Analysis of Factors Influencing Rockfall Runout Distance and Prediction Model Based on an Improved KNN Algorithm

SHUAI HUANG [ID]**[1], YUEJUN LYU[1], (Member, IEEE), YANJU PENG[1], (Fellow, IEEE), AND MINGMING HUANG[2], (Fellow, IEEE)**
[1]Institute of Crustal Dynamics,China Earthquake Administration, Beijing 100085, China
[2]Beijing Meteorological Information Center, Beijing 100089, China

Corresponding author: Shuai Huang (huangshuai3395@163.com)

ABSTRACT The prediction method plays crucial roles in the accurate prediction of rockfall runout range which could improve the protection of endangered residential areas and infrastructure. Recently, the K-nearest neighbor (KNN) algorithm, one of many machine learning techniques, showed good performance in pattern classification. Therefore, the aim of this study was to use the K-nearest neighbor (KNN) algorithm to predict the rockfall runout range which is classified into different subintervals according to the distance from the slope toe. First, we proposed the prediction model of the rockfall runout range based on our improved KNN algorithm which could better offer robustness against different choices of the neighborhood size k, and it is the first work of applying our improved KNN algorithm to rockfall runout range prediction. Second, the shaking table tests of rockfall runout models were conducted for simulating the rockfall process, and the influence laws of factors-including types of an earthquake, peak ground acceleration, vibration frequency, slope angle, slope height, and block mass and block shape-on rockfall distance are investigated. Finally, there is a discussion of the performance of our proposed prediction model based on our improved KNN algorithm in the prediction of rockfall runout range. The extensive experimental results for rockfall runout range prediction demonstrate the effectiveness of our proposed prediction model.

INDEX TERMS Improved KNN algorithm, rockfall runout range, earthquake, shaking table test.

I. INTRODUCTION

Earthquake-induced rockfalls, such as the 1964 Alaska earthquake in the United States, the 1964 Niigata earthquake in Japan, the 1976 Tangshan earthquake and the 2008 Wenchuan earthquake in China, caused the massive destruction of civil structures, which usually resulted in tremendous losses both in lives and properties [1]. To illustrate the rockfall problem, Figure 1 shows a rockfall event after the Wenchuan earthquake which occurred in the Sichuan Province on May 12, 2008. Rockfall hazard mapping requires definition of the runout distance and the range that can be reached by blocks, i.e. the propagation area [2]. Unfortunately, the rockfall runout distance prediction presents a major challenge

due to a large number of influencing factors, which have a highly nonlinear relationship with the rockfall runout distance. Because of the randomness and uncertainties of earthquakes, the prediction of the rockfall runout range should be of practical significance. The aim of this work is to obtain a prediction approach of the rockfall runout range based on machine learning techniques.

Predicting the rockfall runout range, namely, the areas potentially under the threat of rockfall, is still a challenge. Various solutions exist-including · field test, laboratory investigation, and numerical simulation [3]. Preh et al. [4] conducted a drop test, and obtained reasonable stochastic approximations of rock fall trajectories in all dimensions. Jaboyedoff et al. [2] estimated rockfall runout zones using a numerical simulation method, which is proved to be a good way to get first estimations of rockfall runout zones. Hu et al. [5]

The associate editor coordinating the review of this manuscript and approving it for publication was Kathiravan Srinivasan.

FIGURE 1. A rockfall disaster occurred after the Wenchuan earthquake, Sichuan Province, China, May 12, 2008.

proposed formulas for computing the movement distance and bounce height of the falling rock in two-dimensional space. Wang, et al. [6] proposed critical threshold values of rockfall events by probabilistic method. Unfortunately, research on rockfall runout distance and the range upon earthquake events is seldom involved due to the long recurrence period and scare field data of the rockfalls during earthquakes. The shaking table test is a key means of earthquake simulation in the lab [7]. As a consequence, by the shaking table test, we obtained the rockfall data under earthquake in our study. In recent years, the interest in the prediction approach of the rockfall runout distance has grown significantly using various machine learning techniques that yield good predictions [8], [9]. Zhou [1] established the prediction system by a support vector machine method to predict the characteristic parameters of rockfall movement on rock slope, and found that the prediction system based on support vector machine (SVM) can easily estimate the characteristic parameters of rockfall movement without modeling and computing with rockfall software. For this reason, machine learning approaches are being increasingly used for rockfall prediction.

The aim of our study is to investigate rockfall runout range which is essentially a range prediction problem. It is worth noting that the K-nearest neighbor (KNN) algorithm [10]–[12], which is one of the most well-known algorithms in pattern recognition, has been proven to be very effective in classification and prediction. Consequently, the rockfall runout range predictions have been made according to the classification results of the KNN algorithm. A number of variations of the KNN-based approaches have also been developed [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. In this approach, the farthest neighbor is weighted as 0, the closest as 1 and the others are scaled between by a linear mapping. The WKNN algorithm performs well in comparison with the traditional KNN approach, but the classification performance is still impacted by the sensitivity to the choice of the neighborhood size k, particularly in the small sample size cases. Gou et al. [16] presented a dual weighted k-nearest neighbor (DWKNN) rule that extended the linear mapping of Dudani, in which the closest and the farthest neighbors are weighted the same way as the linear mapping, but those between them are assigned smaller values. Although the WKNN and DWKNN algorithms perform

well in comparison with the traditional KNN approach, the sensitivity of the classification performance to the choices of the neighborhood size k still exits. It is also noticed that the exponential of some distance, which is chosen as the weighting scheme, exhibits better classification accuracy and lower variance [17]. Inspired by the effectiveness of the exponential of some distance for classification, we believe that this approach should be a better choice as the weighting scheme. Also, we proposed an improved KNN algorithm from the literature [18] to predict the rainfall grade, and found that the performance of our proposed rainfall grade approach based on our improved KNN algorithm is somewhat better than the other three approaches, namely, DWKNN, WKNN and KNN.

In this paper, we use our improved KNN algorithm to the prediction of the rockfall runout range, which, to our knowledge, is the first study to apply an improved KNN algorithm for the prediction of the rockfall runout range. Extensive experimental results show that our proposed prediction model achieves high prediction performance of the rockfall runout range.

II. OUR IMPROVED KNN ALGORITHM
Our improved KNN algorithm, which exhibits high classification accuracy [18], chose the exponential of some distance as the weighting scheme. Before presenting our improved KNN algorithm, we briefly review of KNN, WKNN and DWKNN algorithm.

A. KNN, WKNN AND DWKNN ALGORITHMS
The k-nearest neighbor algorithm is a powerful nonparametric classifier which assigns an unclassified pattern to the class represented by a majority of its k nearest neighbors. In the general classification problem, let $T = \left\{ x_n \in R^d \right\}_{n=1}^{N}$ denote a training set with M classes including N training samples in d-dimensional feature space. The class label of one sample x_n is c_n. Given a query point x, the KNN rule is carried out as follows.

(1) Find k nearest neighbors from the set T for the unknown query point x, and let $\overline{T} = \left\{ x_i^{NN}, c_i^{NN} \right\}_{i=1}^{k}$ indicate the set of k nearest neighbors for x. The distance between x and the neighbor x_i^{NN} is measured by the Euclidean distance metric, as shown in equation (1).

$$d(x, x_i^{NN}) = \sqrt{(x - x_i^{NN})^T (x - x_i^{NN})} \qquad (1)$$

(2) The class label of the query point x is predicted by the majority voting of its neighbors, as shown in equation (2).

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

where c is a class label and c_i^{NN} is the class label for the i-th nearest neighbor among its k nearest neighbors. $\delta(c = c_i^{NN})$, an indicator function, takes a value of one if the class c_i^{NN} of the neighbor x_i^{NN} is the same as the class c and zero otherwise.

(3) Dudani [15] first introduced a weighted voting method for KNN, calling the distance-weighted k-nearest neighbor rule (WKNN). In WKNN, the closer neighbors are weighted more heavily than the farther ones, using the distance-weighted function. The weighted function of WKNN is shown in Equation (3).

$$w_i = \begin{cases} \dfrac{d_k - d_i}{d_k - d_1}, & d_k \neq d_1 \\ 1, & d_k = d_1 \end{cases}$$
$$d_i = \sqrt{(x - x_i^{NN})^T (x - x_i^{NN})} \quad (3)$$

Accordingly, the classification result of the query is made by the majority weighted voting, as defined in Equation (4).

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

(4) DWKNN is based on 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 DWKNN is defined in Equation (5).

$$w_i = \begin{cases} \dfrac{d_k - d_i}{d_k - d_1} \cdot \dfrac{d_k + d_1}{d_k + d_i}, & d_k \neq d_1 \\ 1, & d_k = d_1 \end{cases} \quad (5)$$

Then, we label the query $\overline{x}$ by the majority weighted vote of k nearest neighbors, the same as Equation (6).

$$c' = \arg\max_c \sum_{(x_i^{NN}, c_i^{NN}) \in \overline{T}} w_i \delta(c = c_i^{NN}) \quad (6)$$

B. OUR IMPROVED KNN ALGORITHM

In the literature [18], we improved the KNN algorithm to further overcome the influence of neighborhood k. Consider a training set $T = \{x_n \in R^d\}_{n=1}^N$ with M classes, where N is the sample numbers of T, and d is the feature dimension. In our proposed improved KNN algorithm, the class label of a query point x is yielded by the following steps.

Find k nearest neighbors for the unknown query point x in the training set T, and let $\overline{T} = \{x_i^{NN}, c_i^{NN}\}_{i=1}^k$ denote the set of k nearest neighbors for x, and the k-nearest neighbors $x_1^{NN}, x_2^{NN}, \cdots\cdots, x_k^{NN}$ are sorted in an increasing order according to their corresponding Euclidean distance to x, described in Equation (1).

Allocate different weights to k nearest neighbors, and the weight w_j of the j-th nearest neighbor is determined as shown in Equation (7).

$$w_j = \begin{cases} \exp(-((\dfrac{d_k - d_j}{d_k - d_1}) \cdot (\dfrac{d_k + d_1}{d_k + d_j}))), & d_k \neq d_1 \\ 1, & d_k = d_1 \end{cases} \quad (7)$$

Accordingly, classify the query point x into the class $\overline{c}$ by a majority weighted voting of its neighbors, as shown in Equation (8).

$$\overline{c} = \arg\max_c \sum_{(x_j^{NN}, c_j^{NN}) \in \overline{T}} w_j \cdot \delta(c = c_j^{NN}) \quad (8)$$

III. ROCKFALL EXPERIMENT

For laboratory study, large natural rockfall is impractical. Consequently, a method that used physical model tests with small-sized rock blocks and artificial slope in combination with shaking table test equipment was applied.

A. EXPERIMENTAL APPARATUS

Natural seismic records can be reproduced by the shaking table test, and the model installed on the shaking table can actually experience the earthquake process. Figure 2 shows the shaking table test equipment to model rockfall under the seismic excitation.
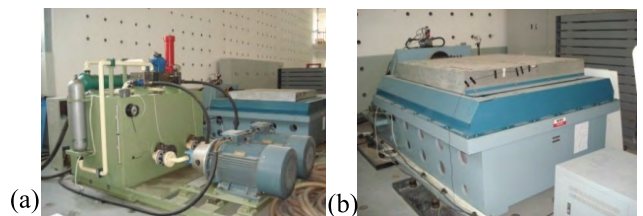

(a)   (b)

FIGURE 2. Shaking table test equipment. (a) Hydraulic system; (b) Shaking table.

As can be seen in Figure 3, the main technical indicators of the shaking table test equipment include a rated working frequency (40 Hz), a maximum acceleration (20 m/s^2), the maximum test load (5000 kg) and dimensions of the shaking table (1.5m × 1.5m).
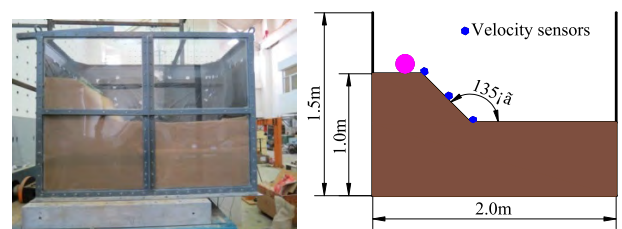

FIGURE 3. Shaking table test model.

B. EXPERIMENTAL MODEL

Model tests of actual projects should meet certain similitude principle. Similitude is a concept applicable to the testing of engineering models. A model is said to have similitude with the real application if the two share geometric similarity, kinematic similarity and dynamic similarity.

The scaling law between our test model and the actual projects follow the Buckingham Pi theorem [19], and the proportional relation for the similarity ratio is developed by Jiang, et al. [20], as shown in Table 1.

TABLE 1. 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	$C_T = C_K^{-1/2} C_\rho^{1/2n} C_L^{\frac{2n-1}{2n}}$	Control variable
Density ρ	$C_\rho = 1$	
Strain level $\gamma/\bar{\gamma}$	$C_{(\gamma/\bar{\gamma})} = 1$	
Dynamic displacement u	$C_u = C_{[\tau]} C_K^{-1} C_\rho^{-1/n} C_L^{\frac{n-1}{n}}$	
Vibration velocity v	$C_v = C_u / C_T = \dfrac{C_{[\tau]}}{C_K^{-1/2} C_\rho^{-\frac{3}{2n}} C_L^{-\frac{1}{2}}}$	
Vibration frequency w	$C_w = 1 / C_T = C_K^{1/2} C_\rho^{-\frac{1}{2n}} C_L^{\frac{1-2n}{2n}}$	
Damping ratio λ	$C_\lambda = 1$	
Dimensionless index n	$n = 2$	

$[\tau]$ is the horizontal shear strength, and $[\tau]$ is computed by Equation (9).

$$[\tau] = \sqrt{(\frac{1+k}{2}\sigma_v \sin\varphi + c\cos\varphi)^2 - (\frac{1-k}{2}\sigma_v)^2} \quad (9)$$

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

In our experiment, similar material was developed based on the slope material in the actual engineering project. The material of the artificial slope is similar to phyllite [21], and the properties of materials simulating slope and block are shown in Table 2.

TABLE 2. Properties of materials simulating slope and block.

| Physical quantity	Slope material	Block material
Unit weight (kN/m^3)	25.4	25.1
Elastic modulus (MPa)	113	2.5
Poisson's ratio	0.31	0.3
Compressive strength (MPa)	0.88	0.41
Tensile strength (MPa)	0.09	0.03
Cohesion (KPa)	29.30	15.7
Internal friction angle (°)	35	34.4
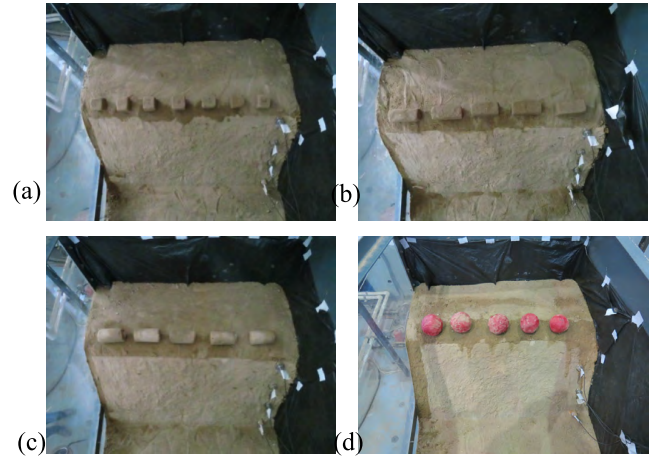
FIGURE 4. Blocks with different shapes. (a) Cubic blocks; (b) cuboidal blocks; (c) cylindrical blocks; (d) spherical blocks.

The slope models were made with varying heights and slope angles. As a sample, the length, width and height of one of the slope models are 2.0m, 1.5m, and 0.5m, as shown in Figure 3.

C. EXPERIMENTAL PROGRAM

As data was acquired in the laboratory test, the shapes of the falling blocks were spherical, cubic, cylindrical and cuboidal respectively. In the laboratory tests, each case was tested 50 times to account for the randomness of the impact. The effect of block shape was investigated by comparing spherical, cubic, cylindrical and cuboidal blocks with the same mass value of 0.28kg. The effect of block mass was examined by releasing spherical blocks on slope top with the different mass values which were 0.01kg, 0.04kg, 0.08kg, 0.16kg, 0.28kg, 0.70kg, 1.30kg and 4.40kg. For effect of slope height and slope angle, tests were performed by controlling the slope height and slope angle, as shown in Table 3.

TABLE 3. Details of experimental program.

Block shape	Block mass (kg)	Slope height(cm)	Slope angle(°)
●	0.01		20
	0.04		30
⬮	0.08	10	40
	0.16	20	45
▰	0.28	30	50
	0.70	40	60
▭	1.30	50	70
	4.40		80

Vibration frequencies and earthquake types are the main external factors in induced rockfall. Thus, we used sine waves with various frequencies, near field and far field seismic waves recorded at different sites (T1-I-1, T2-I-3, T1-II-1, T2-II-3, T1-III-1 and T2-III-3), El Centro seismic wave and

TABLE 4. Test conditions.

Category	Earthquake type	Peak acceleration (g)
SIN-01	Sine wave 2Hz	0.3
SIN-02	Sine wave 4Hz	0.3
SIN-03	Sine wave 6Hz	0.3
SIN-04	Sine wave 8Hz	0.3
SIN-05	Sine wave 10Hz	0.3
SIN-06	Sine wave 12Hz	0.3
SIN-07	Sine wave 14Hz	0.3
T01	T1-I-1	0.3, 0.4, 0.5
T02	T2-I-3	0.3, 0.4, 0.5
T03	T1-II-1	0.3, 0.4, 0.5
T04	T2-II-1	0.3, 0.4, 0.5
T05	T2-III-3	0.3, 0.4, 0.5
T06	T1-III-1	0.3, 0.4, 0.5
T07	T2-III-3	0.3, 0.4, 0.5
T08	Tianjin seismic wave	0.3, 0.4, 0.5
EI-01	El Centro seismic wave	0.3, 0.4, 0.5

Tianjin seismic waves as inputs to the shaking table. The peak accelerations of the seismic waves were adjusted for different values to investigate the influence of peak acceleration on rockfall runout distance. The test conditions are shown in the Table 2.

D. ROCKFALL RUNOUT RANGE CLASSIFICATION

The rockfall runout range is divided into a number of subintervals. The rockfall runout distance is measured from the slope toe, and one fifth of the slope height H is divided into a subinterval, as shown in Figure 5.
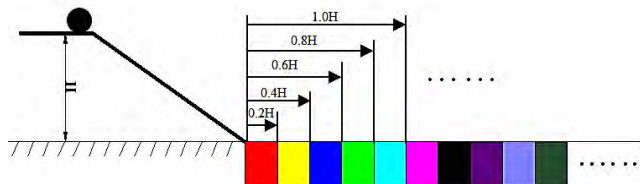

FIGURE 5. Subintervals of the rockfall runout range.

According to the distance to the slope toe, we define different subintervals as different rockfall runout ranges, as shown in Table 5.

In the laboratory tests, each case was tested 50 times to account for the randomness of the impact. The number of all the blocks with the same shape and mass are the population T_n, and the number of the blocks with the same shape and size in different subintervals are sample S_n

$(n = 1, 2, 3 \ldots \ldots)$. Then probability ζ_n of the blocks in the different subintervals could be computed, as shown in Equation (10).

$$\zeta_n = \frac{S_n}{T_n} \qquad (10)$$

where ζ_n is probability of the rock falls in the different subintervals; T_n is the number of all the rock falls with the same shape and size; and S_n is the number of the rock falls with the same shape and size in different subintervals.

IV. EXPERIMENTAL RESULTS

A. INFLUENCE FACTORS ANALYSIS

Rockfall is a common seismic damage phenomenon, and also it is also an important content area of engineering geological evaluation in earthquake locations [22]. There are many influencing factors to rockfall runout which is a very complicated process. In fact, the influencing factors can be roughly divided into three categories [23]: the first is the slope shape, such as slope height and slope angle. The second is the type of the block, such as block mass, block shape. The third is the earthquake, such as peak acceleration and vibration frequency. In our study, the effect of the main factors including peak acceleration, vibration frequency, slope height, slope angle, block mass and block shape on the rockfall runout distance and the range were investigated.

1) EFFECT OF VIBRATION FREQUENCIES

To address the effect of the vibration frequencies on rockfall runout range, the sphere, cube, cylinder and cuboid blocks were used to conduct the shaking table test, as shown in Figure 5. The probability values of the blocks falling into the different subintervals affected by different vibration frequencies are shown in Figure 6.

Figure 6 shows the probability values of the blocks with various mass and shapes in different rockfall runout range. It should be noted that the rockfall runout distance increases with increasing block mass. The rockfall runout distance increases with increasing block mass when the block shape is sphere. The main reason is that the block with the maximum mass has the greatest gravitational potential energy, and the gravitational potential energy transforms to the kinetic energy when the blockfalls in an earthquake event. The maximum rockfall runout range is 1.8~2.0H when the block shape is spherical. Taking one value of the cylinder block mass as an example, when the block mass is 0.28kg, the probability of the blocks falling into the subinterval 1.2~1.4H is the highest. Consequently, 1.4H can be used as the critical protection range for the major construction and infrastructure projects when the cylinder block mass is 0.28kg.
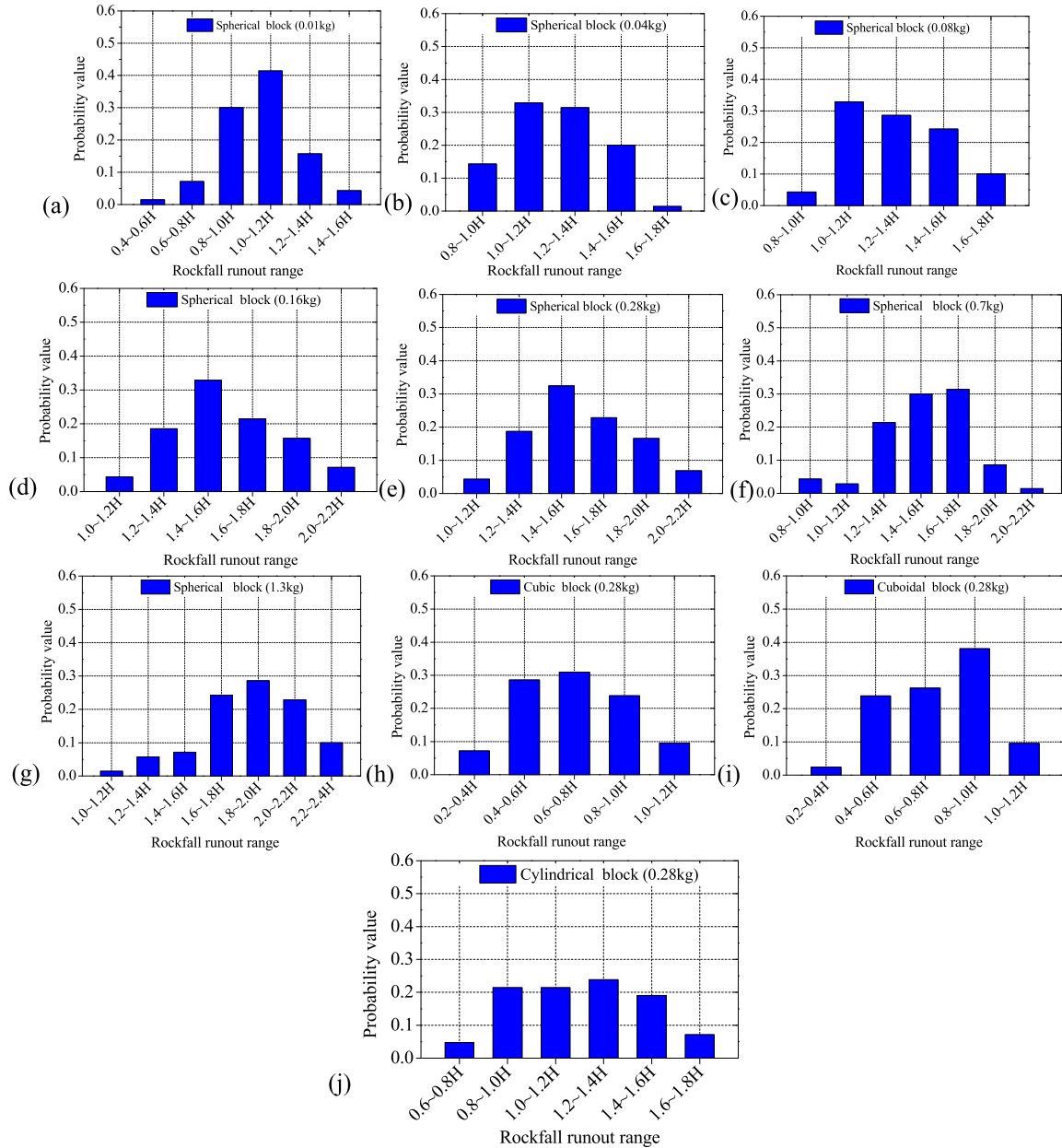
TABLE 5. Definition of the subinterval.

Rockfall range	0~0.2H	0.2~0.4H	0.4~0.6H	0.6~0.8H	0.8~1.0H	1.0~1.2H	1.2~1.4H	1.4~1.6H
Subinterval	0	1	2	3	4	5	6	7

FIGURE 6. Probability values of the blocks falling into the different subintervals. (a) Sphere block (0.01kg); (b) Sphere block (0.04kg); (c) Sphere block (0.08kg); (d) Sphere block (0.16kg); (e) Sphere block (0.28kg); (f) Sphere block (0.70kg); (g) Sphere block (1.3kg); (h) Cube block (0.28kg); (i) Cuboid block (0.28kg); (j) Cylind er block (0.28kg).

2) EFFECT OF EARTHQUAKE TYPES

In this section, we conduct the shaking table test to investigate the influence laws of earthquake types on the rockfall runout range. The earthquakes include near field and far field seismic waves recorded at different sites. The probability values of the blocks falling into the different subintervals affected by different earthquakes are shown in Figure 7.

As can be seen in Figure 7, the rockfall runout distance increases with the increasing of the block mass under earthquakes with different spectral characteristics and different site types. The rockfall runout distance increases with the increasing of the block mass under different earthquakes

which shows that the influence of block mass on rockfall runout distance is more significant than the influence of earthquake type. The maximum rockfall runout range is 1.6~1.8H when the block shape is spherical. Taking one value of the cylinder block mass as an example, when the block mass is 0.28kg, the probability of the blocks falling into the range 1.0~1.2H is the highest. Consequently, 1.2H can be used as the critical protection range for the major construction and infrastructure projects under earthquakes when the cylinder block mass is 0.28kg.

In the above text, we could not be certain whether the near field earthquake or the far field earthquake has the
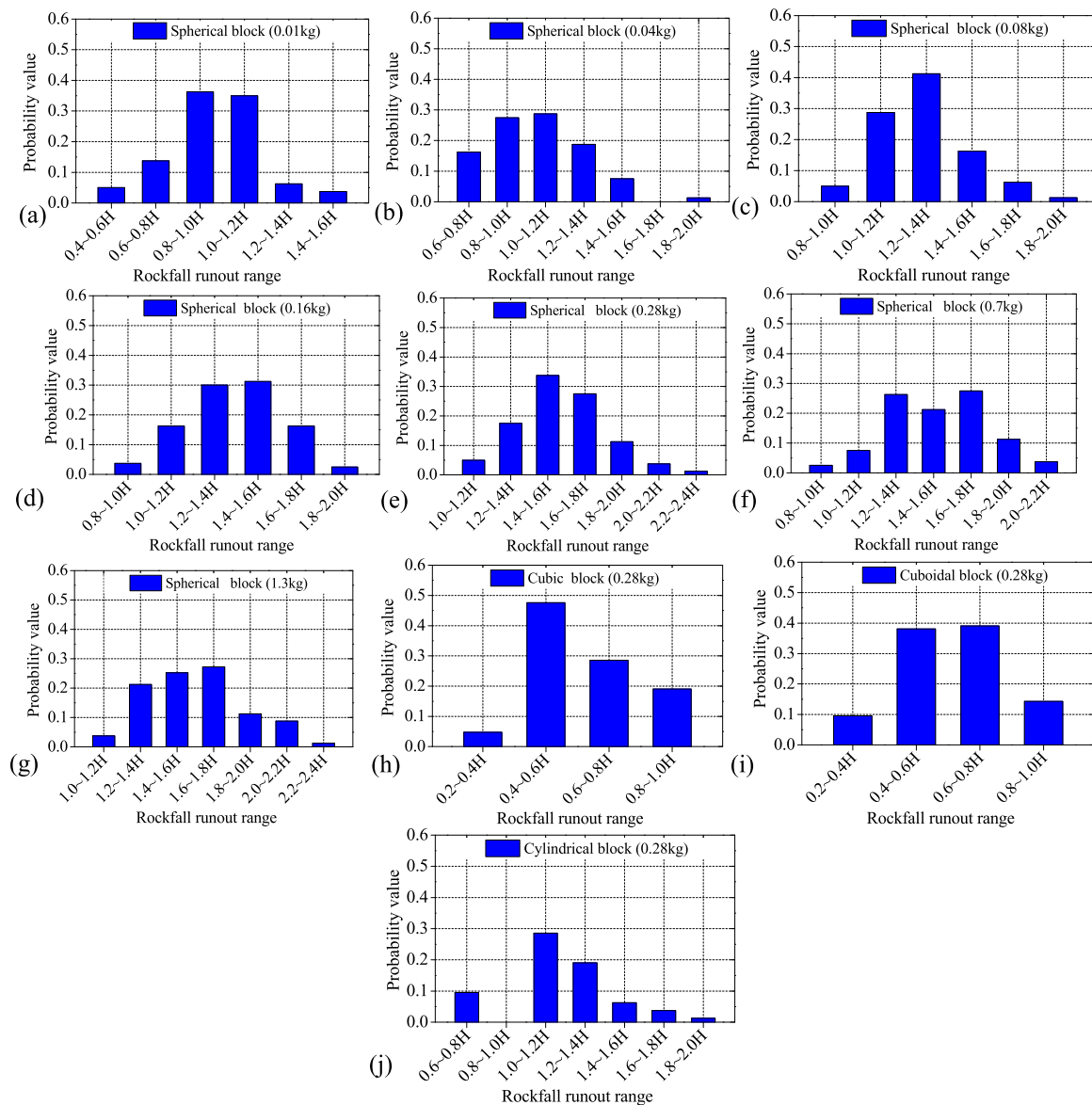
FIGURE 7. Probability values of the blocks falling into the different subintervals. (a) Sphere block (0.01kg); (b) Sphere block (0.04kg); (c) Sphere block (0.08kg); (d) Sphere block (0.16kg); (e) Sphere block (0.28kg); (f) Sphere block (0.70kg); (g) Sphere block (1.3kg); (h) Cube block (0.28kg); (i) Cuboid block (0.28kg); (j) Cylinder block (0.28kg).

greater influence on the rockfall runout distance. Consequently, we also conducted the shaking table test to see the influence laws of near field earthquake T2-I-1 and far field earthquake T1-I-1 on the rockfall runout distance. The rockfall runout distance is the distance average of the blocks in the subinterval with the highest probability value, and various blocks are marked with different identification number (ID), as shown in Figure 8.

In addition, peak acceleration is also a major trigger for rockfall and it is also a key factor affecting the rockfall runout distance. We use seismic wave El Centro to conduct the shaking table test to see the influence laws of the peak accelerations on rockfall runout distance, as shown in Figure 9

As shown in Figure 9, we can observe that the rockfall runout distance increases linearly with the increasing of peak accelerations which shows that the effect of peak acceleration on rockfall runout distance is very significant. Taking the sphere block as an example, the rockfall runout distance was 85cm when the peak acceleration was 0.5g. The rockfall runout distance was 62cm when the peak acceleration was 0.3g. When the peak acceleration was 0.5g, the rockfall runout distance increases by 37% compared with that when the peak acceleration was 0.1g. The rockfall runout distance with different block shapes from the largest to the smallest are spherical block, cylindrical block, cubic block, and cuboidal block. In addition, we found that the effect of the peak acceleration on sphere and cylinder blocks is greater
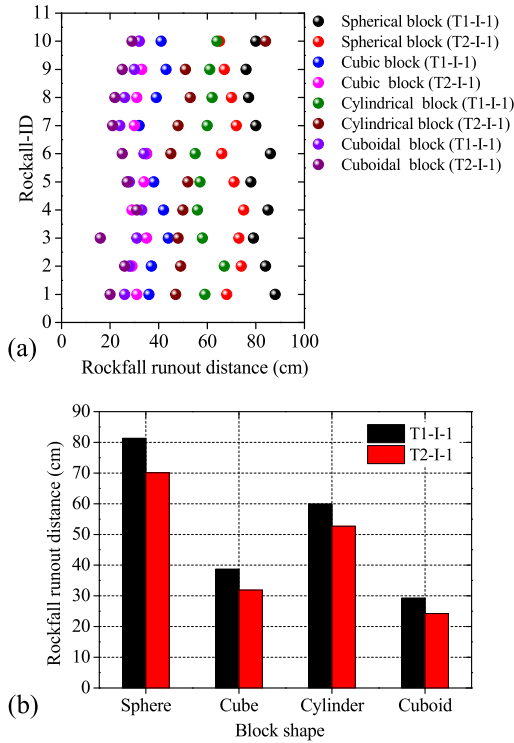
FIGURE 8. Rockfall runout distance under near field and far field earthquakes. (a) Runout distance; (b) Mean values of the runout distance.
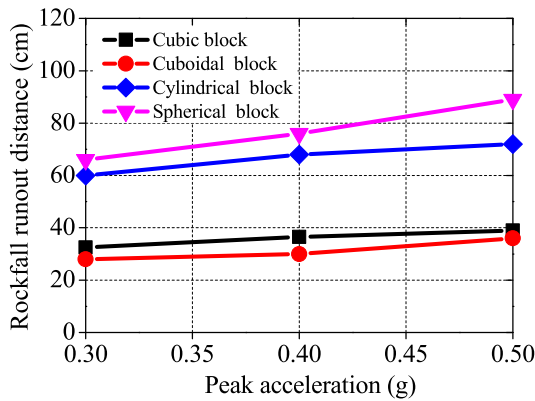

FIGURE 9. Rockfall runout distance under El Centro earthquake with different peak accelerations.

FIGURE 10. Influence laws of slope height and angle on runout distance. (a) Runout distance with different slope heights; (b) Runout distance with different slope angles.

increases first and then decreases. The runout distance increases with the increasing of slope angle until the angle increases to 60°, and the runout distance decreases as the slope angle continues to increase. However, the rockfall runout distance constantly increases with the increasing of the slope height. Taken the sphere block as an example, when the slope height was 50cm, the rockfall runout distance increases by 65% compared with that when the slope height was 10cm. Consequently, more attention should be paid to this phenomenon in engineering protection design. In addition, the block shapes with the maximum rockfall runout distance from the largest to the smallest are sphere block, cylinder block, cube block and cuboid block under the effect of different slope heights and angles.

B. COMPARISON OF SHAKING TABLE TEST RESULT AND THEORY RESULT

In this section, for evaluating the accuracy of the shaking table test results, we compare the rockfall runout distance computed by analytical solution of the basic equation of rockfall motion with the shaking table test results. There are three motion behaviors of rockfall including a sliding motion, rotational motion and bouncing motion [24], [25]. In our shaking test, the motion of the blocks mainly displays sliding motion and bouncing motion. The basic equations of the sliding and jumping motions are as follows.

than cube and cuboid blocks with the increasing of the peak accelerations.

3) EFFECT OF SLOPE HEIGHT AND ANGLE

To address the effect of the slope height and angle on the runout distance, the slope height was set to 10cm, 20cm, 30cm, 40cm and 50cm; the slope angle was set to 30°, 40°, 50°, 60°, 80° and 90°. The influence laws of slope height and angle on runout distance are shown in Figure 10.

As shown in Figure 10, by conducting a series of tests, we found that as for the slope angle influence, with the increasing of the slope angle, the rockfall runout distance
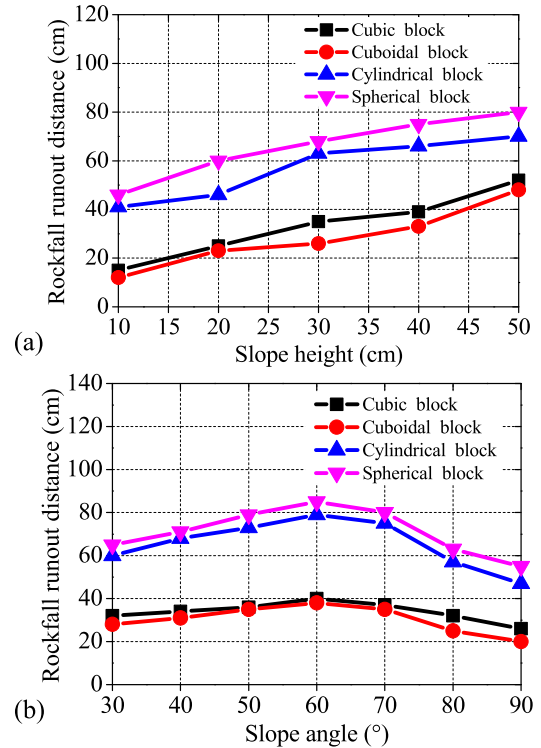
1) JUMPING MOTION

The basic equation of the jumping motion is presented in Equation (11).

$$x_d = x_0 + \frac{2V_0{}^2 \cos^2 \beta}{g}(\tan \theta - \tan \beta)$$
$$y_d = y_0 + (x_d - x_0) \tan \theta \tag{11}$$

where (x_d, y_d) is the setting point coordinates of the rockfall; (x_0, y_0) is the coordinates of a point on the slope; θ is the slope angle; and β is the angle between the tangents of the trajectory direction and the slope normal.

In addition, velocity (V_x, V_y) and displacement (x, y) along x and y are presented in Equation (12) and Equation (13).

$$V_x = V_0 \cos \beta e^{-a_k t}, \quad V_y = \frac{g}{a_k} + (V_0 \sin \beta - \frac{g}{a_k})e^{-a_k t} \tag{12}$$

$$x = x_0 + V_0 \cos \beta \frac{1 - e^{-a_k t}}{a_k}$$
$$y = y_0 + \frac{gt}{a_k} + (V_0 \sin \beta - \frac{g}{a_k})\frac{1 - e^{-a_k t}}{a_k} \tag{13}$$

where a_k is the air resistance factor; (x, y) is the displacement of the rock fall along x and y; and (V_x, V_y) is the velocity of the rock fall along x and y.

2) SLIDING MOTION

The basic equation of the sliding motion is presented in Equation (14) and Equation (15).

$$V = V_0 + gt(\sin \theta - \mu' \cos \theta) \tag{14}$$
$$S = V_0 t + \frac{1}{2}gt^2(\sin \theta - \mu' \cos \theta) \tag{15}$$

where V is the sliding velocity after t seconds; S is the sliding distance along the slope after t seconds.

To compare the rockfall runout distance computed by analytical solution with the shaking table test results, we use T1-II-1, T2-II-1 and El Centro seismic waves whose peak accelerations were 0.3g as the input loads of the shaking table. The slope height was 40cm, and the slope angle was 45°. The block shape was spherical and the block mass was 0.28kg. The comparison results are shown in the Figure 11.

As can be seen in Figure 11, the formula calculation results are in good agreement with the experiment results except that there is certain discrete which is in the permissible range. Due to errors in the process of experiment, such as, operational errors, measurement errors, etc., the rockfall runout distance distance is different each time, however, the error is within the allowable range. The comparison results show the credibility and the accuracy of the shaking table test results.

V. PREDICTION MODEL OF THE ROCKFALL RUNOUT RANGE BASED ON OUR IMPROVED KNN ALGORITHM

A. 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 a normalization [26].
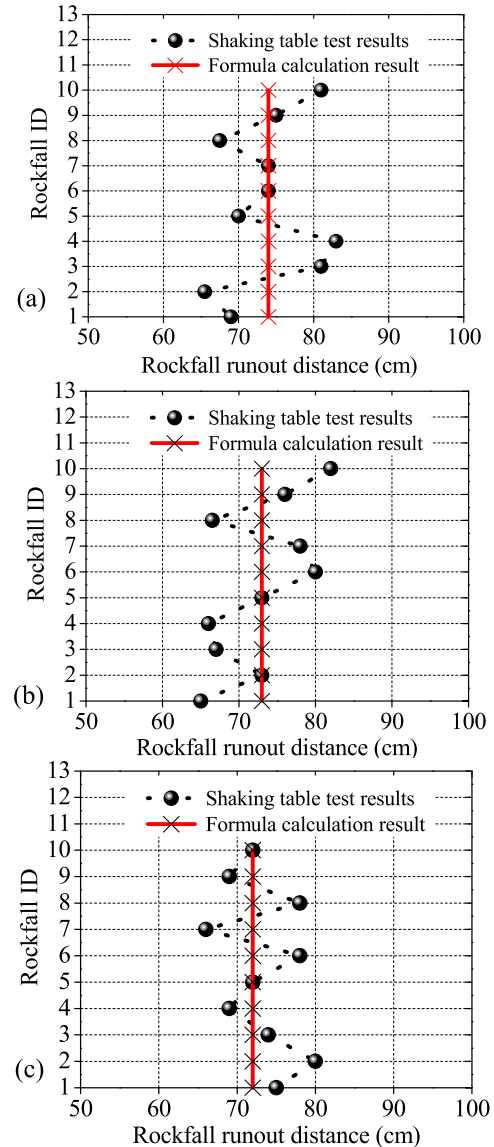
FIGURE 11. Comparison results of shaking table test and formula calculation. (a) T1-II-1; (b) T2-II-1; (c) El Centro.

We compute the upper and lower bound of each predictor, and the process for the used normalization is represented as Equation (16), Equation (17) and Equation (18).

$$\bar{y}_i = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \tag{16}$$

$$y_{\min} = \min\{y\} = \min_{j=1,\cdots\cdots,n}\{y_j\} \tag{17}$$

$$y_{\max} = \max\{y\} = \max_{j=1,\cdots\cdots,n}\{y_j\} \tag{18}$$

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

Accordingly, the value of each predictor is normalized to between 0 and 1 based on the Equation (16), Equation (17) and Equation (18).

B. PREDICTION MODEL

The prediction model of the rockfall runout range based on our improved algorithm can be expressed as follows.

Let X denote a set of sand liquefaction sample, and suppose X is $X = \{x_n \in R^m\}_{n=1}^N$, where x_i represents the feature of the i-th sand liquefaction sample, N is the total number of features, and m is the feature dimension. In addition, let y_i represents the sand liquefaction grade, and $y_i \in \{0, 1, 2, 3\}$, $i = 1, 2, \cdots, N$. Therefore, the sample set of the prediction model is shown in Equation (19).

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ x_N & y_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdot & \cdot & \cdot & x_{1m} & y_1 \\ x_{11} & x_{12} & \cdot & \cdot & \cdot & x_{2m} & y_2 \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ x_{N1} & x_{N2} & \cdot & \cdot & \cdot & x_{Nm} & y_N \end{bmatrix} \tag{19}$$

Given the unknown sample $x = (x_1 x_2 \cdots x_m)$, our proposed rockfall runout range prediction model based on our improved algorithm can be expressed as Equation (20).

$$\bar{y} = \arg\min_{w_i} f_i(x) = \arg\min_{w_i} d(x, \bar{x}_i^{PNN}) \tag{20}$$

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

In our prediction model, we used the 70 typical rockfall cases and 17 typical rockfall cases as training samples and testing samples respectively. Thus the training dataset contains 70 samples and testing set is 17.

VI. ROCKFALL RUNOUT RANGE PREDICITON
In this section, we will investigate our improved prediction approach based on the KNN algorithm. The following prediction experiments will show whether our proposed prediction model will achieve better prediction performance. For the purpose of comparison our proposed approach, we have also built four other prediction approaches including rockfall runout range prediction models based on KNN algorithm [14], WKNN algorithm [15], DWKNN algorithm [16] and SVM algorithm [26].

A. CRITERIA FOR OUR PREDICTION APPROACH PERFORMANCE
The accuracy, computed based on the percentage of all test samples classified correctly, is used to evaluate the prediction performance of the rockfall runout range. Accuracy tells us about the number of samples which are correctly predicted, and it is defined as follows.

$$Accuracy = \frac{\# \text{ test samples predicted correctly}}{\# \text{ test samples}}$$

where the *# test samples* denotes the total number of test samples; the *# test samples predicted correctly* is the number of test samples that are predicted correctly.
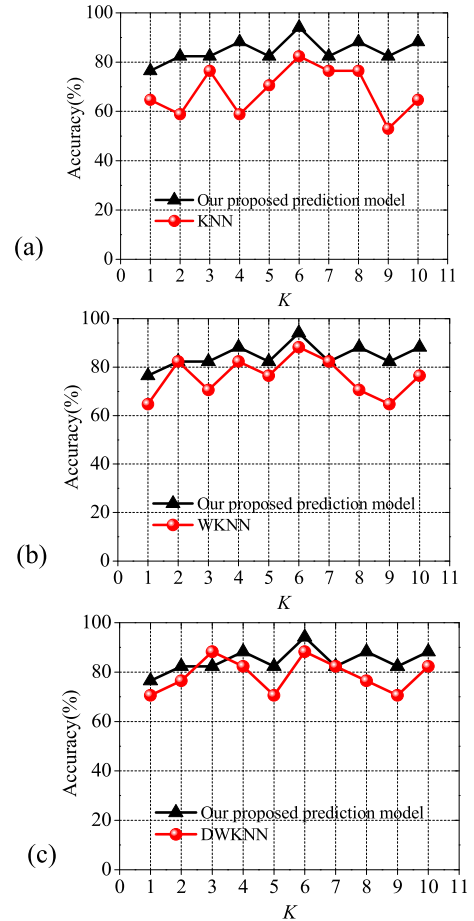
FIGURE 12. The prediction results between our proposed approach with other approach. (a) KNN algorithm; (b) WKNN algorithm; (c) DWKNN algorithm.

B. PREDICTION OF THE ROCKFALL RUNOUT RANGE
In this section, we conduct extensive experiments on runout range prediction to evaluate the performance of our proposed prediction model. We are interested to see how the performance changes if we modify the value of the neighborhood size k. Inspired by Ref. [16], the parameter k ranges from 1 to 10 with an interval of 1. Table 6 details the training samples which are from 70 typical rockfall experiment cases, and Table 7 shows the test samples that are from 17 typical rockfall experiment cases. This prediction experiment is implemented in eclipse 3.7.2 by Java language programming, and the hardware environment is Inter Core i7-6700 CPU 3.40GHz.

As can be seen in Table 7, our proposed prediction model based on the improved KNN algorithm achieves the best performance compared with the experimental results. The prediction accuracy is up to 94.12% which demonstrate that our proposed prediction model could be used for rockfall runout range prediction before the major projects construction near the slope.

TABLE 6. Training samples of the rockfall.

No.	Block mass (kg)	Slope height (cm)	Slope angle (°)	Peak acceleration (g)	Block shape	Vibration frequency (Hz)	Experimental results
1	1.3	50	30	0.5	1	2	6
2	0.7	50	30	0.5	3	6	5
3	0.04	10	30	0.5	4	4	2
4	0.7	40	40	0.55	3	8	5
5	0.04	20	40	0.6	2	6	3
6	0.04	20	40	0.6	2	14	3
7	1.7	50	40	0.55	1	12	7
8	0.01	30	30	0.6	2	14	3
9	1.3	40	60	0.4	1	12	7
10	0.08	30	60	0.5	2	10	3
11	0.7	50	40	0.5	1	2	7
12	1.3	50	50	0.55	3	2	7
13	1.3	40	50	0.35	3	2	7
14	0.7	40	50	0.35	1	4	7
15	0.7	50	50	0.35	3	6	6
16	0.7	50	40	0.55	1	6	7
17	0.28	30	40	0.5	3	8	6
18	1.3	40	30	0.5	1	8	7
19	1.3	50	60	0.45	1	10	7
20	0.16	40	45	0.4	2	12	5
21	0.16	30	45	0.55	2	14	4
22	0.08	30	45	0.55	2	2	3
23	0.04	30	45	0.35	2	4	3
24	0.7	50	45	0.5	1	6	7
25	1.3	50	50	0.5	1	8	7
26	0.01	10	30	0.5	4	10	0
27	0.04	20	40	0.65	4	12	3
28	0.7	50	50	0.5	1	2	7
29	0.28	30	40	0.55	3	14	5
30	0.04	20	70	0.6	4	14	2
31	0.16	40	45	0.55	3	12	5
32	0.08	20	80	0.6	2	10	2
33	0.7	30	40	0.55	3	8	7
34	0.28	20	30	0.5	3	6	5
35	1.3	50	45	0.55	1	4	7
36	0.16	50	70	0.6	2	8	3
37	0.16	30	80	0.6	4	6	3
38	0.7	50	20	0.5	1	10	7
39	1.3	50	20	0.5	1	12	7
40	0.08	10	20	0.55	4	14	1
41	0.08	20	20	0.55	4	2	2
42	0.01	20	30	0.55	4	2	0
43	0.16	30	20	0.6	2	4	4
44	0.7	50	60	0.45	1	12	7
45	0.16	30	70	0.6	2	6	3
46	0.16	40	40	0.5	3	10	6
47	0.28	30	40	0.6	3	8	5
48	0.7	30	45	0.6	3	8	5
49	0.7	30	30	0.6	3	8	5
50	1.3	40	45	0.5	3	10	6
51	1.3	40	45	0.55	1	10	6
52	1.3	40	60	0.55	3	12	6
53	0.16	30	80	0.55	2	2	3
54	0.7	50	50	0.5	1	10	7
55	0.01	10	20	0.5	4	2	1
56	0.04	20	30	0.6	2	2	2
57	0.28	30	30	0.6	2	4	4
58	0.28	30	50	0.3	2	4	4
59	0.7	40	60	0.3	3	6	5
60	0.7	40	60	0.45	3	6	5
61	0.1	50	45	0.5	3	8	5
62	0.4	50	45	0.5	4	2	2
63	0.8	50	45	0.5	4	8	3
64	0.01	50	45	0.5	4	4	2
65	0.01	50	45	0.5	4	4	2
66	0.01	50	45	0.5	4	4	2
67	13	50	45	0.5	2	6	3
68	7.0	40	40	0.35	2	6	3
69	7.0	40	50	0.35	3	8	4
70	7.0	40	60	0.35	3	10	4

Exegesis: Sphere block, cube block, cylinder block, and cuboid block are defined as 1, 2, 3 and 4 respectively.

TABLE 7. Test samples of the rockfall.

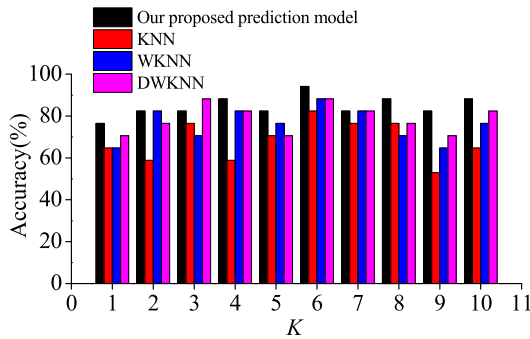| No.	Block mass (kg)	Slope height (cm)	Slope angle (°)	Peak acceleration (g)	Block shape	Vibration frequency (Hz)	Experimental results	Prediction results
1	7.0	40	70	0.35	2	4	3	3
2	7.0	40	80	0.35	2	2	3	3
3	0.01	10	45	0.5	4	2	2	4
4	7.0	20	45	0.5	3	8	5	5
5	0.01	30	45	0.5	2	4	3	3
6	0.01	40	45	0.3	3	2	6	6
7	0.04	40	45	0.3	1	4	6	6
8	0.08	40	45	0.3	1	6	6	6
9	0.16	40	45	0.3	1	8	7	7
10	0.28	40	45	0.3	1	10	7	7
11	0.28	40	45	0.3	2	12	3	3
12	0.28	40	45	0.3	3	14	4	4
13	0.28	50	30	0.3	1	2	5	5
14	0.28	30	70	0.4	1	4	4	4
15	0.28	20	60	0.5	1	6	7	7
16	0.70	30	30	0.4	4	8	2	2
17	1.30	20	60	0.3	4	10	5	5

FIGURE 13. Rockfall runout range prediction results with different neighborhood sizes.

The test results of the rockfall runout range based on our improved KNN algorithm are compared with the KNN, WKNN and DWKNN algorithms, as shown in Figures 12 and 13.

As can be seen in Figure 12 and Figure 13, it can be found that the accuracy of our proposed prediction approach based on the improved KNN algorithm is somewhat better than that of the KNN, WKNN and DWKNN algorithms in almost all of the test cases. That is to say, our proposed prediction approach almost gives an improvement over the other approaches with the increasing of the neighborhood size k. Taking one value of the neighborhood size k as an example, our proposed approach obtains an accuracy of 94.12% when the neighborhood size is fixed as 6. This result suggests that the prediction approach based on our improved KNN algorithm has the robustness to the sensitivity of different choices of the neighborhood size k with the satisfactory prediction performance to some degree.

TABLE 8. Rockfall runout range prediction results comparison of different approaches.

Algorithm		Accuracy (%)
	Our proposed	94.12
SVM	Linear kernel	88.23
	Polynomial kernel	76.47
	Radial basis function kernel	82.35
	Sigmoid kernel	70.59
	String kernel	64.71

Furthermore, we also conducted experiments on runout range prediction to compare our proposed prediction approach with the prediction approach based on the SVM algorithm. In this experiment, the neighborhood size k is fixed as 6. Detailed comparison results are shown in Table 8.

As can be seen in Table 8, our proposed prediction approach based on our improved KNN algorithm obtains a performance of 94.12%; while the prediction approach based on SVM with linear kernel, polynomial kernel, radial basis function kernel, sigmoid kernel and string kernel achieve a prediction accuracy of 88.23%, 76.47%, 82.35%, 70.59% and 64.71%, respectively. Consequently, it can be found that the performance of our proposed prediction approach is somewhat better than SVM with linear kernel, SVM with polynomial kernel, SVM with radial basis function kernel, SVM with sigmoid kernel and SVM with string kernel.

VII. CONCLUSIONS

(1) We used our improved KNN algorithm to construct a rockfall runout range prediction model, which is the first

work of applying the improved KNN algorithm to runout range prediction. Moreover, we evaluated the performance of our proposed prediction model based on our improved KNN algorithm by conducting extensive experiments on runout range prediction, and the experimental results demonstrated the effectiveness of our proposed prediction model.

(2) The effect of the main factors including slope height, slope angle, peak acceleration of earthquake, block mass and block shape on rockfall runout distance and the range were conducted via a series of shaking table tests. We discovered that the rockfall runout distance increases first and then decreases with the increasing of the slope angle, while the rockfall runout distance increases with the increasing of slope height, peak acceleration and block mass. The rockfall runout distance with different block shapes from the largest to the smallest are sphere block, cylinder block, cube block, and cuboid block. These obtained rockfall runout distance data can be helpful for rockfall protection design.

(3) We compared our proposed prediction model based on the improved KNN algorithm with four other prediction approaches, namely, the KNN, WKNN, DWKNN and support vector machine (SVM) algorithms. It can be found that the accuracy of our proposed prediction approach based on the improved KNN algorithm is somewhat better than that of the KNN, WKNN and DWKNN algorithms in almost all the test cases. In our future work, we aim to design selecting methods of predictors to improve the performance of rockfall distance prediction.

REFERENCES

[1] Z. Aihong, W. Shuaiwei, Y. Ying, and Y. Chao, "Analysis on characteristic parameters of rock slope rockfall movement and SVM prediction model," *J. Highway Transp. Res. Develop.*, vol. 34, no. 3, pp. 21–27, 2017.

[2] M. Jaboyedoff and V. Labiouse, "Preliminary estimation of rockfall runout zones," *Nature Hazards Earth Syst. Sci.*, vol. 11, no. 3, pp. 819–828, 2011.

[3] J. Hu *et al.*, "Experimental study on parameters affecting the Runout range of rockfall," *Adv. Civil Eng.*, vol. 2018, pp. 1–9, Mar. 2018.

[4] A. Preh, A. Mitchell, O. Hungr, and B. Kolenpratc, "Stochastic analysis of rock fall dynamics on quarry slopes," *Int. J. Rock Mech. Mining Sci.*, vol. 80, pp. 57–66, Dec. 2015.

[5] H. T. Hu, "Research on the collapse and falling stone," *J. Railway Eng. Soc.*, vol. 22, no. s1, pp. 387–391, 2005.

[6] W. Rong, D. Jie, and F. Yi, "Collapse distance of rock fall of slope under the earthquake," *Met. Mine*, no. 11, pp. 21–26, 2016.

[7] S. Lijun, L. Shuangqing, W. Yang, and X. Xingqian, "Application research of GPR in detection of earthquake-induced deformation of shaking table model slope," *J. Eng. Geol.*, vol. 23, pp. 728–733, 2015.

[8] Y. Liu, H. Yu, and P. Zhong, "The application of rough set neural networks of GSS-PSO in the risk evaluation of collapse and rockfall disasters," in *Proc. 2nd Int. Conf. Intell. Comput. Technol. Automat.*, Changsha, China, Oct. 2009, pp. 481–484.

[9] A. M. Fanos and B. Pradhan, "A novel rockfall hazard assessment using laser scanning data and 3D modelling in GIS," *CATENA*, vol. 172, pp. 435–450, Jan. 2019.

[10] L. Suping and Y. Shuxia, "Load prediction of HVAC system based on optimization parameter of phase space reconstruction," *Refrigeration*, vol. 42, no. 3, pp. 59–64, 2013.

[11] G. Jianping, Q. Wenmo, Z. Yi, S. Xiangjun, Z. Yongzhao, and Q. Weihua, "Locality constrained representation-based k-nearest neighbor classification," *Knowl.-Based Syst.*, vol. 167, pp. 38–52, 2019.

[12] G. Jianping, M. Hongxing, Q. Weihua, Z. Shaoning, R. Yunbo, and Y. Hebiao, "Generalized mean distance-based k-nearest neighbor classifier," *Expert Syst. Appl.*, vol. 115, pp. 356–372, 2019.

[13] M. Hongxing, G. J. Gou, W. Xili, K. Jia, and Z. S. Zeng, "Sparse coefficient-based k-nearest neighbor classification," *IEEE Access*, vol. 5, pp. 16618–16634, 2017.

[14] G. Jianping, Q. Yi, X. Yong, M. Q. Mao, and Z. Yongzhao, "A local mean representation-based k-nearest neighbor classifier," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 3, p. 29, 2019.

[15] S. A. Dudani, "The distance-weighted k-nearest-neighbor rule," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, no. 4, pp. 325–327, Apr. 1976.

[16] G. Jianping, L. Du, Z. Yuhong, and X. Taisong, "A new distance-weighted k-nearest neighbor classifier," *J. Inf. Comput. Sci.*, vol. 9, no. 6, pp. 1429–1436, 2012.

[17] N. Li, H. Kong, Y. Ma, G. Gong, and W. Huai, "Human performance modeling for manufacturing based on an improved KNN algorithm," *Int. J. Adv. Manuf. Technol.*, vol. 84, no. 1, pp. 473–483, 2016.

[18] M. M. Huang, R. S. Lin, S. Huang, and T. F. Xing, "A novel approach for precipitation forecast via improved K-nearest neighbor algorithm," *Adv. Eng. Inform.*, vol. 33, pp. 89–95, Aug. 2017.

[19] B. Louis, "The Pi theorem of dimensional analysis," *Arch. Rational Mech. Anal.*, vol. 1, no. 1, pp. 35–45, 1957.

[20] L. W. Jiang, L. K. Yao, and J. Wang, "Similitude for shaking table model test on side slope relating to dynamic characteristics and strength," *J. Transp. Sci. Eng.*, vol. 25, no. 2, pp. 1–7, 2009.

[21] Z. Qian, "Instability mechanism and control measures of bedding steep slope in high intensity seismic area," Shandong Univ., Jinan, China, Tech. Rep., May, 2015.

[22] H. B. Li, X. W. Li, W. Z. Li, S. L. Zhang, and J. W. Zhou, "Quantitative assessment for the rockfall hazard in a post-earthquake high rock slope using terrestrial laser scanning," *Eng. Geol.*, vol. 248, no. 8, pp. 1–13, 2019.

[23] L. Qing *et al.*, "Evaluation model of rockfall trajectory," *J. Natural Disasters*, vol. 12, no. 2, pp. 79–84, 2003.

[24] P. Ashayer, "Application of rigid body impact mechanics and discrete element modeling to rockfall simulation," Univ. Toronto, Toronto, ON, Canada, Tech. Rep., 2007.

[25] W. Yujia, L. Shuxue, D. Bingbing, and Z. Min, "Analysis and prediction of factors influencing rockfall distance," *Subgrade Eng.*, vol. 196, no. 1, pp. 47–51, 2018.

[26] J. H. Seo, Y. H. Lee, and Y. H. Kim, "Feature selection for very short-term heavy rainfall prediction using evolutionary computation," *Adv. Meteorol.*, vol. 2014, pp. 1–15, Jan. 2014.

SHUAI HUANG was born in Feicheng, Shandong, China, in 1987. He received the B.S. degree in civil engineering from the Shandong University of Science and Technology, Taian, in 2009, and the M.S. and Ph.D. degrees in disaster prevention and reduction engineering and protective engineering from the University of Science and Technology Beijing, Beijing, in 2015.

From 2015 to 2017, he was a Research Assistant with the Institute of Crustal Dynamics, China Earthquake Administration. Since 2018, he has been an Assistant Professor with the Institute of Crustal Dynamics, China Earthquake Administration. He is the author of one book and over 40 articles. His research interests include earthquake resistance and disaster reduction of urban infrastructure, seismic response mechanism and seismic fortification of marine engineering structures, and seismic liquefaction mechanism of land and sea sites.

YUEJUN LYU (M'63) received the B.S. degree in probability statistics from the University of Science and Technology of China, Hefei, Anhui, in 1989, and the Ph.D. degree in mechanical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2006. He is currently pursuing the Ph.D. degree in solid geophysics with the China University of Geosciences, Beijing, China, in 2008.

From 1995 to 1998, he was a Research Assistant with the Institute of Crustal Dynamics, China Earthquake Administration. Since 2003, he has been a Professor with the Institute of Crustal Dynamics, China Earthquake Administration. His research interests include seismic activity model construction, analysis of seismic hazard, zonation map compilation for earthquake fortification, and seismic risk estimating.

YANJU PENG (F'78) received the B.S. degree in inorganic normal material engineering from the Taiyuan University of Technology, Taiyuan, Shanxi, in 1998, and the Ph.D. degree in solid earth geophysics from the China University of Geosciences, Beijing, China, in 2009.

From 2002 to 2004, she was a Research Assistant with the Institute of Crustal Dynamics, China Earthquake Administration. Since 2012, she has been a Professor with the Institute of Crustal Dynamics, China Earthquake Administration. Her research interests include site soil dynamic characteristics and earthquake site effects, and seismic response of soft site and its applications to the determination of design ground motion parameters for offshore structures.

MINGMING HUANG (F'86) received the B.S. degree in mathematics and applied mathematics from Shandong Normal University, Jinan, Shandong, in 2009, and the Ph.D. degree in control science and engineering from the University of Science and Technology Beijing, Beijing, in 2016.

Since 2016, she has been a Research Assistant with the Beijing Meteorological Information Center. Her research interests include precipitation forecast, the application of machine learning method in engineering, and pattern recognition.

● ● ●