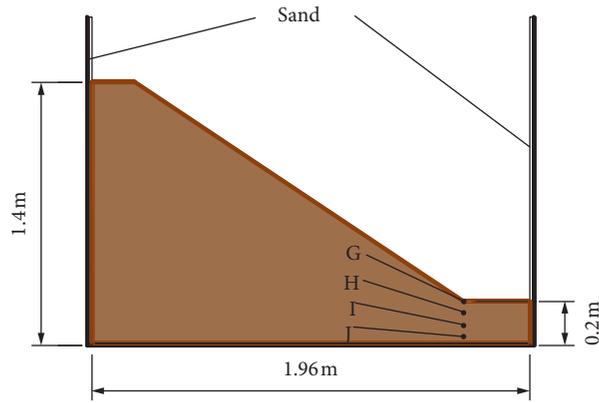


FIGURE 10: Layout of the monitoring sensor.

TABLE 7: Parameters of the earthquake motions.

Seismic waves	Names	Magnitude	Distance to epicenter (km)	Record location
T1-II-1	Hyūganada earthquake (1968)	7.5	100	Foundation of the Itajima bridge
T2-II-1	Kobe (1995)	7.2	16	JR Takatori Station

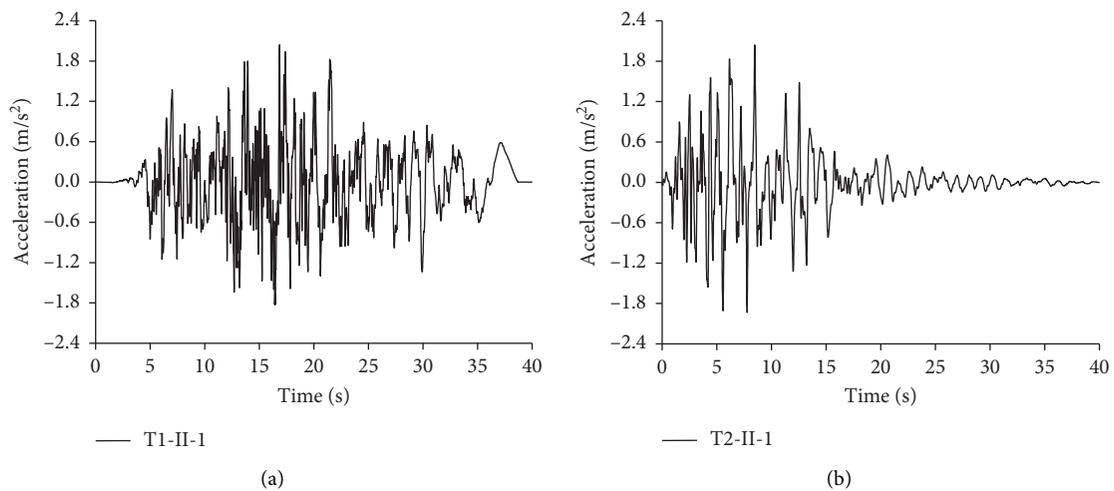


FIGURE 11: Time history of earthquakes. (a) T1-II-1. (b) T2-II-1.

TABLE 8: Values of the influencing factors in the shaking table test.

Test number	Slope angle (°)	Slope height (m)	Cohesion (MPa)	Internal friction angle (°)	Unit weight (kN/m ³)	Groundwater level (m)	Earthquake intensity	Rainfall intensity (mm)
1#	45	40	0.01	35.00	18	20	4	20
2#	45	40	0.01	35.00	18	20	5	20
3#	45	40	0.01	35.00	18	20	6	30
4#	45	40	0.01	35.00	18	20	7	30
5#	45	40	0.01	35.00	18	20	7	30
6#	40	40	0.02	30.00	19	18	7	40
7#	40	45	0.02	30.00	17	16	4	40
8#	40	45	0.02	32.00	18	16	5	40
9#	40	45	0.04	32.00	17	18	6	50
10#	40	45	0.04	38.00	19	16	7	50
11#	50	45	0.04	38.00	19	20	6	50
12#	50	50	0.06	36.00	17	14	6	60
13#	50	50	0.06	36.00	18	14	7	60

TABLE 9: Similarity coefficient between our test model and the actual projects.

Physical quantity	Similarity coefficient	Exegesis
Geometric dimensioning L	C_L	Control variable
Acceleration a	$C_a = C_{[\tau]} C_\rho^{-2/n} C_L^{-1}$	Control variable
Time T	$C_T = C_K^{-1/2} C_\rho^{1/2n} C_L^{2n-1/2n}$	Control variable
Density ρ	$C_\rho = 1$	
Strain level $\gamma/\bar{\gamma}$	$C_{(\gamma/\bar{\gamma})} = 1$	
Dynamic displacement u	$C_u = C_{[\tau]} C_K^{-1} C_\rho^{-1/n} C_L^{n-1/n}$	
Vibration velocity v	$C_v = C_u/C_T = C_{[\tau]} / C_K^{1/2} C_\rho^{-3/2n} C_L^{-1/2n}$	
Vibration frequency w	$C_w = 1/C_T = C_K^{1/2} C_\rho^{-1/2n} C_L^{-1-2n/2n}$	
Damping ratio λ	$C_\lambda = 1$	
Dimensionless index n	$n = 2$	

TABLE 10: Similarity coefficients.

Physical quantity	Similarity constants
Dynamic shear strength	4.85
Length	12
Density	1
Acceleration	0.40
Time	6.45
Strain	1
Dynamic displacement	16.80
Dynamic shear modulus	3.46
Frequency	0.16



FIGURE 12: Simulation of slope groundwater.

To assess the stability of the slope along the Sichuan-Tibet railway, we chose 16 slope cases to predict the slope stability using our proposed prediction model and the finite element method, respectively. And the predicted results obtained by our proposed prediction model and the finite element method are compared as shown in Table 6.

As shown in Table 6, our proposed prediction approach based on the KNN algorithm almost achieves the best performance compared with the finite element method. The prediction accuracy is up to 93.75% which demonstrates that our proposed prediction model could be used for slope stability discrimination for the engineering geological hazard safety assessment.

4.2. Comparison of Our Prediction Model Results with Shaking Table Test Results. In this section, we mainly conduct the effect of the earthquake on slope stability, and other influencing factors remain as a constant value. Thus, we

conduct shaking table test which could reproduce the failure process of the slope under the real earthquakes to see the prediction performance of our proposed prediction model. Figure 8 shows shaking table test equipment to model the slope failure process under the seismic excitation.

As can be seen in Figure 8, the main technical indicators of shaking table test equipment include rated working frequency (40 Hz), the maximum acceleration (20 m/s²), the maximum test load (5000 kg), and dimensions of the shaking table (1.5 m × 1.5 m).

The size of the test model is 1.96 m × 0.96 m × 1.2 m. The slope rate is 1 : 1.5. To keep the sandy soil uniform, the sandy slope is repeatedly stirred. The sponge whose thickness is 20 mm is used to reduce the reflection of seismic waves at the border of the slope. The test model is shown in Figure 9.

The dynamic pore water pressure change of the slope under the earthquake is the main cause of the slope failure. In order to obtain the dynamic pore water pressure, many sensors are deployed at the slope toe. The layout of monitoring points of the test model is shown in Figure 10.

The far-field seismic wave (type I: T1-II-1) and the near-field seismic wave (type II: T2-II-1) are applied to the slope stability analysis. The parameters of the earthquake motions are shown in Table 7, and the acceleration-time histories of the seismic waves are shown in Figure 11. According to the code for seismic design of railway engineering (GB50111-2006) [26] of China, the peak accelerations of the seismic waves are adjusted to 4 degrees, 5 degrees, 6 degrees, and 7 degrees.

In the laboratory test, the effect of different influencing factors on slope stability is investigated. We could determine the damage degree of the slope with different stability levels through the shaking table test. The values of the influencing factors in the shaking table test are shown in Table 8.

The scaling law between our test model and the actual projects follows the Buckingham Pi theorem [27], and the proportional relation for the similarity ratio is developed by Jiang et al. [28]. Poisson's ratio μ of the soil in the test is 0.35, the coefficient $K = \mu/(1 - \mu)$ of the lateral pressure is 0.54, and the dimensionless index $n = 2$. Other similarity coefficients [29] based on the similarity principle are shown in Table 9.

$[\tau]$ is the horizontal shear strength, and $[\tau]$ is computed by

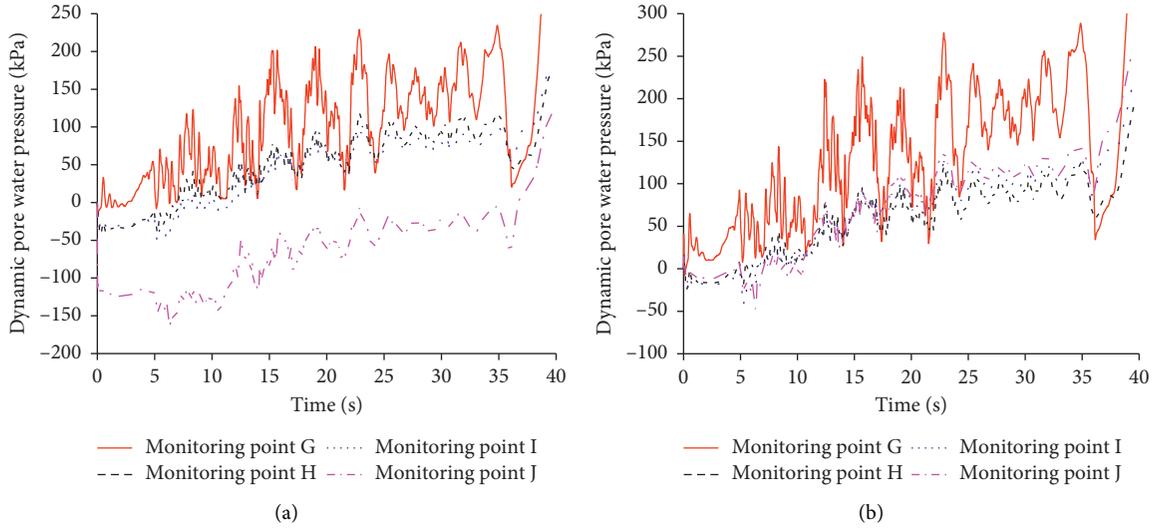


FIGURE 13: The dynamic pore water pressure under T1-II-1. (a) 14 m. (b) 20 m.

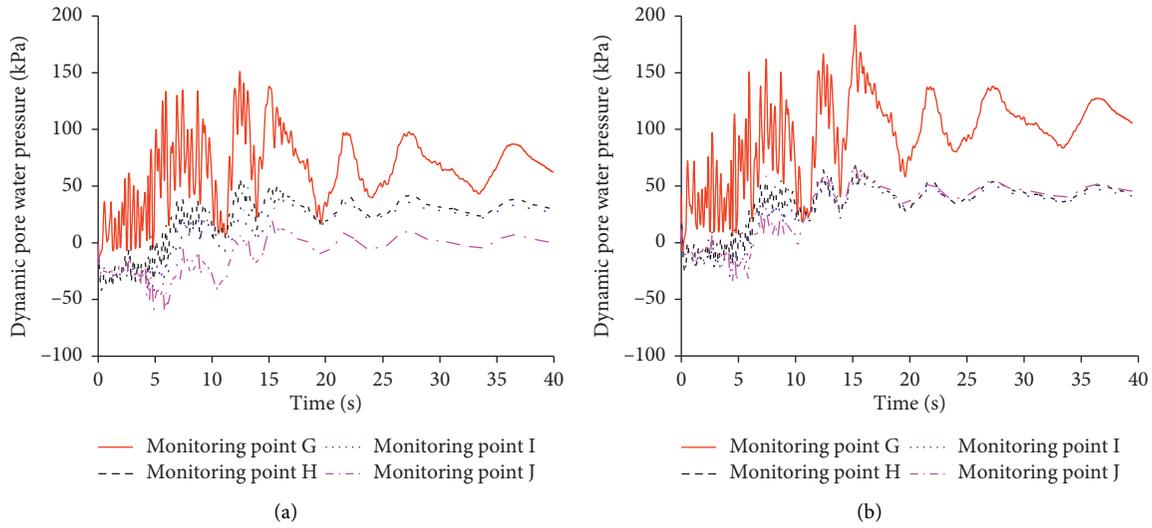


FIGURE 14: The dynamic pore water pressure under T2-II-1. (a) 14 m. (b) 20 m.

$$[\tau] = \sqrt{\left(\frac{1+k}{2}\sigma_v \sin \varphi + c \cos \varphi\right)^2 - \left(\frac{1-k}{2}\sigma_v\right)^2}, \quad (18)$$

where σ_v is the normal pressure stress (i.e., the geostatic stress caused by burial depth), φ is the internal friction angle, c is cohesion, and k is the lateral pressure coefficient.

In our experiment, similar material was developed based on the slope material of the Sichuan-Tibet railway. The similarity coefficients simulating the slope are shown in Table 10.

Simulation of underground water level determines the accuracy of test results. Test program of the simulation of groundwater level is shown in Figure 12.

As shown in Figure 12, water is injected at the left side of the slope. The height of the groundwater level on the right

side is always controlled at the height of the slope toe by turning on the tap. The stable seepage field inside the slope will be formed after a long time seepage of water. Then, the seismic excitation is input into the model.

Before prediction, we analyze the development laws of the dynamic pore water pressure inside the slope, which is a more intuitive indicator of slope failure. The dynamic pore water pressure of the slope at different positions is calculated at the groundwater levels of 14 m and 20 m (before scaling using the equation in Table 10), as shown in Figures 13 and 14.

As shown from Figures 13 and 14, the dynamic pore water pressure rises sharply within a short time under near-field earthquakes and far-field earthquakes. The rising pore water pressure is too late to dissipate and shows large fluctuations. Especially, the dynamic pore water pressure values of the monitoring point G are greater than the dynamic pore water pressure values of other monitoring points. The dynamic pore water pressure of the slope toe is

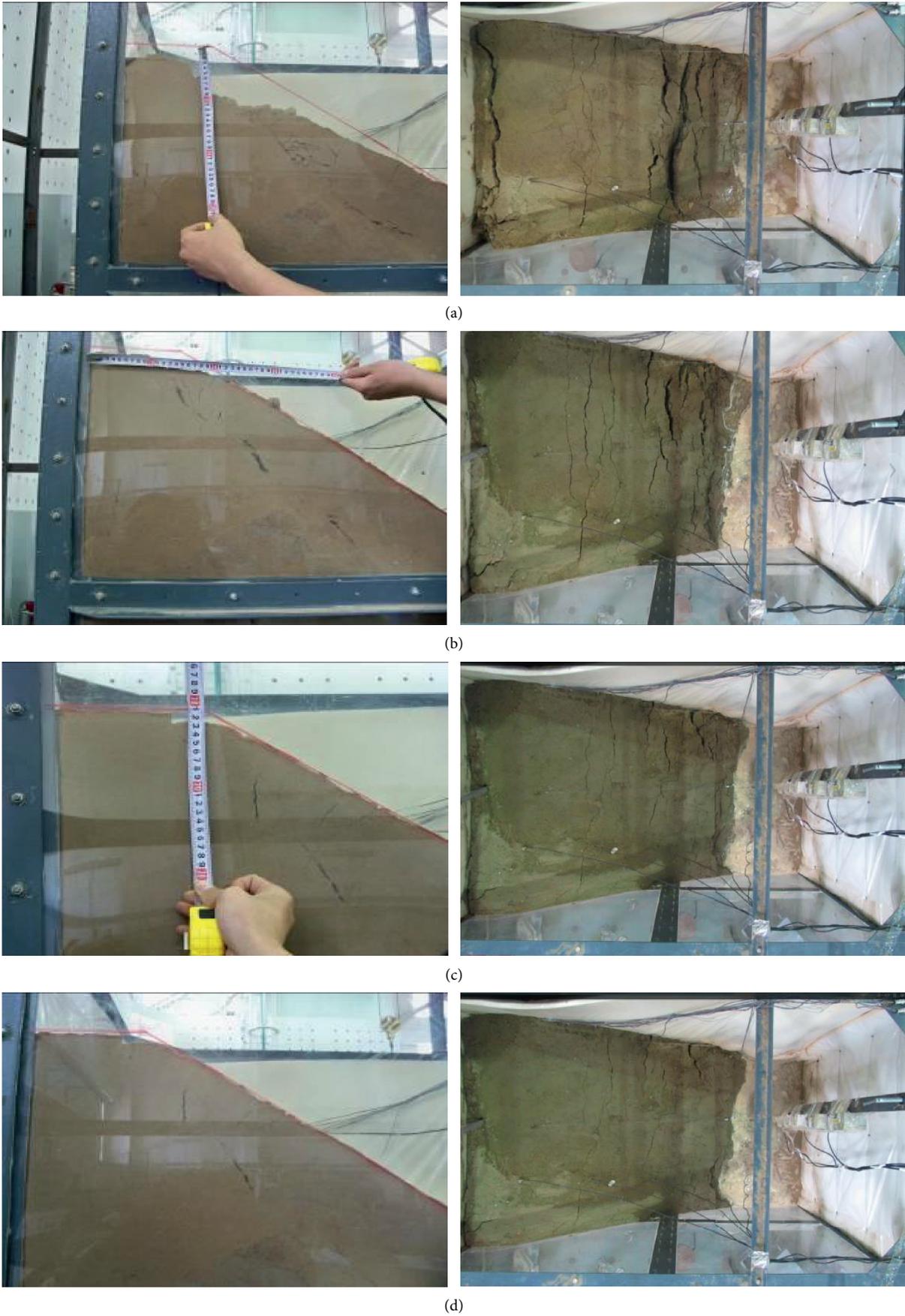


FIGURE 15: Plastic failure states with different stability levels. (a) Grade I. (b) Grade II. (c) Grade III. (d) Grade IV.

TABLE 11: Slope stability prediction results of different methods.

Test number	Shaking table test results	Our proposed prediction model results	Comparison results
1#	V	V	√
2#	IV	IV	√
3#	III	III	√
4#	II	II	√
5#	I	I	√
6#	I	I	√
7#	IV	IV	√
8#	IV	IV	√
9#	III	III	√
10#	II	II	√
11#	III	II	×
12#	III	III	√
13#	II	II	√

greater influenced by the seepage force, which shows that the slope toe is the position which is most easy to have the plastic damage. The production of dynamic pore water pressure under the earthquake could decrease the strength of the soil slope, make the effective stress act on the soil skeleton change, limit its deformation, and cause the destruction of the slope. Shear slide occurs at the slope toe position under the dynamic pore water pressure action; thus, slope toe should be as the key protection position in actual engineering.

Grade V of the slope stability indicates that the slope is in a stable state; thus, we determined the level of the slope damage for I, II, III, and IV grades, as shown in Figure 15.

As can be seen in Figure 15, through the shaking table test, we can more directly determine the failure degree of the slope at different stability grades. The slope slides with the increasing of earthquake intensities, and the destruction begins at the slope toe. The cracks in the slope spread gradually when the slope stability grades vary from IV to I. Meanwhile, the position at the top of the slope has an obvious settlement phenomenon. The sliding surface is approximately circular in shape. And the slope is particular instability when the slope stability grade is I; thus, the shaking table test could accurately determine the slope stability grades.

Then, we compare the shaking table test results with the prediction results of our proposed prediction model, and the comparison results are shown in Table 11.

As can be seen in Table 11, our proposed prediction model based on the improved KNN algorithm achieves the best performance compared with the shaking table test results. The prediction accuracy is up to 92.30% which demonstrates that our proposed prediction model could be used for slope stability prediction before the major project construction near the slope.

5. Conclusions

- (1) We improved the KNN algorithm and established a prediction model of the slope stability. And the performance of our proposed prediction model is evaluated by conducting extensive experiments on slope stability grade prediction, and the experimental results demonstrate the effectiveness of our proposed prediction model.

- (2) We used our proposed prediction model to evaluate the stability of actual slope engineering, and the evaluation results using the finite element method match well with the predicted results of our proposed prediction model, which shows that our proposed prediction approach is an effective method to predict the slope stability.
- (3) The progressive failure process of the slope is conducted by the shaking table test, and the failure degree of the slope at different stability grades is determined. Our proposed prediction model could determine the failure degree of the slope by comparing the experiment result, which further demonstrates the effectiveness of our proposed prediction model of the slope stability.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was financially supported by the research grant from the Institute of Crustal Dynamics, China Earthquake Administration (no. ZDJ2019-10), the National Natural Science Foundation of China (Grant no. 51708516), the Young Elite Scientists Sponsorship Program by CAST (2018QNRC001), the Hebei Province Natural Science Fund (no. E2019210126), and the Fundamental Research Funds for the Central Universities (FRF-BD-19-004A).

References

- [1] J. Li and F. Wang, "Study on the forecasting models of slope stability under data mining," in *Proceedings of the 12th Biennial International Conference on Engineering*, ASCE, Honolulu, HI, USA, pp. 765–776, March 2010.
- [2] C. I. Sachpazis, "Detailed slope stability analysis and assessment of the original carsington earth embankment dam

- failure in the UK,” *Electronic Journal of Geotechnical Engineering*, vol. 21, no. 11, pp. 4247–4273, 2016.
- [3] J. Gao, D. Jinqian, C. Yanxia, and W. Zhigang, “Study on prediction model to estimate stability of high loess slope considering water content,” *Journal of Northwest A&F University*, vol. 35, no. 6, pp. 202–206, 2007.
 - [4] Y. Lin, K. Zhou, and J. Li, “Prediction of slope stability using four supervised learning methods,” *IEEE Access*, vol. 6, pp. 31169–31179, 2018.
 - [5] H. Zhao, S. Yin, and Z. Ru, “Relevance vector machine applied to slope stability analysis,” *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 36, no. 5, pp. 643–652, 2012.
 - [6] P. Samui and D. P. Kothari, “Utilization of a least square support vector machine (LSSVM) for slope stability analysis,” *Scientia Iranica*, vol. 18, no. 1, pp. 53–58, 2011.
 - [7] J. Hu, K. Wang, and J. Dong, “Forecasting of slope stability based on BCABC-SVM,” *Journal of Shenyang University of Technology*, vol. 38, no. 2, pp. 223–227, 2016.
 - [8] C. Li and L. Jiang, “k-nearest neighbor estimation of slope stability,” *Journal of Zhejiang Normal University*, vol. 38, no. 1, pp. 73–77, 2015.
 - [9] J. Xiong and Z. Li, “Research on forecasting of rock slope stability based on IEA-PNN,” *Chinese Journal of Rock Mechanics and Engineering*, vol. 24, no. 1, pp. 4924–4928, 2005.
 - [10] S. Li and S. Yao, “Load prediction of HVAC system based on optimization parameter of phase space reconstruction,” *Refrigeration*, vol. 42, no. 3, pp. 59–64, 2013.
 - [11] J. Gou, W. Qiu, Z. Yi, X. Shen, Y. Zhan, and W. Ou, “Locality constrained representation-based k-nearest neighbor classification,” *Knowledge-Based Systems*, vol. 167, pp. 38–52, 2019.
 - [12] J. Gou, H. Ma, W. Ou, S. Zeng, Y. Rao, and H. Yang, “A generalized mean distance-based k-nearest neighbor classifier,” *Expert Systems with Applications*, vol. 115, pp. 356–372, 2019.
 - [13] L. Dong, G. Geng, and M. Zhou, “Design of auto text categorization classifier based on boosting algorithm,” *Computer Applications*, vol. 27, no. 2, pp. 384–386, 2007.
 - [14] J. Gou, W. Qiu, Y. Zhang, Y. Xu, Q. Mao, and Y. Zhan, “A local mean representation-based k-nearest neighbor classifier,” *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 3, p. 29, 2019.
 - [15] S. A. Dudani, “The distance-weighted k-nearest-neighbor rule,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 6, no. 4, pp. 325–327, 1976.
 - [16] J. Gou, D. Lan, Y. Zhang, and T. Xiong, “A new distance-weighted k-nearest neighbor classifier,” *Journal of Information and Computing Science*, vol. 9, no. 6, pp. 1429–1436, 2012.
 - [17] J. Zheng, *Improvement on feature selection and prediction algorithm for text prediction*, Ph.D. thesis, Xidian University, Xi an, China, 2012.
 - [18] S. Liu and Z. Zhang, “A multi-stage prediction KNN algorithm based on center vector,” *Computer Engineering and Science*, vol. 39, no. 9, pp. 1758–1764, 2017.
 - [19] H. Ma, J. Gou, X. Wang, J. Ke, and S. Zeng, “Sparse coefficient-based k-nearest neighbor classification,” *IEEE Access*, vol. 5, pp. 16618–16634, 2017.
 - [20] J. Gou, L. Wang, B. Hou, J. Lv, Y. Yuan, and Q. Mao, “Two-phase probabilistic collaborative representation-based classification,” *Expert Systems with Applications*, vol. 133, pp. 9–20, 2019.
 - [21] S. Huang, Y. Lyu, Y. Peng, and M. Huang, “Analysis of factors influencing rockfall runout distance and prediction model based on an improved KNN algorithm,” *IEEE Access*, vol. 7, pp. 66739–66752, 2019.
 - [22] S. Huang, M. Huang, and Y. Lyu, “A novel approach for sand liquefaction prediction via local mean-based pseudo nearest neighbor algorithm and its engineering application,” *Advanced Engineering Informatics*, vol. 41, pp. 1–12, 2019.
 - [23] J. Xiong, *The forecasting method of slope stability based on grey-BP neural network*, Ph.D. thesis, Wuhan University of Science and Technology, Wuhan, China, 2012.
 - [24] J.-H. Seo, Y. H. Lee, and Y.-H. Kim, “Feature selection for very short-term heavy rainfall prediction using evolutionary computation,” *Advances in Meteorology*, vol. 2014, Article ID 203545, 15 pages, 2014.
 - [25] M. Huang, R. Lin, S. Huang, and T. Xing, “A novel approach for precipitation forecast via improved k-nearest neighbor algorithm,” *Advanced Engineering Informatics*, vol. 33, pp. 89–95, 2017.
 - [26] Ministry of Railways of the People’s Republic of China, *Code for Seismic Design of Railway Engineering (GB50112-2009)*, China Planning Press, Beijing, China, 2006.
 - [27] B. Louis, “The pi theorem of dimensional analysis,” *Archive for Rational Mechanics and Analysis*, vol. 1, no. 1, pp. 35–45, 1957.
 - [28] L. Jiang, L. Yao, and J. Wang, “Similitude for shaking table model test on side slope relating to dynamic characteristics and strength,” *Journal of Transport Science and Engineering*, vol. 25, no. 2, pp. 1–7, 2009.
 - [29] D. Wei and C. Hu, “Scaling of an impacted reticulated dome using partial similitude method,” *Latin American Journal of Solids and Structures*, vol. 16, no. 2, pp. 1–12, 2019.