Transportation Geotechnics 37 (2022) 100878

Contents lists available at ScienceDirect

Transportation Geotechnics

journal homepage: www.elsevier.com/locate/trgeo

Novel hybrid MFO-XGBoost model for predicting the racking ratio of the rectangular tunnels subjected to seismic loading

Van-Quang Nguyen [a,b,1], Viet-Linh Tran [b,1], Duy-Duan Nguyen [b], Shamsher Sadiq [c], Duhee Park [a,*]

[a] Department of Civil and Environmental Engineering, Hanyang University, Seoul 04763, South Korea
[b] Department of Civil Engineering, Vinh University, Vinh 461010, Viet Nam
[c] Department of Civil Engineering, Mirpur University of Science and Technology (MUST), Mirpur 10250, AJK, Pakistan

ARTICLE INFO

Keywords:
Dynamic analysis
Extreme gradient boosting
Moth-flame optimization
Racking ratio
Rectangular tunnel
Web application

ABSTRACT

This study proposes a novel hybrid MFO-XGBoost model that integrates the moth-flame optimization (MFO) algorithm and the extreme gradient boosting (XGBoost) to predict the racking ratio of rectangular tunnels subjected to seismic loading. For this purpose, a nonlinear finite difference model of soil-tunnel considering a realistic partial-slip condition is developed and validated against centrifuge test results. Then, 2040 dynamic simulations subjected to 85 ground motions are analyzed to cover a comprehensive suite of soil-tunnel configurations. Based on the generated database, the MFO-XGBoost model is constructed to capture the relationship between various effective parameters and the racking ratio of the rectangular tunnel. The obtained results are compared with those of four existing models to evaluate the performance of the proposed MFO-XGBoost model. The comparison reveals that the proposed MFO-XGBoost model captures well the numerical results of the racking ratio and outperforms other models. Among twelve input variables, parameters with primary and secondary influences are identified. Finally, a web application is built based on the proposed MFO-XGBoost model to calculate the racking ratio of rectangular tunnels, which is computationally more effective compared with alternative procedures.

Introduction

Tunnels play an essential role in the urban transport infrastructure. The tunnels performed better than above-ground structures during past severe seismic events [1–3]. However, damage to tunnels observed in recent earthquakes [4–6] demonstrates that their seismic performances should be carefully evaluated. The seismic response of rectangular tunnels has been extensively studied using experiments [7–11], numerical simulations [12–16], and analytical methods [17–19]. Nonetheless, the results of these works have not been widely applied in the design practice. Conventional design methods [2,20–23] are still used due to their simple implementation. Wang [20] proposed a simplified static frame analysis method for evaluating the seismic response of rectangular tunnels using the racking ratio (*R*), which is defined as the ratio of racking deformation of the structure to the free-field racking deformation. Representative empirical correlation between *R* and the flexibility ratio (*F*) was presented based on a series of dynamic analyses, where *F* characterizes the relative stiffness of soil and structure. The no-slip interface was assumed, and a linear soil model was applied. Because of the small number of cases that were performed, the influences of the structure types, soil profiles, buried depths, and input ground motions were not quantified. Penzien [21] proposed analytical *F-R* equations considering the effects of soil-structure interface conditions and the Poisson's ratios. Anderson [22] and Zhang and Liu [23] recommended a *F-R* relation by fitting the *F-R* results from numerical analyses. However, recent studies demonstrate that the deformation of rectangular tunnels during seismic shaking is not only a pure racking but also a coupled racking-rocking mode [24–28]. Tsinidis and Pitilakis [29] developed a new set of *F-R* relations for different soil-tunnel configurations accounting for the effects of the rocking response using a constant linear shear wave velocity profile. However, the slip contact interface was not simulated, only presenting outputs for the unrealistic no-slip condition. A wide range of ground motion characteristics was not considered. Furthermore, the constant shear wave velocity profile fails to capture

* Corresponding author.
E-mail address: dpark@hanyang.ac.kr (D. Park).
[1] Co-first authors.

https://doi.org/10.1016/j.trgeo.2022.100878
Received 11 July 2022; Received in revised form 1 October 2022; Accepted 1 October 2022
Available online 8 October 2022

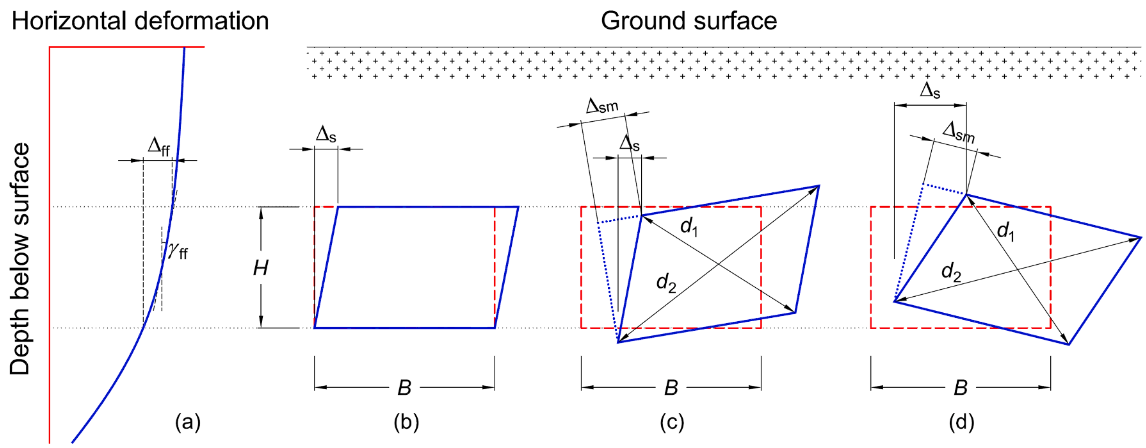

Fig. 1. Typical free-field ground distortion imposed on a underground box tunnels: (a) soil deformation profile, (b-d) structure racking deformation for the case of $F = 1$, $> 1F$, $F < 1$, respectively (modified after Hashash, Hook [2], Wang [20], Tsinidis and Pitilakis [29]).

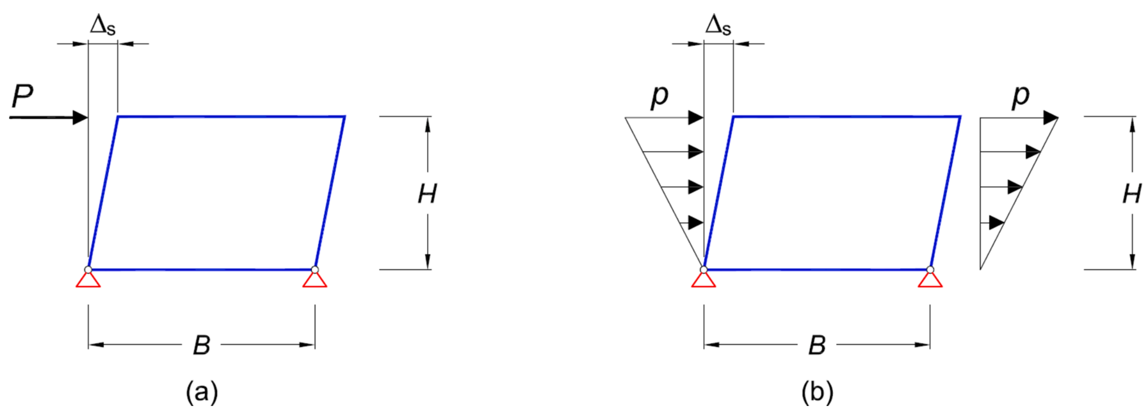

Fig. 2. Simplified frame analysis model: (a) pseudo-concentrated force for deep tunnels, (b) pseudo-triangular pressure distribution for shallow tunnels (modified after Wang [20]).

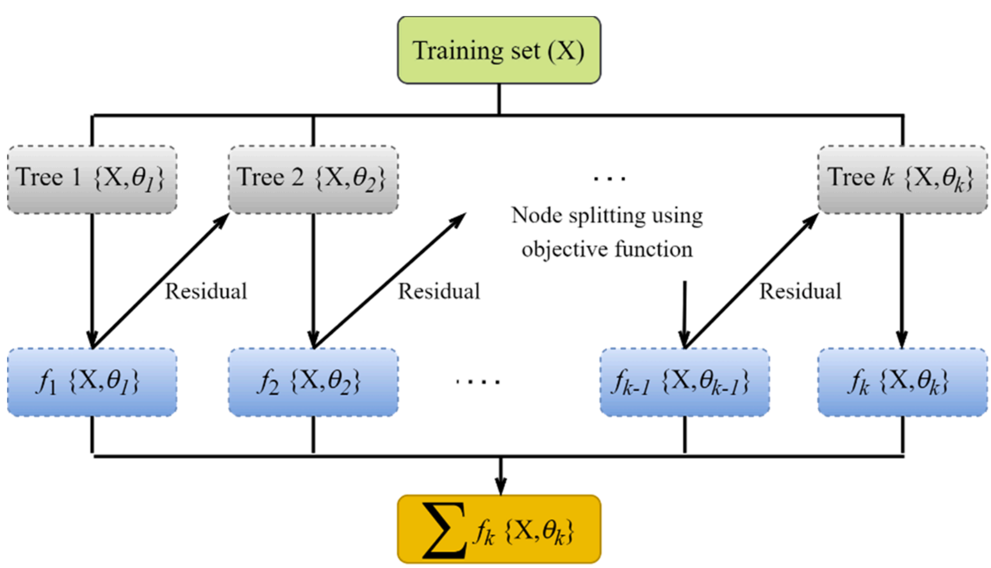

Fig. 3. Typical flowchart of XGBoost algorithm.

the variation of the soil stiffness along with the tunnel height, which may influence the seismic response of tunnels.

Recently, machine learning (ML) methods have been successfully applied to complex problems in geotechnical engineering [30–33], including landsides [34–36], excavations [37–41], slopes [42–45], dams [46–49], characterization of soil/rock properties [50–59], and pile foundations [60–63]. Several studies used ML algorithms for predicting the response of soil-tunnel interaction. Goh, Zhang [64] utilized

Table 1
Properties of tunnel structure models [86].

Parameters	Designed target properties	Centrifuge model properties (prototype scale)
Height (m)	8	8
Width (m)	14	14
Wall, slab thickness (m)	0.8	0.57
Material	Reinforced concrete	6061 Aluminum
Density (kg/m^3)	2400	2700
Young's modulus	2.50×10^7	6.89×10^7
Poission's ratio	0.2	0.33

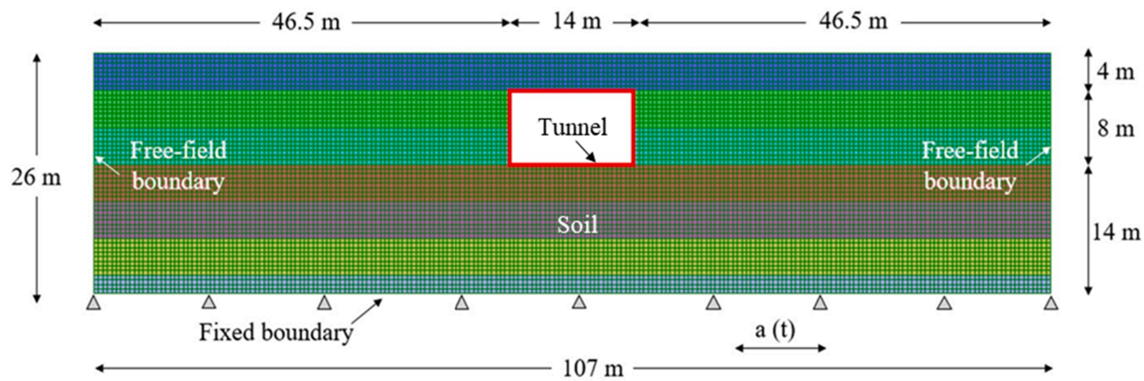

Fig. 4. The numerical simulation model.

Table 2
Parameters used for Darendeli [93] formulation.

Parameter	Assumed value
Lateral at-rest earth pressure coefficient (K_o)	0.46
Plastic index (PI)	0
Over consolidation ratio (OCR)	1
Excitation frequency	1
Number of cycle loading	10

Table 3
Interface element properties.

Parameters	Value
Normal stiffness, K_n (Pa/m)	10^{10}
Shear stiffness, K_s (Pa/m)	10^{10}
Friction angle (degree)	33

multivariate adaptive regression splines (MARS) to predict the surface settlement due to tunneling. Zhou, Shi [65] used random forest; whereas Zhang, Li [66] combined artificial neural network (ANN), support vector machine (SVM), extreme gradient boosting (XGBoost), and MARS approach. Zheng, Yang [67] applied the MARS model to estimate the earthquake-induced uplift displacement of the circular tunnel. Wang, Wang [68] employed the single shot detector algorithm to detect the crack of tunnel lining. Zhang, Li [69] evaluated the lining bending moment for twin tunnels based on MARS and the decision tree method. Wang, Li [70] developed a dynamic regression model by using the bidirectional long short-term memory (Bi-LSTM) with light gradient boosting machine (LGBM) to predict the advance speed and torque during the shield tunneling process. Although various studies on soil-tunnel interaction used ML, a model for estimation of *R* of rectangular tunnels induced by earthquake loading has not yet been developed. Moreover, several studies have also shown that hyperparameters significantly affect ML model performance [71–74]. ML models with default parameters have the major disadvantage of overfitting or underfitting because they introduce bias and variance [72,75]. It eventually leads to poor generalizability and inaccuracy when

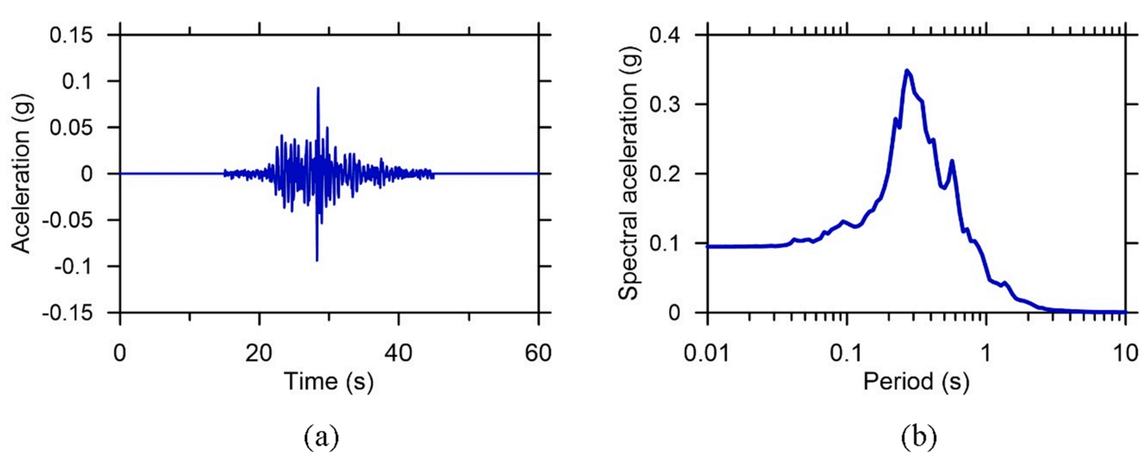

Fig. 5. Input motion at the base of centrifuge container and used in the numerical model: (a) acceleration time history and (b) response spectra [86].

Table 4
Characteristics of input motion [86].

Event	Station	Year	PGA (g)	T_p (s)	I_a (m/s)	D_{5-95}(s)
Loma Prieta	Santa Cruz	1989	0.1	0.1	11.3	0.6

predicting new data samples.

This paper proposes a novel hybrid MFO-XGBoost model that integrates the moth-flame optimization (MFO) algorithm and the extreme gradient boosting (XGBoost) for predicting the racking ratio of rectangular tunnels. The XGBoost model is used to train the model, whereas the MFO algorithm is used to optimize the hyper-parameters of the XGBoost model. To do so, a database from 2040 numerical simulations is generated to develop the MFO-XGBoost model. The effect of different soil-tunnel configurations and ground motion characteristics are considered. The proposed racking ratio also accounts for the nonlinear behavior of the surrounding soil and the frictional contact interface between the soil and tunnel. The numerical model is validated against centrifuge test measurements before performing a parametric study. The results of the proposed MFO-XGBoost model are compared with the models presented in published studies. Finally, a web application (WA) is built based on the MFO-XGBoost model for practical application.

Simplified frame analysis method and exiting *F*-*R* relationships

The response of the rectangular tunnels subjected to earthquake excitations can be calculated using a dynamic or pseudo-static analysis. Although the dynamic analysis method is recognized to most closely represent the seismic response of tunnels [15,76,77], it is not widely used in practice because of its significant computational cost. Instead, the pseudo-static method is most often performed to design of rectangular tunnels, using either continuum or frame models. However, performing a continuum analysis using either finite element or finite difference methods also requires pronounced computational efforts. Instead, the simplified frame analysis method, which is reported to produce reasonable estimates of the tunnel response ([29] and also easiest to perform the analysis, is frequently utilized in tunneling design practice, especially during the initial design stages. The method accounts for the soil-structure interaction effect using the parameter *F*, which represents the relative stiffness between the tunnel and surrounding ground. The step-by-step design procedure is described as follows:

(1) Estimate the maximum free-field shear strain (γ_{ff}) or free-filed ground distortion (Δ_{ff}) corresponding to the top and bottom elevations of the tunnel (Fig. 1a).

(2) Calculate the relative stiffness (i.e. the flexibility ratio, *F*) between the surrounding ground and the tunnel:

$$F = \left(\frac{G_m}{K_s}\right)\left(\frac{B}{H}\right) \tag{1}$$

where G_m is the strain-compatible shear modulus of the surrounding ground; *B* and *H* are the width and height of the tunnel, respectively; K_s is the racking stiffness of the tunnel. K_s can be obtained by applying a unit concentrated lateral force at the roof of the tunnel while restraining the translation at the base in the frame analysis. K_s is defined as the ratio of the applied force to the resulting lateral displacement.

(3) Determine the racking ratio (*R*) based on the flexibility ratio (*F*) from step (2) and *F*-*R* relation. *R* is defined as the normalized structure

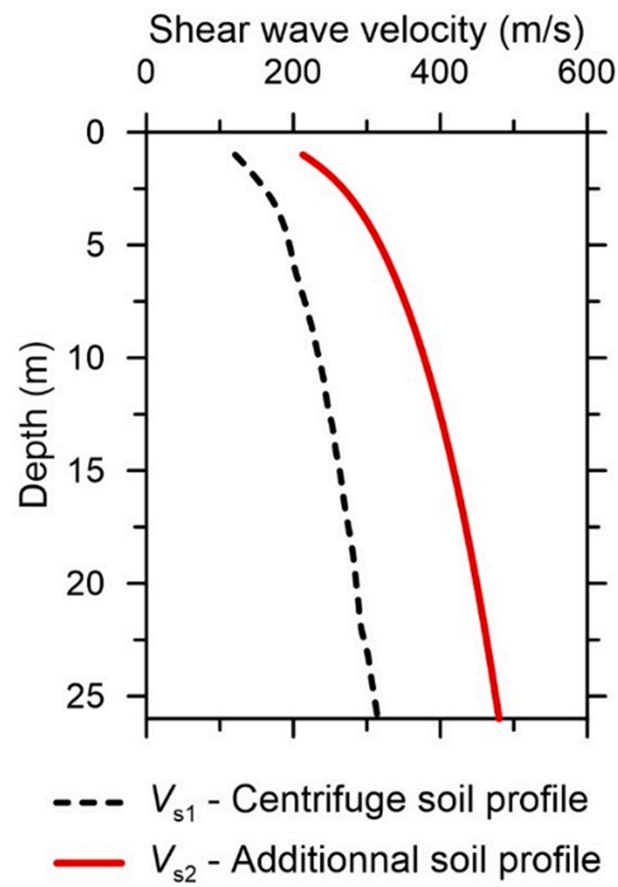

Fig. 7. Soil profiles used in the centrifuge test and parametric studies.

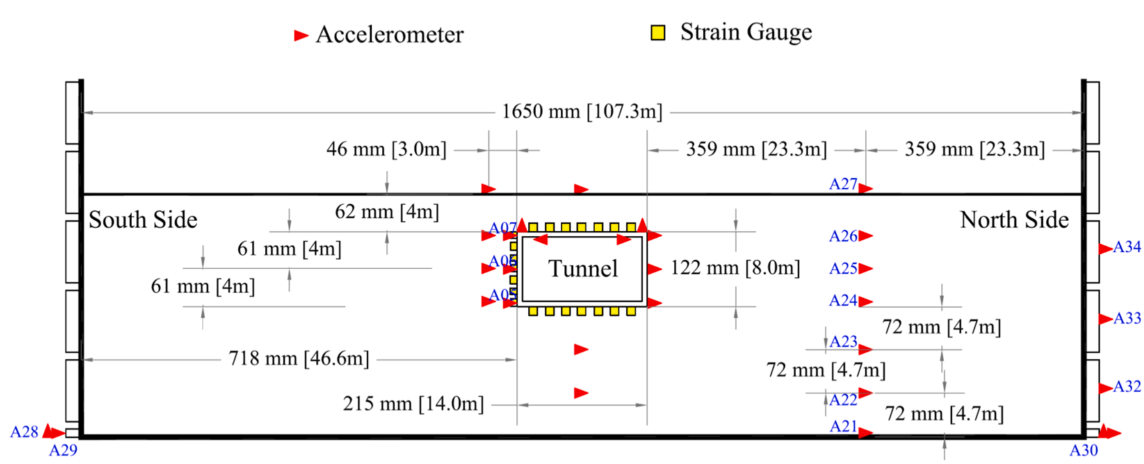

Fig. 6. Instrumentation layout of the centrifuge test (dimension in model scale) [86].

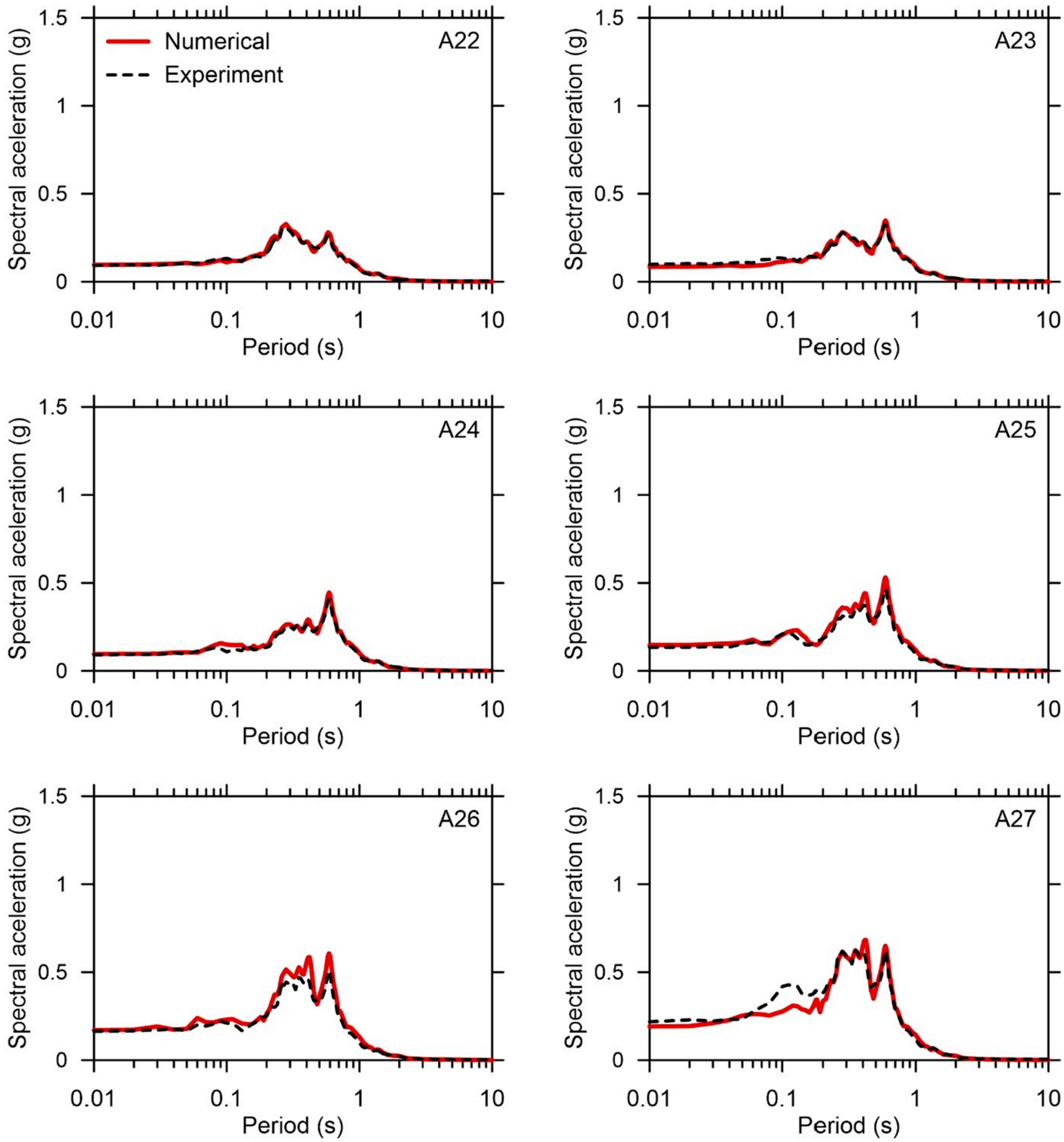

Fig. 8. The free-field response comparison between numerical results and centrifuge test data at various depths (A22-A27).

racking deformation (Δ_s) with respect to the free-field ground deformation (Fig. 1b):

$$R = \frac{\Delta_s}{\Delta_{ff}} \tag{2}$$

(4) Compute the racking deformation of the structure:

$$\Delta_s = R \times \Delta_{ff} \tag{3}$$

(5) The seismic demands are obtained by imposing Δ_s from step (4) to the structure using the simplified frame analysis, as shown in Fig. 2.

Several studies have proposed the *F-R* relationships. Wang [20] performed 25 linear soil-structure analyses for a range of tunnel types, soil profiles, soil cover depths, input motions, and Poisson's ratio. The no-slip interface condition was used. Penzien [21] established analytical *F-R* correlations considering the effect of soil-structure interface condition and ground Poisson's ratio (ν_s) as follows:

No-slip interface condition:

$$R = [4(1-\nu_s)F/(3-4\nu_s+F) \tag{4}$$

Full-slip interface condition:

$$R = [4(1-\nu_s)F/(2.5-3\nu_s+F) \tag{5}$$

Anderson [22], and Zhang and Liu [23] proposed following design curves for R by fitting numerical analysis results:

$$\text{Anderson [22]}, R = \frac{2F}{1+F} \tag{6}$$

$$\text{Zhang and Liu [23]}, R = \frac{1.75F}{1+F} \tag{7}$$

Tsinidis and Pitilakis [29] developed a new set of *F-R* relationships accounting for a wide range of soil-tunnel configurations. "Actual" structure racking deformation (Δ_{sm}) that accounts for the rocking rotation (Fig. 1c-d) is presented as follows:

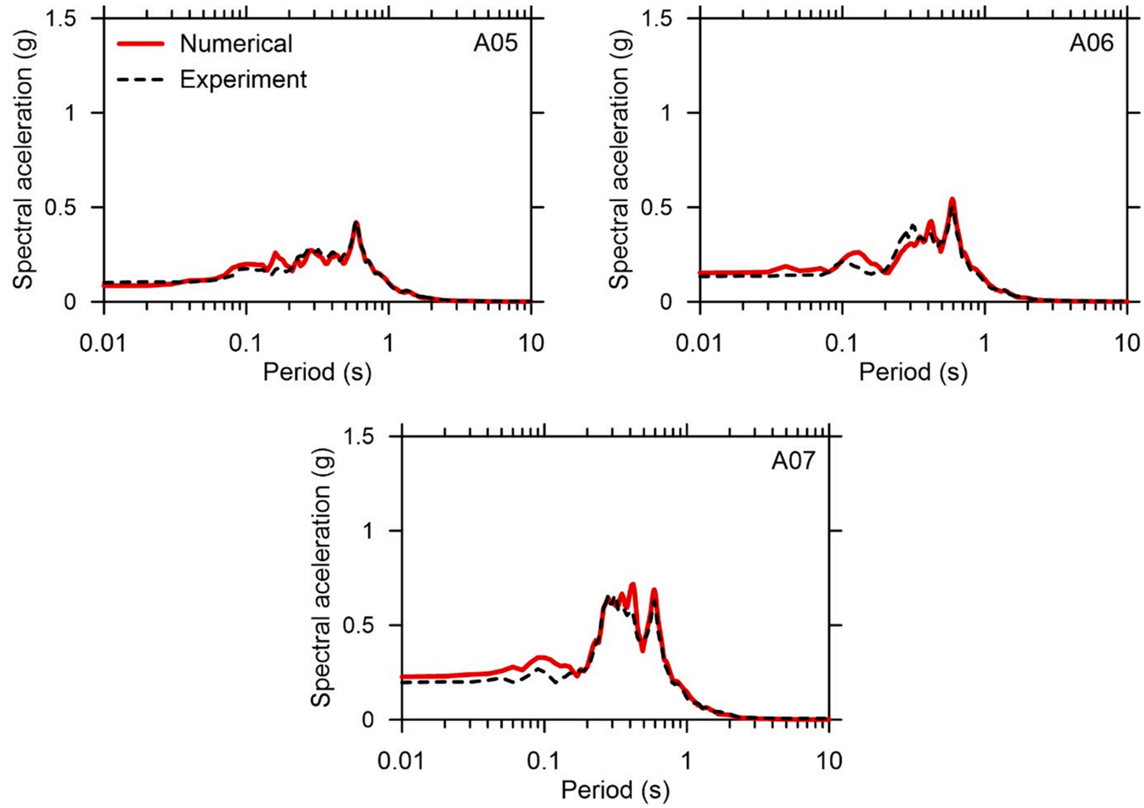

Fig. 9. The tunnel response comparison between numerical results and centrifuge test data at various depths (A05-A07).

$$R = \frac{\Delta_{sm}}{\Delta_{ff}} \tag{8}$$

Overview of XGBoost and MFO

XGBoost

XGBoost is an ensemble ML method proposed by Chen and Guestrin [78] and widely used in many studies [79–83]. This algorithm performs well on diverse datasets and gives the most accurate results. For the initial productivity prediction, each tree's results are accumulated based on the following equation:

$$\hat{y}_i = \varphi(x_i) = \sum_{k=1}^{K} f_k(x_i), f_k \in F \tag{9}$$

where $\hat{y}_i$ is the XGBoost's prediction, f_k is the regression tree's output, F is the regression tree's space.

The XGBoost uses the objective function of the loss function (l) and the regular term (Ω) to minimize the gap between the actual and predicted values (equation (10)). Herein, a regularization parameter is used to prevent model complexity and thus, reduce overfitting.

$$Obj = L(\varphi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{10}$$

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda||\omega||^2 \tag{11}$$

where γ is the leaf's complexity, λ is the penalty parameter, and $||\omega||$ is the vector score on the leaves.

The loss function is approximated by Taylor expansion, as follows.

$$L^{(t)} = \sum_{i=1}^{n}\left[g_i f_i(x_i) + \frac{1}{2}h_i f_t^2(x_i)\right] + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T}\omega_j^2$$

$$= \sum\nolimits_{j=1}^{T}\left[\left(\sum_{i\in I_j} g_i\right)\omega_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right)\omega_j^2\right] + \gamma T \tag{12}$$

where g_i and h_i are the first derivative and second derivative of the loss function, I_i is the total set of leaf nodes; t is the t^{th} iteration.

The following greedy algorithm is used to compare the variation of the objective function for each feature at each node before and after splitting:

$$L_{split} = \frac{1}{2}\left[\frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i + \lambda} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i + \lambda} - \frac{\left(\sum_{i\in I} g_i\right)^2}{\sum_{i\in I} h_i + \lambda}\right] - \lambda \tag{13}$$

where I_L and I_R are the instance sets of left and right leaf nodes after the split; I is the total set, $I = I_L \bigcup I_R$. Fig. 3 shows the flowchart of the XGBoost algorithm.

MFO

MFO is a new population-based algorithm proposed by Mirjalili [84]. MFO was inspired by the natural navigation technique of moths when they move at night. The MFO mathematical formulation is briefly introduced below.

First, the set of moths is expressed in a matrix as follows:

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,d} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ m_{n,1} & m_{n,2} & .. & m_{n,d} \end{bmatrix} \tag{14}$$

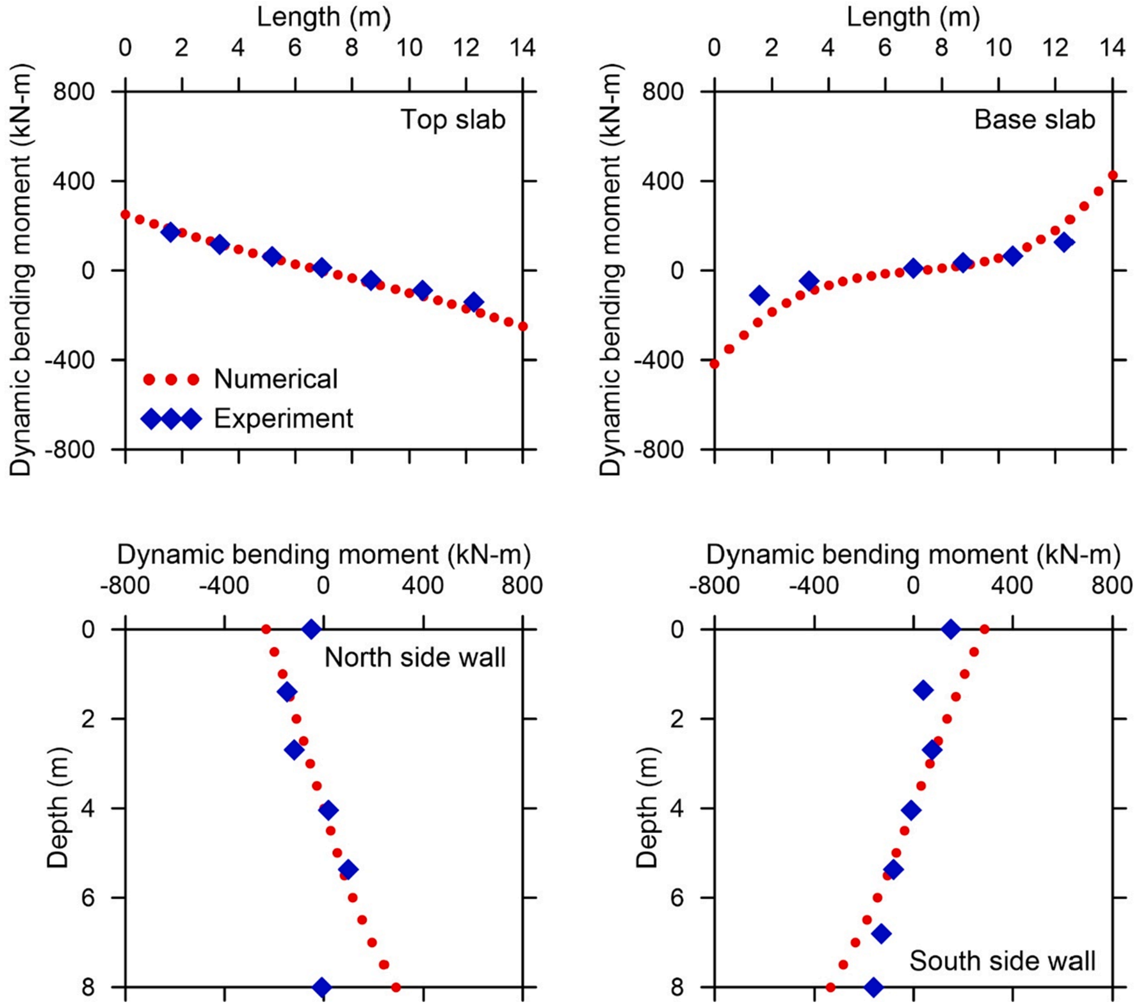

Fig. 10. The bending moment comparison between numerical results and centrifuge test data at the maximum bending moment step of each wall.

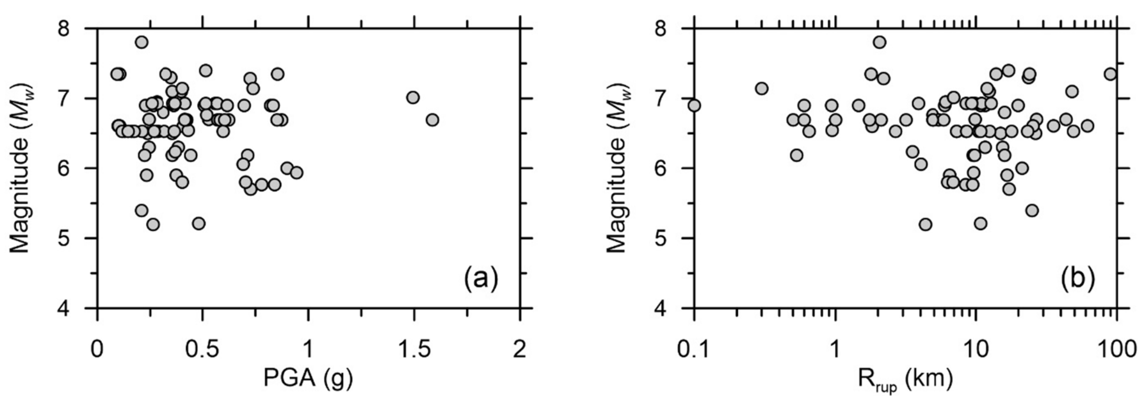

Fig. 11. Moment magnitude distribution of selected ground motions with (a) PGA and (b) rupture distance.

where d indicates the number of variables, n is the number of moths.

Second, the set of flames is expressed in a matrix as follows:

$$F = \begin{bmatrix} F_{1,1} & F_{1,2} & \cdots & F_{1,d} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ F_{n,1} & F_{n,2} & .. & F_{n,d} \end{bmatrix} \tag{15}$$

As the dimension of M and F is the same, there is an array that stores the corresponding fitness values as follows:

$$OM = \begin{bmatrix} OM_1 \\ \vdots \\ \vdots \\ OM_n \end{bmatrix} \text{ and } OF = \begin{bmatrix} OF_1 \\ \vdots \\ \vdots \\ OF_n \end{bmatrix} \tag{16}$$

In the MFO algorithm, moth and flame both are solutions. Flames are the best position for moths, and moths are the actual search agents that move around the search space. Each moth seeks around a flame and updates its position to find a better result using the equation below [84]:

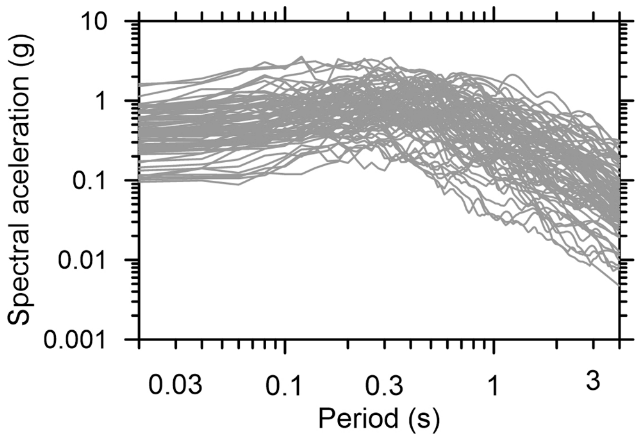

Fig. 12. Acceleration response spectra of the selected ground motions.

Table 5
Ground motion intensity measures selected for model development.

No.	Intensity measure (unit)	Notation	Reference
1	Peak ground acceleration (g)	PGA	Kramer [101]
2	Peak ground velocity (m/s)	PGV	Kramer [101]
3	Peak ground displacement (m)	PGD	Kramer [101]
4	Ratio of PGV/PGA (s)	PGV/ PGA	Kramer [101]
5	Root-mean-square of acceleration (g)	A_{rms}	Housner and Jennings [102]
6	Root-mean-square of velocity (m/s)	V_{rms}	Housner and Jennings [102]
7	Root-mean-square of displacement (m)	D_{rms}	Housner and Jennings [102]
8	Arias intensity (m/s)	I_a	Arias [103]
9	Characteristic intensity ($m^{1.5}/s^{2.5}$)	I_c	Park, Ang [104]
10	Specific energy density (m^2/s)	SED	–
11	Cumulative absolute velocity (m/s)	CAV	Kramer [101]
12	Acceleration spectrum intensity (g*s)	ASI	Housner [105]
13	Velocity spectrum intensity (m)	VSI	Housner [105]
14	Housner spectrum intensity (m)	HI	Housner [105]
15	Sustained maximum acceleration (g)	SMA	Nuttli [106]
16	Sustained maximum velocity (m/s)	SMV	Nuttli [106]
17	Effective peak acceleration (g)	EPA	Benjamin [107]
18	Spectral acceleration at T_1 (g)	$S_a(T_1)$	Shome, Cornell [108]
19	Spectral velocity at T_1 (m/s)	$S_v(T_1)$	–
20	Spectral displacement at T_1 (m)	$S_d(T_1)$	–
21	A95 parameter (g)	A_{95}	Sarma and Yang [109]
22	Predominant period (s)	T_p	Kramer [101]
23	Mean period (s)	T_m	Rathje, Abrahamson [110]

Table 6
Analysis case matrix for data generation purposes.

Aspect ratio, B/H	Buried depth, h (m)	Soil profiles	A total of ground motion records
1, 1.5, 2, 3	3, 6, 12	V_{s1}, V_{s2}	85

$$M_i = S(M_i, F_j) \tag{17}$$

where M_i is the i^{th} moth, F_j is the j^{th} flame, S is the spiral function. The main update mechanism of the moth is expressed as:

$$S(M_i, F_j) = D_i.e^{bt}.\cos(2\pi t) + F_j \tag{18}$$

where b is a constant, t is a random number $\in [-1,1]$, D_i is the distance of the i^{th} moth for the j^{th} flame. D_i is calculated as below:

$$D_i = |F_j - M_i| \tag{19}$$

It is important to note that iterations will gradually decrease the number of flames leading to balances in the exploitation and exploration of the search area. The number of flames is calculated as:

$$flameno = round\left(N - l^* \frac{(N-1)}{T}\right) \tag{20}$$

where N and l are the maximum numbers of flames and the current number of iterations, respectively, and T is the maximum iterations.

Numerical model and validation

In this section, an accurate numerical model is developed and validated against experimental results. This study performed numerical simulations to exhibit the tunnel and soil response using a two-dimensional finite-difference analysis program, FLAC^{2D} version 7.0 [85]. The numerical model is then validated against the measurements from the experiment carried out by Gillis [86].

Underground tunnel model

The overburden of the rectangular tunnel is 4 m. The cross-section dimensions of the tunnel are 14 m and 8 m in width and height, respectively. The thickness of the side wall, top, and bottom slabs is 0.8 m. The tunnel structure was modeled using beam elements with a length of 0.5 m. The input parameters used for the structural elements are listed in Table 1.

Soil domain model

The dimension of the computational model was set to 107 × 26 m (width × height) to simulate the experiment, as presented in Fig. 4. The soil medium was modeled using plane-strain quadrilateral elements. The shear wave velocity profile and properties of soil are discussed in section 4.6. The element size, $\Delta l = 0.5$ m, was selected based on the following recommendation of Kuhlemeyer and Lysmer [87]:

$$\Delta l \leq \frac{\lambda}{10} \frac{\lambda}{8} \tag{21}$$

where λ is the wavelength of propagated wave corresponding to maximum frequency of interest. The *Sig*3 model was employed to simulate the nonlinear behavior of soil. The model, which is available in the FLAC^{2D} program and has been widely used in previous studies [88–92], is defined as follows [85]:

$$M_s = \frac{\eta}{1 + \exp(-(L - x_o)/\theta)} \tag{22}$$

where M_s is the shear modulus reduction factor, L is $log(\gamma)$, γ is the shear strain, and x_o, η, and θ are the parameters of curve-fitting. The parameters of the *Sig*3 model were chosen to match the curves of [93], adjusted to fit the shear strength at the middle of each soil layer. The parameters for the Darendeli formulation used to generate the shear modulus reduction and damping curves are listed in Table 2.

The Rayleigh damping was used to model small strain damping, is expressed as follows [94]:

$$[C] = \alpha[M] + \beta[K] \tag{23}$$

where $[C]$ is the damping matrix, $[M]$ is the mass matrix, $[K]$ is the stiffness matrix, α and β are the Rayleigh coefficients, which are determined through:

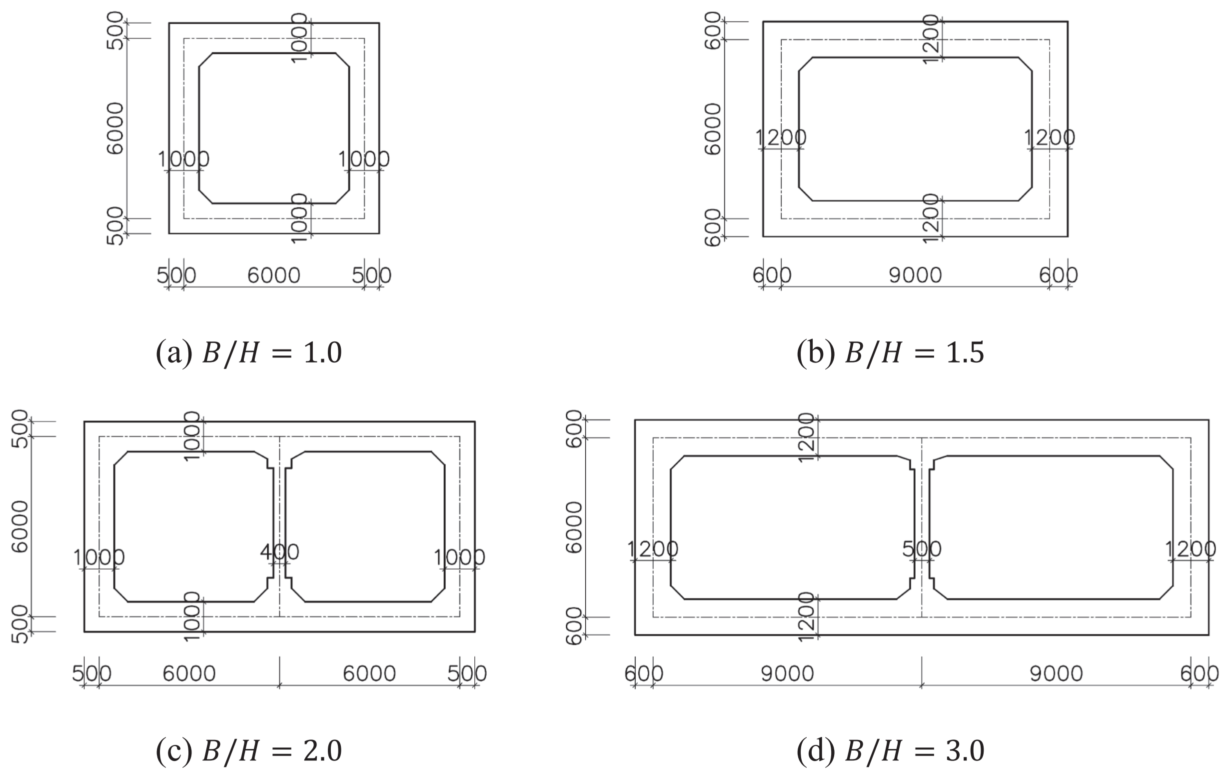

Fig. 13. Tunnel types used in the dynamic simulations.

$$\alpha = \frac{4\pi\xi f_m f_n}{f_m + f_n}$$

$$\beta = \frac{\xi}{\pi(f_m + f_n)} \tag{24}$$

where ξ is the damping ratio, f_m and f_n are the natural frequency at m^{th} and n^{th} modes, respectively. The Rayleigh coefficients α and β should be chosen such that the effect of the frequency-dependent damping is minimal [95]. In this study, 1st and 5th mode site frequencies were used for f_m and f_n based on the recommendation of Kwok, Stewart [95].

Soil-tunnel interface

The soil-structure interaction was simulated using the interface elements. The interface option *UNBONED* in the FLAC^{2D} program was used in this study. This contact interface can model a realistic partial-slip condition [96], considering the gapping and the slipping phenomena between soil and tunnel under loading. Parameters for the interface element include normal and shear springs stiffness (K_n and K_s). As recommended in the FLAC^{2D} manual [85], K_n and K_s are calculated as follows:

$$K_n = K_s = 10\max\left[\frac{K^{int} + \frac{4}{3}G^{int}_{max}}{\Delta Z_{min}}\right] \tag{25}$$

where K^{int} and G^{int}_{max} are the bulk and shear modulus of the stiffest neighboring zone, respectively, and ΔZ_{min} is the smallest width of an adjoining zone in the normal direction. The max[-] notation implies that the maximum value over all zones adjacent to the interface. Considering the tunnel material properties as the stiffest neighbor zone in equation (25), K_n and K_s values are greater than 10^{11} Pa/m. A large value of K_n and K_s should not be used due to increased analysis time [85]. Therefore, the obtained values of K_n and K_s were reduced to 10^{10} Pa/m in this study [97]. This reduction did not significantly affect the results, whereas analysis time was dramatically shortened. Properties of the interface elements are shown in Table 3.

Boundary condition

The free-field boundary was applied for lateral boundaries to absorb reflected waves. The bottom boundary was fixed to simulate the rigid

Table 7
Statistical characteristics of input parameters.

Parameters	Min	Max	Mean	SD	COV
B/H	1	3	1.875	0.740	0.394
h(m)	3	12	7	3.472	0.535
F	0.057	7.439	1.883	1.637	0.870
PGA (g)	0.093	1.585	0.547	0.277	0.606
PGV (m/s)	0.050	1.480	0.513	0.339	0.662
PGD (m)	0.011	1.765	0.213	0.245	1.150
PGV/PGA (s)	0.031	0.384	0.123	0.068	0.549
A_{rms} (g)	0.014	0.157	0.061	0.033	0.550
V_{rms} (m/s)	0.008	0.329	0.089	0.064	0.717
D_{rms} (m)	0.002	0.618	0.056	0.079	1.408
I_a (m/s)	0.117	11.822	2.498	2.544	1.018
I_c ($m^{1.5}/s^{2.5}$)	0.010	0.342	0.096	0.073	0.767
SED (m^2/s)	0.003	5.875	0.468	0.805	1.718
CAV (m/s)	2.662	35.988	11.065	6.653	0.601
ASI (g*s)	0.083	1.074	0.366	0.202	0.550
VSI (m)	0.190	5.801	1.787	1.160	0.649
HI (m)	0.161	5.845	1.660	1.130	0.681
SMA (g)	0.063	0.734	0.310	0.164	0.527
SMV (m/s)	0.037	0.859	0.313	0.190	0.607
EPA (g)	0.085	1.644	0.445	0.268	0.602
S_a (T_1) (g)	0.157	3.274	0.945	0.573	0.606
S_v (T_1) (m/s)	7.235	169.620	45.692	29.999	0.657
S_d (T_1) (m)	0.278	10.257	2.747	2.062	0.751
A_{95} (g)	0.090	1.573	0.451	0.275	0.609
T_p (s)	0.040	1.240	0.378	0.204	0.538
T_m (s)	0.148	1.583	0.618	0.260	0.421
R(output)	0.043	3.252	1.012	0.594	0.587

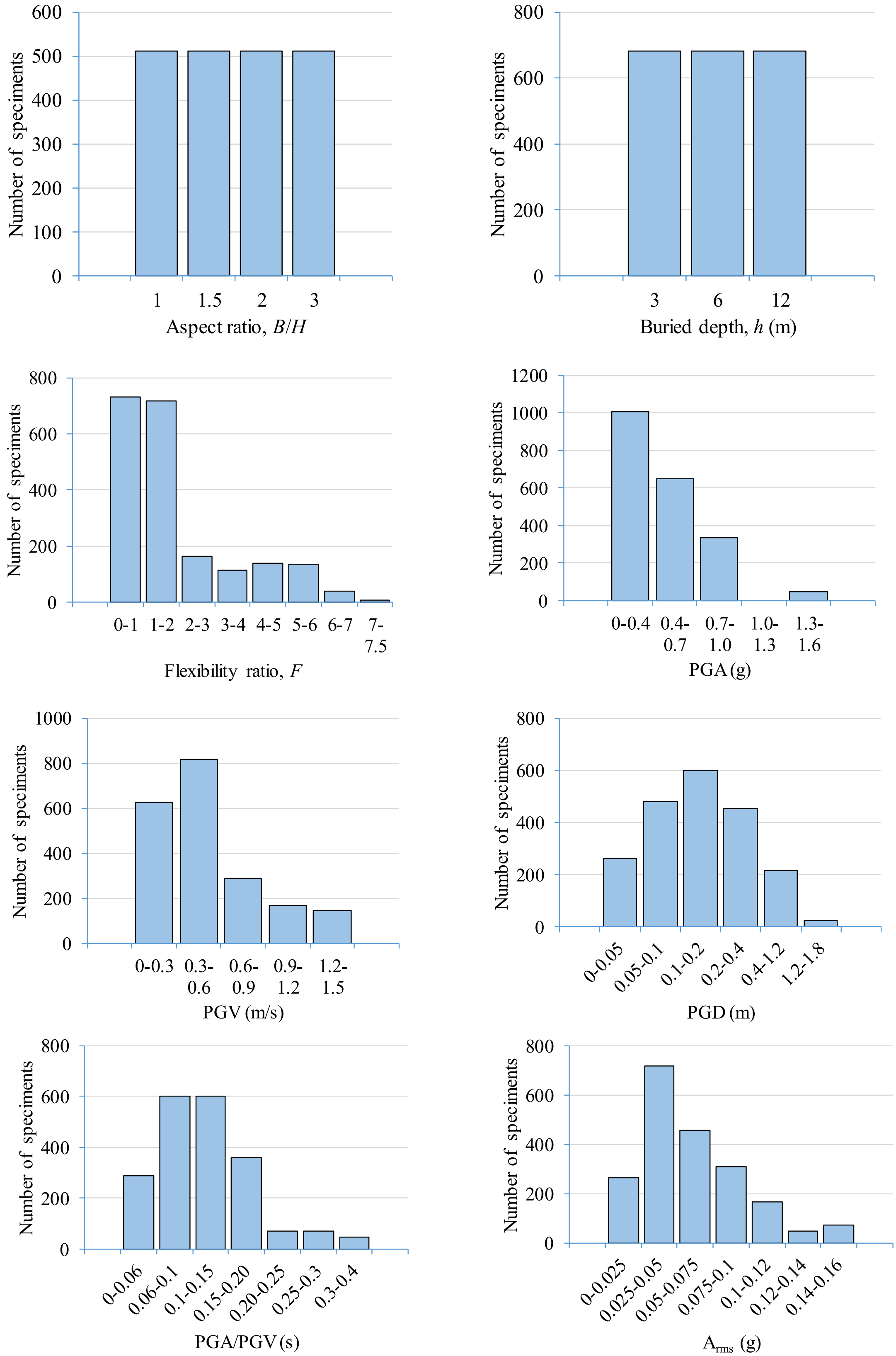

Fig. 14. Histograms of input parameters.

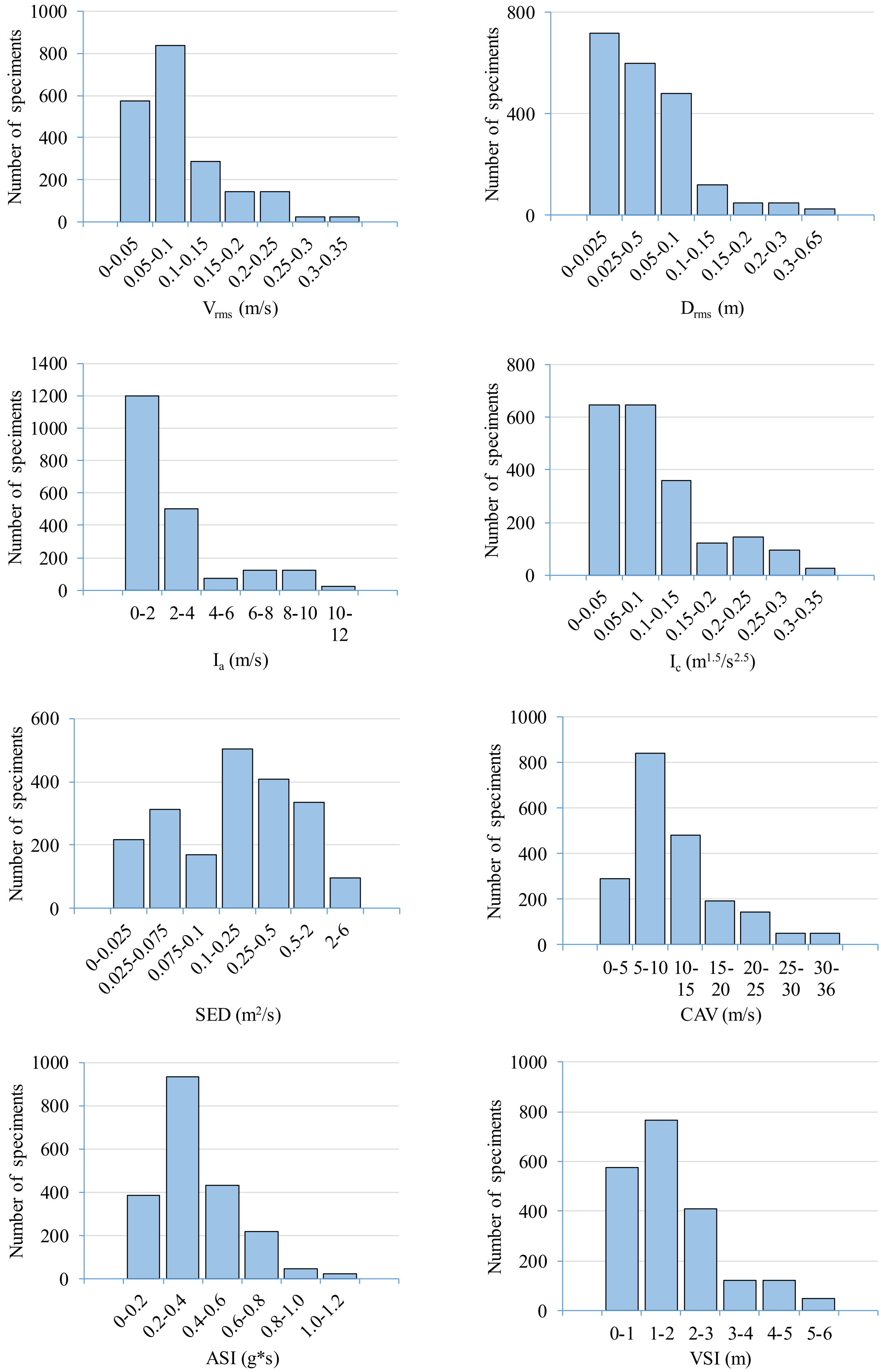

Fig. 14. (*continued*).

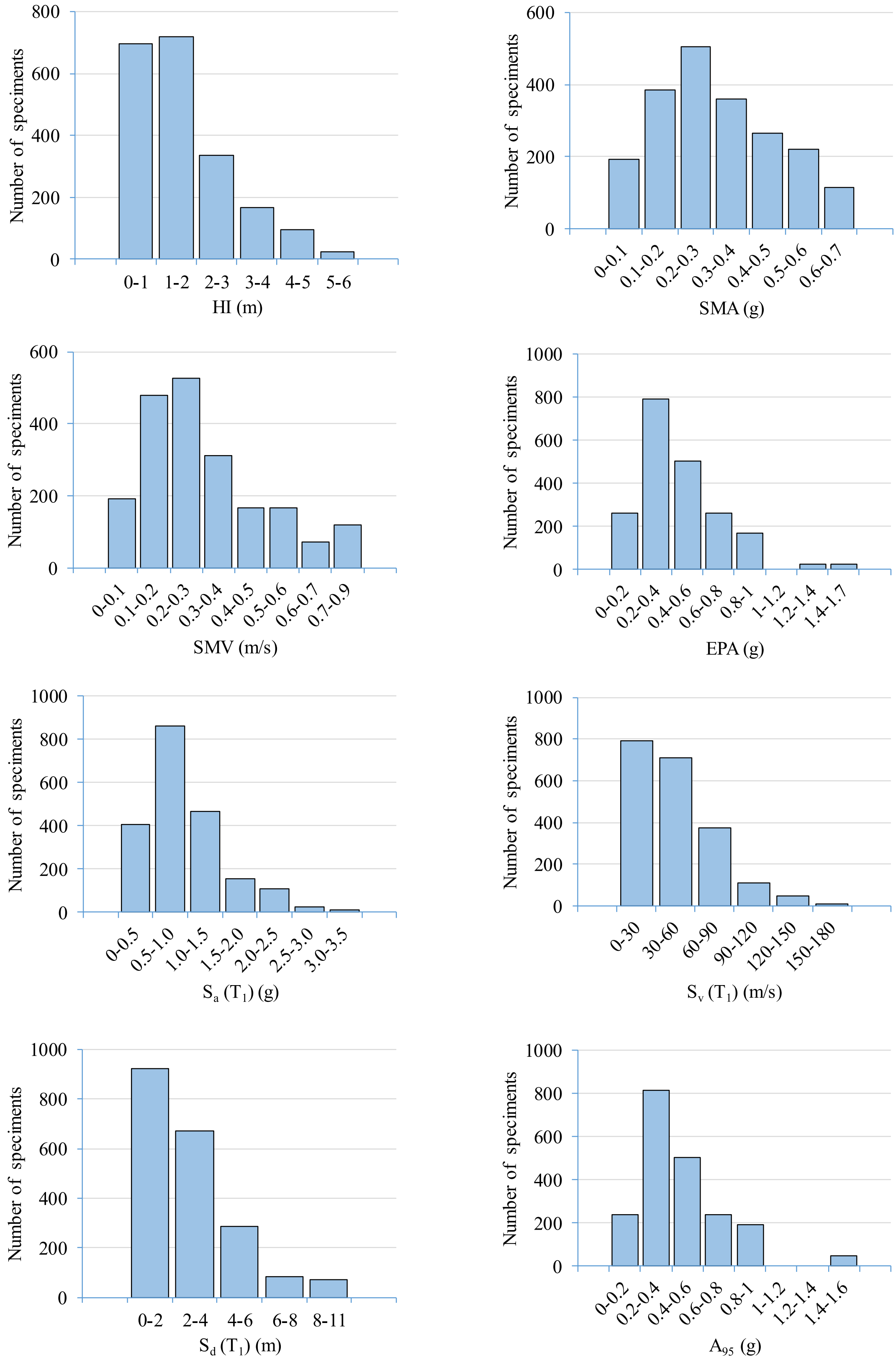

Fig. 14. (*continued*).

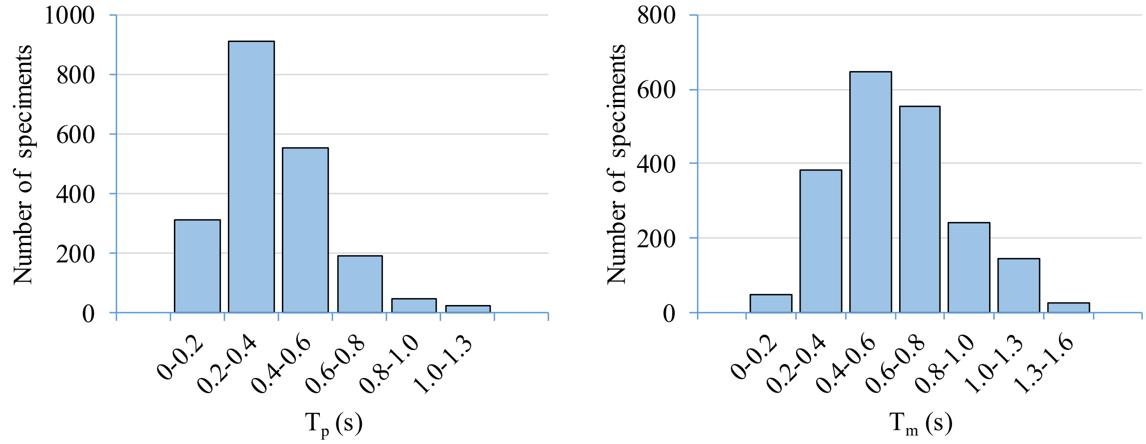

Fig. 14. (*continued*).

boundary used in the experiment. The acceleration time history of the input motion was defined at the base of the numerical model.

Input ground motion

The acceleration time history and response spectra of input motion are shown in Fig. 5. The properties of input motion such as peak ground acceleration (PGA), predominant period (T_p), arias intensity (I_a), significant duration (D_{5-95}) are summarized in Table 4.

Validation of the numerical model

The centrifuge test conducted by Gillis [86] was used to validate the numerical model. The model of the centrifuge test is presented in Fig. 6 with a 1/65 scale. Accelerometers and strain gauges were installed to measure the acceleration and bending moment time history of the tunnel structure. Uniform medium-dense dry Nevada sand (mean grain sizeD_{50} = 0.14, uniformity coefficient C_u = 2.07, minimum void ratio e_{min} = 0.53, maximum void ratio e_{max} = 0.9, specific gravity G_s = 2.66) was adopted in the centrifuge test. The sand was dry pluviated into the centrifuge container to reach a relative density of 55 %. The unit weight of sand was 15.3 kN/m^3. Due to small-strain shear wave velocities were measured at two depth levels (8 m and 21.3 m), the following power law [98] was used to develop shear wave velocity profile:

$$G_{max} = G_o\left(\frac{\sigma_o}{P_a}\right)^{0.5} \tag{26}$$

where G_{max} is the maximum shear modulus, G_o is the fitting parameter, σ_o is the mean effective stress, and P_a is the atmospheric pressure. The shear wave soil profile of the centrifuge test and numerical model are shown in Fig. 7 as V_{s1}. For the data generation purpose, which is presented in section 5, an additional homogeneous soil profile with a unit weight of 15.3 kN/m^3 was adopted, denoted as V_{s2} in Fig. 7. This profile is stiffer than the centrifuge profile (V_{s1}). For V_{s1} and V_{s2} profiles, effective friction angles applied were 30° and 35°, respectively. The cohesion was set to zero, considering the sand deposit used in this study. The Poisson's ratio of 0.3 was used for both soil profiles.

The centrifuge model of the tunnel structure was made from aluminum to represent a reinforced concrete tunnel. Tunnel properties of the design and centrifuge test model are presented in Table 1. The input motion shown in Fig. 5 was used in the centrifuge test. Further details on the experiment are described in the study of Gillis [86]. Digitized data is available in Gillis, Dashti [99] (https://doi.org/10.4231/D3JQ0SW10).

Comparisons between numerical results and centrifuge test records are presented in Figs. 8–10. The results from the numerical analysis were extracted at depths corresponding to those recorded in the centrifuge test. Fig. 8 compares free-field response spectra at selected depths (A22-A27). The calculated responses provide exceptional fits with the recordings at all depths. Fig. 9 presents the calculated and measured spectra acceleration comparison at the top (A07), middle (A06), and bottom (A05) of the tunnel. As can be seen from the figure, calculated and measured tunnel responses also produce favorable agreements with the recordings. The comparison of the calculated and measured bending moments of the tunnel are shown in Fig. 10. The results of numerical analysis closely match the centrifuge recording except at the slab and wall connections. One possible reason is that the connections were simulated as rigid in the numerical model, whereas they are not perfectly rigid in the centrifuge model [100]. The data shown in Fig. 8 and Fig. 10 demonstrate that the numerical model captures the dynamic response of the centrifuge test properly, and it is reliable for performing parametric investigations.

Data generation

Although the measured data from the experiment can be considered to be the most accurate, it is difficult to cover a wide range of soil-tunnel layouts and motions with the centrifuge model tests. Therefore, a series of dynamic analyses were performed using the validated numerical model with different soil-tunnel configurations to develop a database for building the MFO-XGBoost model. The dimensions of the rectangular tunnel cross-sections were varied such that the aspect ratio (B/H) ranges from 1 to 3. A center column with a spacing of 3.0 m was used for tunnels with aspect ratios greater than 2 to match the typical tunnel design. The cross-section of center column are 0.4 m × 1.0 m and 0.5 m × 1.0 m for B/H = 2.0 and B/H = 3.0, respectively. The thickness of slabs and side walls is 1.0 m for B/H = 1.0 and B/H = 2.0, and 1.2 m for B/H = 1.5 and B/H = 3.0. The tunnel structures modeled are shown in Fig. 11. Three buried depths (i.e. 3, 6, and 12 m) were used to represent both the shallow and deep box tunnels. Two shear wave velocity profiles were used, as displayed in Fig. 7. A total of 85 earthquake ground motions were selected from the NGA-west2 database (https://ngawest2.berkeley.edu). The peak ground acceleration (PGA) varies from 0.093 to 1.585g. The moment magnitude (M_w) and rupture distance (R_{rup}) range from 5.2 to 7.8 and 0.1 to 89.76 km, respectively, as displayed in Fig. 11. The response spectra of selected ground motions are shown in Fig. 12. The ground motions cover a wide range of 23 intensity measures (IMs) to consider the effect of the earthquake on the tunnel response. The details of 23 IMs are summarized in Table 5. Notably, T_1 in this table is the fundamental period of the tunnel structure. Due to the tunnel structures are enclosed within the soil domain, the fundamental period of the structure is approximate to the natural period of the soil profile. The analysis case matrix is presented in Table 6. It should be noted that the

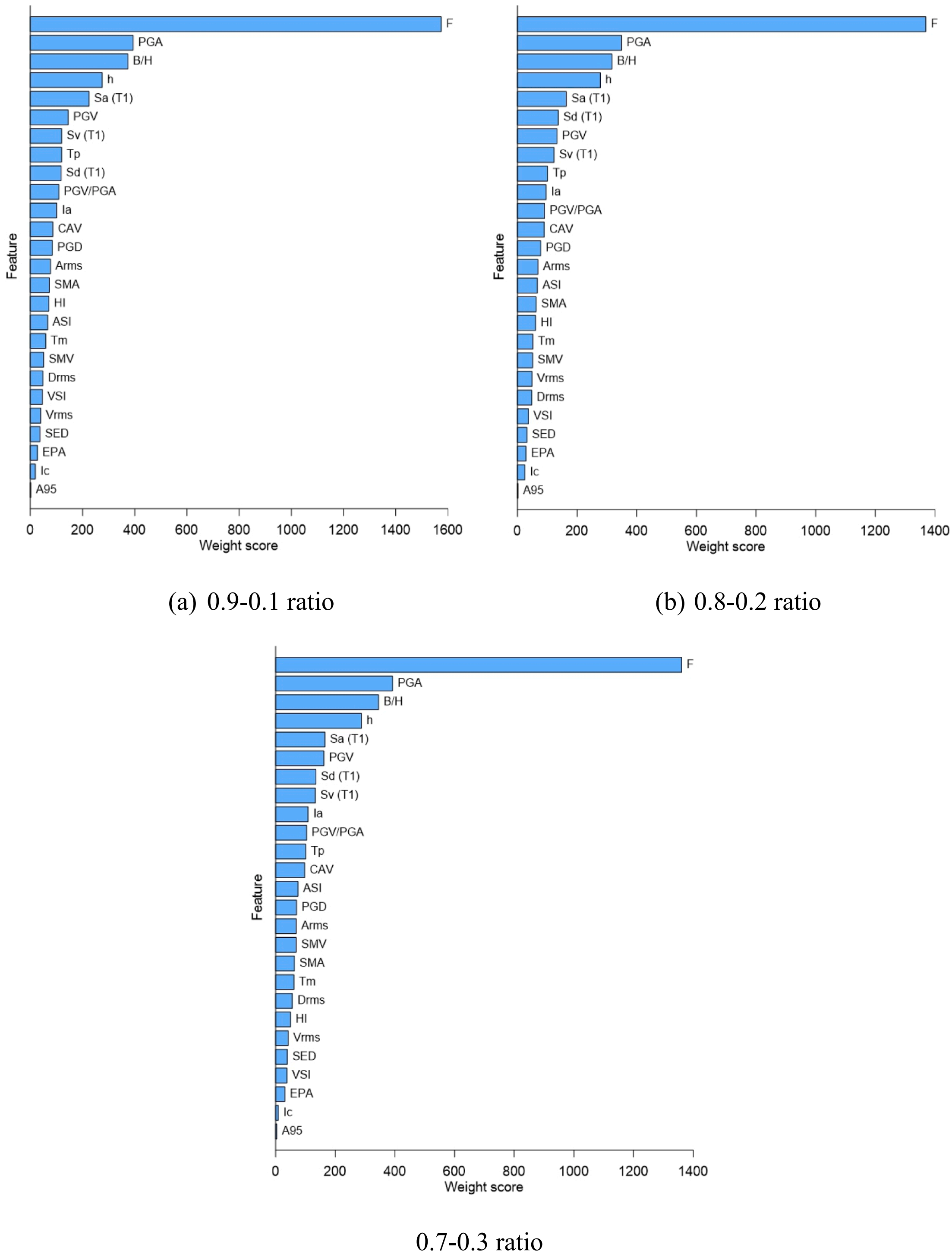

Fig. 15. Feature selection.

numerical model is only validated to one tunnel-soil profile layout. Because the fundamental seismic tunnel-soil-interface interaction mechanism is expected not to vary with tunnel dimension and soil profile, other tunnel-soil configurations were not validated against model test measurements (see Fig. 13).

A total of 2040 data points was obtained from numerical analyses. It should be noted that the equivalent shear modulus (G_m) at the mid-height of the tunnel used in equation (1) was calculated as follows:

$$G_m = \gamma_{eff} G_{max} = 0.65\gamma_{ff} G_{max} \tag{27}$$

where γ_{eff} is the effective free-field shear strain at mid-height of the tunnel. The maximum free-field soil shear strain (γ_{ff}) was determined for each motion by dividing the free-field maximum racking displacement by the structure height ($\gamma_{ff} = \Delta_{ff}/H$). Δ_{sm} in equation (7) is calculated as follows [29,111]:

$$\Delta_{sm} = \frac{d_1^2 - d_2^2}{4B} \tag{28}$$

where d_1 and d_2 are the lengths of two diagonal edges of the deformed tunnel section at the maximum distortion time step during ground

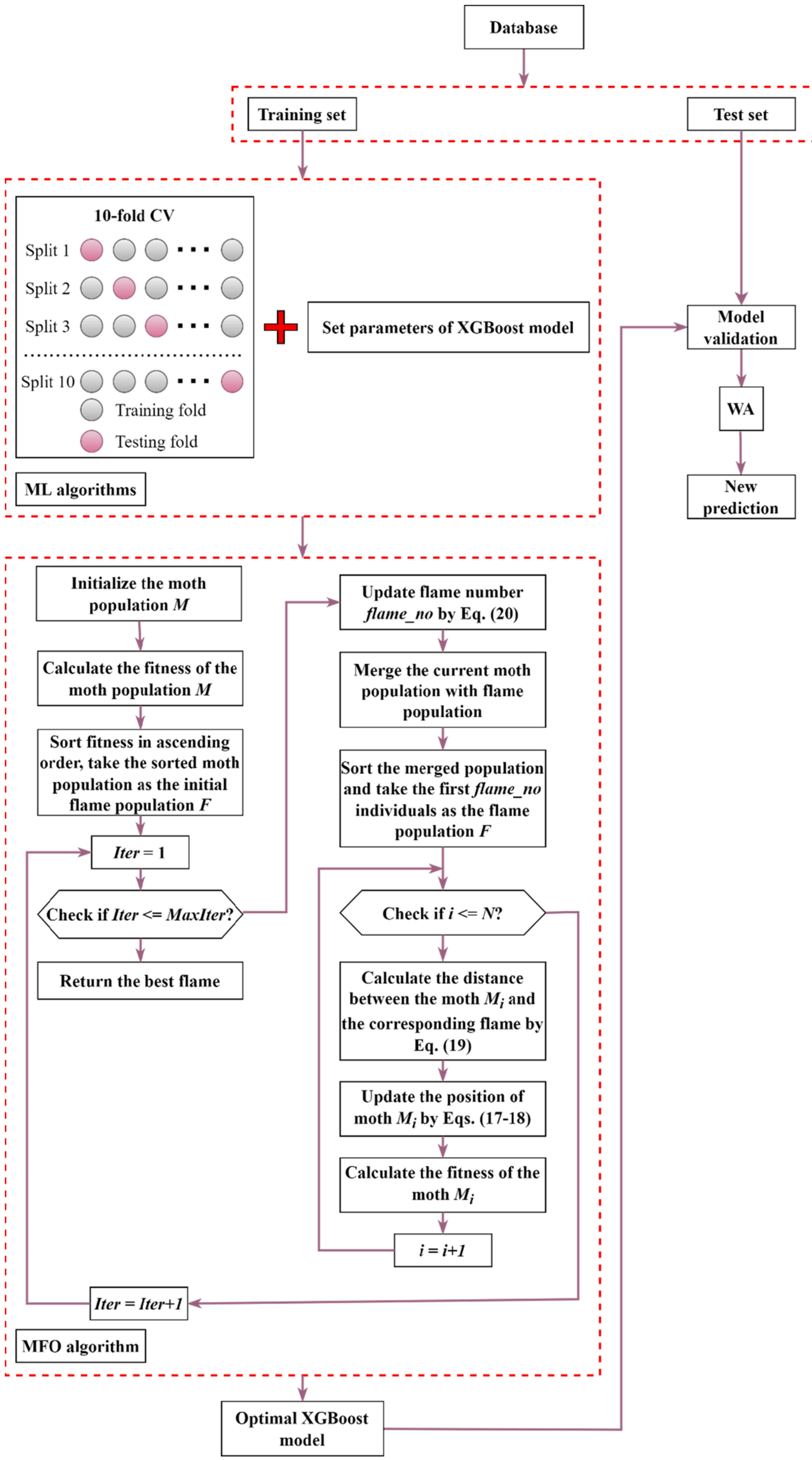

Fig. 16. Flowchart for developing MFO-XGBoost models to predict the racking ratio.

Table 8
Crucial parameters and their ranges of the XGBoost model.

Model	Parameter and range				
XGBoost	gama	learning_rate	max_delta_step	max_depth	min_child_weight
	(0.0–1.0)	(0.01–1.0)	(1–10)	(1–10)	(0.0–1.0)
	n_estimators	reg_alpha	reg_lambda	subsample	
	(5–100)	(0.0–1.0)	(0.0–1.0)	(0.0–1.0)	

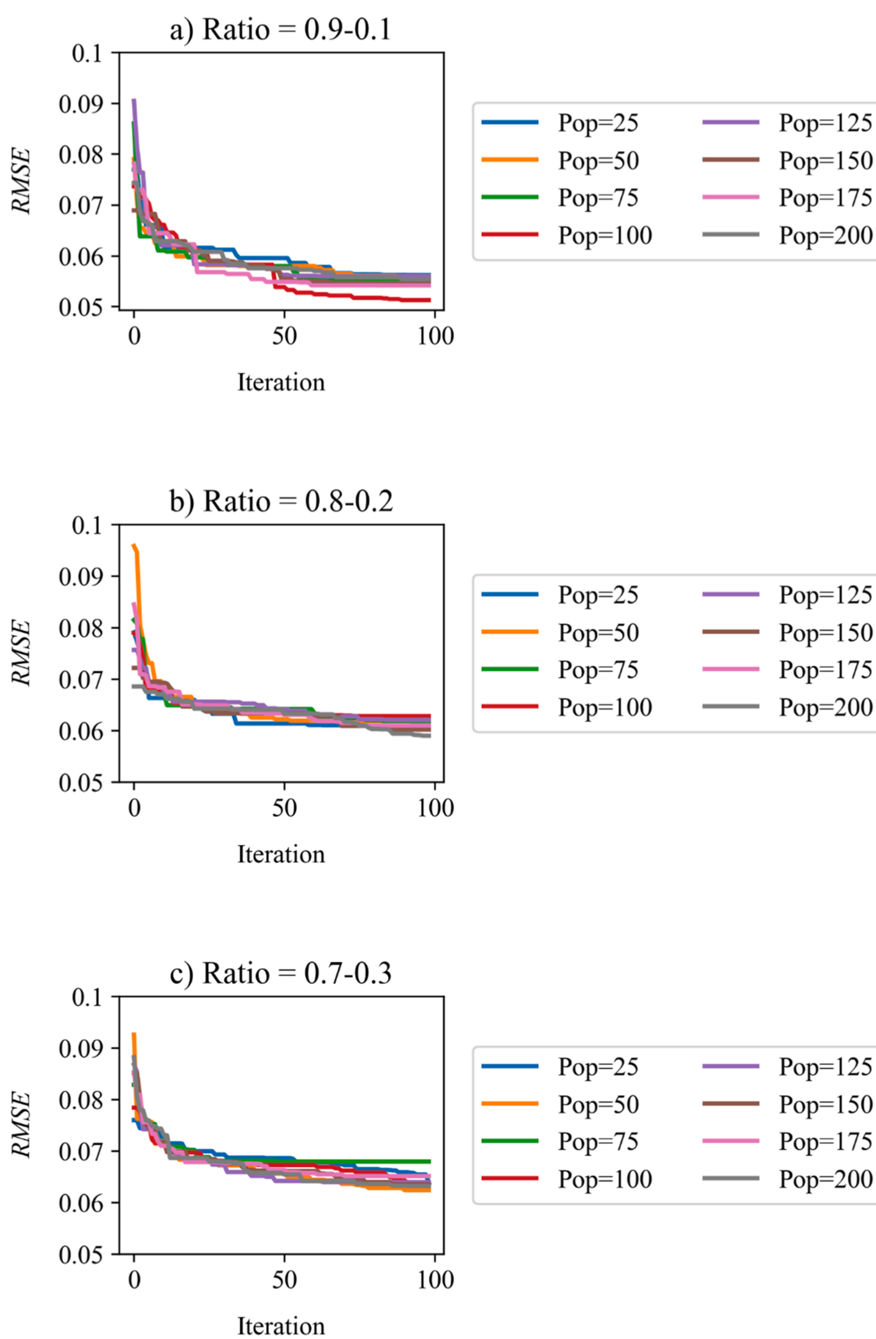

Fig. 17. Convergence curve of MFO-XGBoost models with different population sizes.

shaking, as presented in Fig. 1c-d.

The statistical characteristics of input parameters are listed in Table 7. The selection of input parameters to develop the ML model is discussed in section 7. Notably, SD and COV are the standard deviation and the coefficient of variation, respectively. Fig. 14 depicts input parameter histograms based on 2040 data points.

Performance measures

Coefficient of determination (R^2), root mean square error (*RMSE*), and mean absolute error (*MAE*) are used to evaluate the performance of the predictive models. The formulation of these parameters are expressed as:

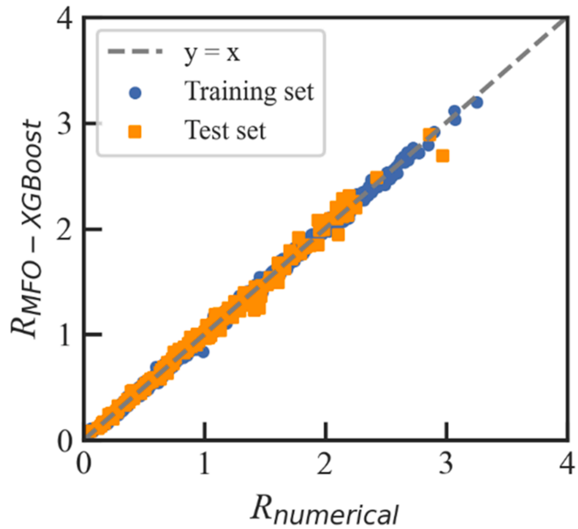

Fig. 18. Regression of the proposed MFO-XGBoost model.

Table 9
Performance of the MFO-XGBoost model with different training-test data ratios.

Ratio	Pop size	Training data			Test data		
		R^2	*RMSE*	*MAE*	R^2	*RMSE*	*MAE*
0.9–0.1	100	0.998	0.023	0.017	0.992	0.051	0.034
0.8–0.2	200	0.998	0.029	0.021	0.990	0.059	0.039
0.7–0.3	50	0.998	0.023	0.017	0.989	0.062	0.041

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(t_i - o_i)^2}{\sum_{i=1}^{n}(t_i - \bar{t})^2} \tag{29}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(t_i - o_i)^2} \tag{30}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|t_i - o_i| \tag{31}$$

where t_i is the target value of i^{th} sample, o_i is the output value of i^{th} sample, $\bar{t}$ is the average of actual values, and n is the number of samples.

The value of R^2 is used to assess the degree of correlation between the observed and predicted data. Meanwhile, the value of *RMSE* represents the error size. Finally, *MAE* measures the error of the predictive method. Generally, a higher value of R^2 and lower values of *RMSE* and *MAE* indicate a good performance of the model. An expected value of $R^2 = 1.0$ represents a perfect case.

Training MFO-XGBoost model

Feature selection is a technique of data preprocessing. Its idea is to choose a number of input parameters that significantly affect the results using predefined evaluation criteria. This reduces the computation time, thus increasing model efficiency.

Fig. 15 shows the results of the feature selection analysis of R by the XGBoost algorithm. This figure ranks the importance of features by calculating their weight scores during the model training. Three training–testing ratios (i.e. 0.9–0.1, 0.8–0.2, 0.7–0.3 ratios) are considered. A higher value of the weight score indicates a more important feature. As can be seen from the figure, F has the most significant influence on R, followed by PGA, B/H ratio, h, S_a (T_1), and PGV. Meanwhile, EPA, I_c, and A_{95} are revealed to be the least important.

A series of trials were performed to evaluate which set of input parameters provide both accurate and efficient predictions. It is demonstrated that use of twelve most influential input parameters provide almost identical results compared with those incorporating all twenty six parameters. These input parameters are F, PGA, B/H ratio, h, S_a (T_1), PGV, S_v (T_1), T_p, S_d (T_1), PGV/PGA, I_a, CAV. The selected input parameters are in agreement with the findings of Wang [20], Penzien [21], Tsinidis and Pitilakis [29], Zhang, Zhao [112], Nguyen, Thusa [113], Du and Wang [114], and Zhang, Shokrabadi [115].

This section establishes MFO-XGBoost models to find the relationship between the input variables (i.e. F, PGA, B/H ratio, h, S_a (T_1), PGV, S_v (T_1), T_p, S_d (T_1), PGV/PGA, I_a, CAV) and an output variable (i.e. racking ratio, R). The flowchart for developing the MFO-XGBoost models to predict the racking ratio is shown in Fig. 16. Firstly, the database was generated and divided into training and test sets. Secondly, the MFO algorithm and cross-validation technique are implemented to optimize the hyper-parameters of the XGBoost model using the training set. Then, the MFO-XGBoost model's performance is evaluated using the test set. Finally, the web application (WA) is developed based on the MFO-XGBoost model for new prediction. The following section introduces detailed descriptions of the procedure.

To speed up the learning process and improve the accuracy, the input and output variables are standardized in the range (−1,1), which are presented as follows:

$$X^N = 2 \times \frac{(X - X_{min})}{(X_{max} - X_{min})} - 1 \tag{32}$$

where X is the original sample, X^N is the normalized sample, X_{min} and X_{max} are the minimum and maximum values of each variable, respectively.

To consider the effect of training and test partitions, three training and test ratios (i.e. 0.9–0.1, 0.8–0.2, 0.7–0.3 ratios), idential to those used in the feature selection process, are applied. The training set is used to find the optimal parameters using the XGBoost model integrated with the MFO algorithm, while the performance of the models is then evaluated using the test set. Moreover, cross-validation (CV) is used to ensure that the ML model performs favorably on unseen data (test data) and avoid overfitting. This method randomly divides the training data

Table 11
Prediction accuracy of the different models.

No	Predicted model	R^2	*RMSE*	*MAE*(%)
1	Penzien [21] no-slip	0.897	0.273	0.23
2	Penzien [21] full-slip	0.891	0.329	0.290
3	Anderson [22]	0.854	0.291	0.235
4	Zhang and Liu [23]	0.854	0.308	0.210
5	MFO-XGBoost	0.998	0.028	0.019

Table 10
Optimal parameters of the MFO-XGBoost model.

Model	Optimal parameters				
MFO-XGBoost	gama	learning_rate	max_delta_step	max_depth	min_child_weight
	0.00168	0.36299	5.672232	6	0.09939
	n_estimators	reg_alpha	reg_lambda	subsample	
	100	0.22554	0.79266	0.95628	

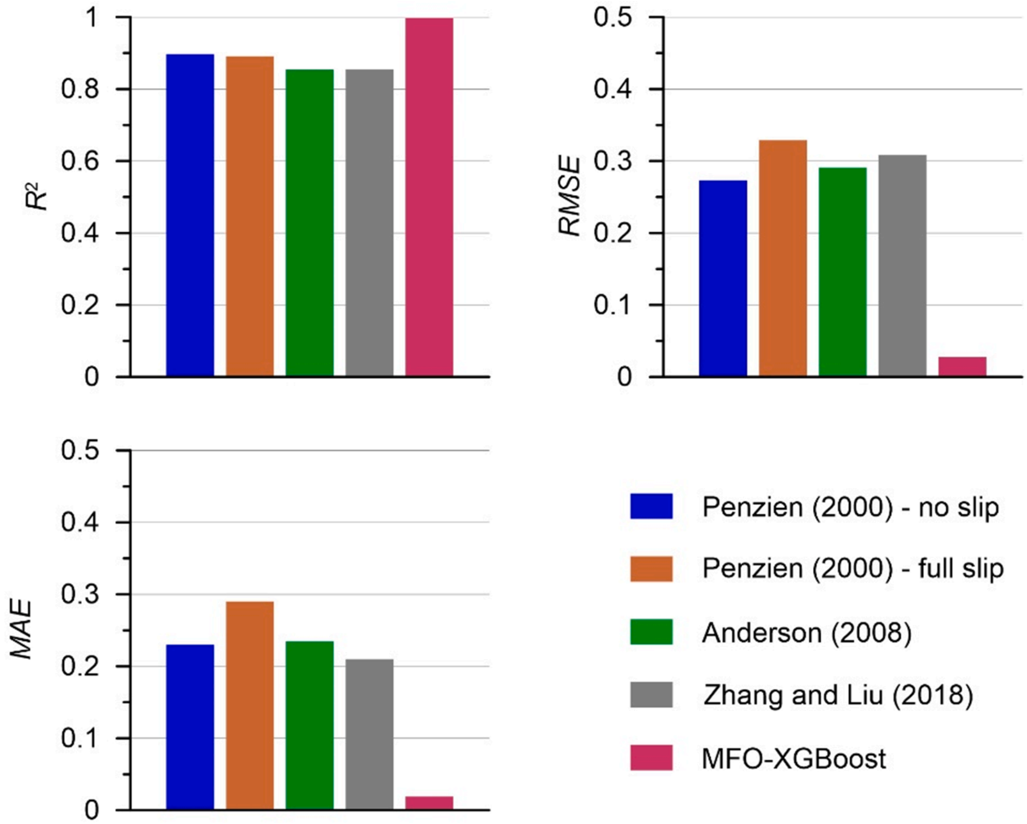

Fig. 19. Comparison of prediction accuracy parameters of models.

Table 12
Statistical analysis for the ratio of predicted to numerical results.

No.	Predicted model	Statistical properties of $R_{\text{prediction}}/R_{\text{numerical}}$		
		SD	Mean	COV
1	Penzien [21] no-slip	0.286	1.306	0.219
2	Penzien [21] full-slip	0.325	1.398	0.232
3	Anderson [22]	0.383	1.282	0.299
4	Zhang and Liu [23]	0.335	1.121	0.299
5	MFO-XGBoost	0.047	1.002	0.047

into k subsamples, called folds, of roughly equal size. Therefore, the first model is estimated using $k-1$ folds as a training dataset, and the remaining fold of training data is used to calculate the prediction accuracy metric. The accuracy is then expressed as an average accuracy acquired by the k models in k validation rounds. Therefore, the average *RMSE* of 10 folds is used as the objective function for the optimization process in this study.

The main hyperparameters that need to be tuned for the XGBoost model are:

- gama: is minimum loss reduction, which is required to make a further partition on a leaf node of the tree

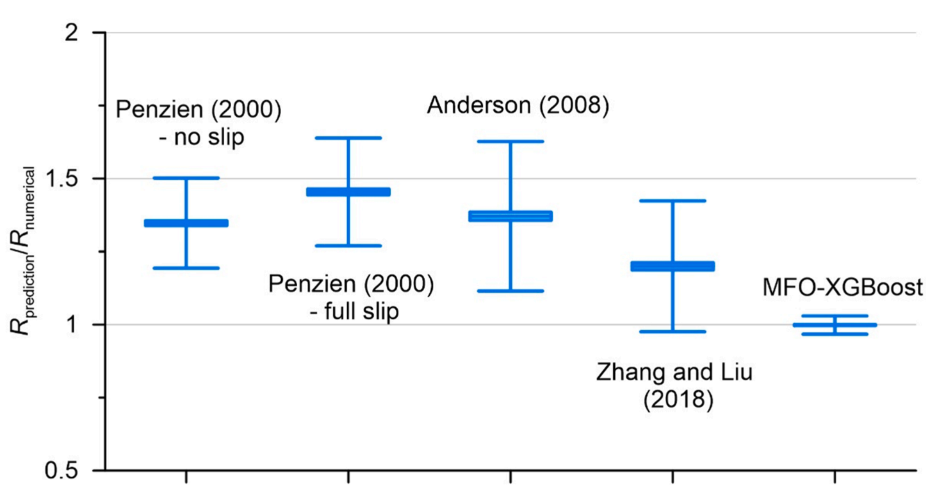

Fig. 20. Deviation of the ratio of predicted to numerical results.

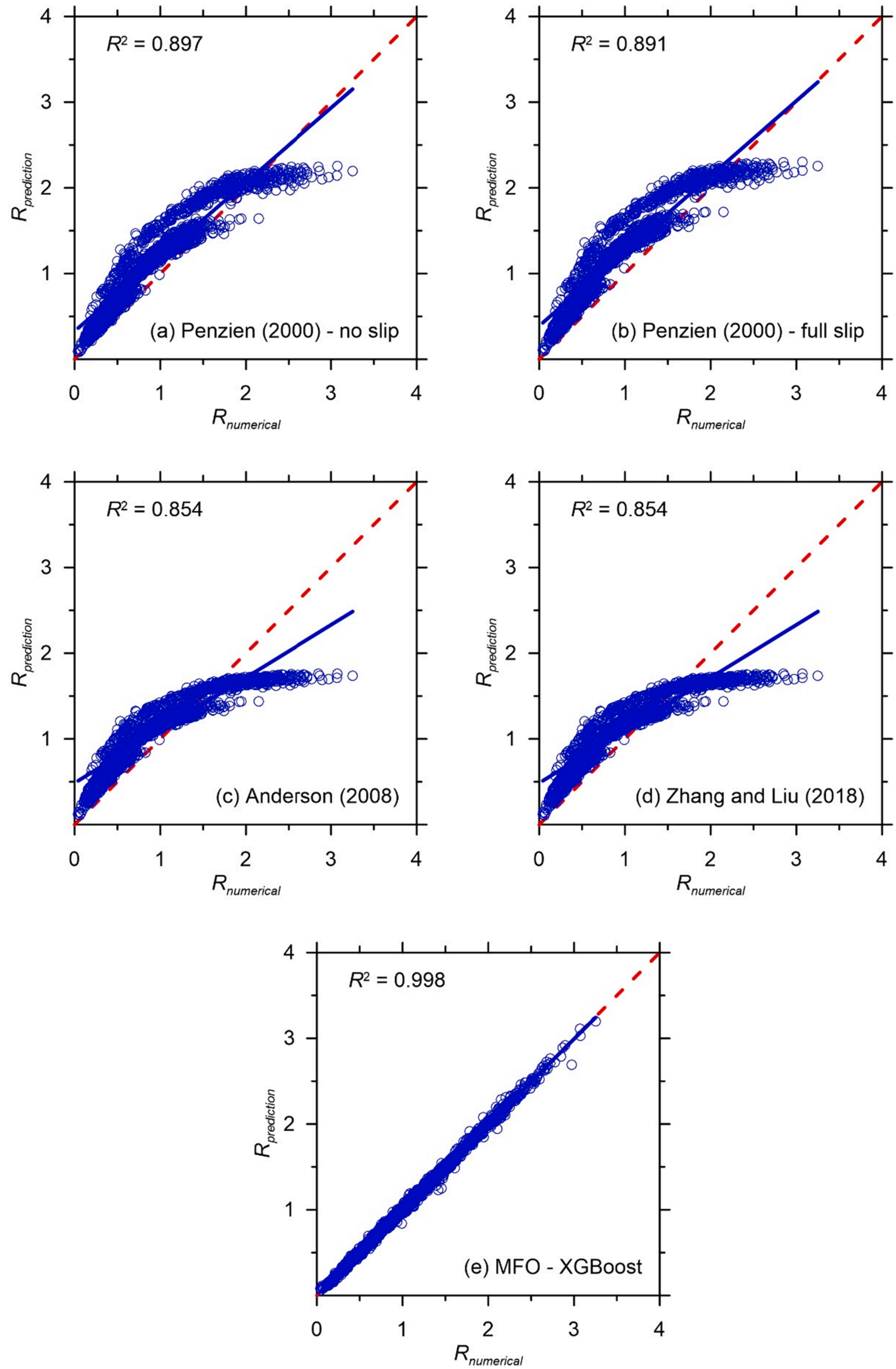

Fig. 21. Comparison racking ratio between numerical results and predicted methods.

- learning_rate: is step size shrinkage, which is used to prevent overfitting
- max_delta_step: allows estimating each tree's weight
- max_depth: is the maximum depth of a tree
- min_child_weight: is the minimum sum of instance weight (hessian) needed in a child
- n_estimators: is the number of gradient boosted trees
- reg_alpha and reg_lambda: are regularization terms for weights

- subsample: is the subsample ratio of the training instances

The ranges of these parameters are listed in Table 8.

After a series of tests, it is found that after 100 iterations, the fitness values become stable. Increasing the number of iterations causes corresponding increment in the calculation time, without an improvement in the prediction accuracy. Therefore, the number of iterations is set to 100 in this study. Moreover, several population sizes (i.e. 25, 50, 75, 100, 125, 150, 175, and 200) are selected for the optimization process. Fig. 17 shows the convergence curves of the MFO-XGBoost models for each population size. Table 9 shows the best model performances of three training–testing ratios.

The results show that three cases perform well in training and test data. However, the 0.9–0.1 ratio with a population size of 100 provides the best prediction with the highest value of R^2 and lowest values of *RMSE* and *MAE* for both training and test data. The R^2, *RMSE*, and *MAE* values for the training data are 0.998, 0.023, and 0.017, respectively. For the test data, the R^2, *RMSE*, and *MAE* values are 0.992, 0.051, and 0.034, respectively. Therefore, the training-test ratio of 0.9–0.1 is chosen in this study. The optimal parameters of the MFO-XGBoost model are presented in Table 10.

The results of the proposed MFO-XGBoost model for predicting the racking ratio are shown in Fig. 18.

Results and discussions

Table 11 and Fig. 19 shows the values of three indicators of different models. Accordingly, the coefficient of determination of the MFO-XGBoost model is the highest ($R^2 = 0.998$) compared with those of Penzien [21] no-slip ($R^2 = 0.897$), Penzien [21] full-slip ($R^2 = 0.891$), Anderson [22] ($R^2 = 0.854$), and Zhang and Liu [23] ($R^2 = 0.854$), respectively. Moreover, the *RMSE* and *MAE* values of the MFO-XGBoost model are much lower and smallest compared with the others. Notably, in Table 11, because of considering the effect of earthquake IMs, the MFO-XGBoost model used more input variables than the other models.

The ratios of the previous study to numerical results were calculated and statistically analyzed to compare the prediction performance. The statistical analysis results, including mean, standard deviation (SD), and the coefficient of variation (COV), are listed in Table 12. Accordingly, the mean value of the MFO-XGBoost model is closer to 1.0 than those of three previous studies. The coefficients of variation are 0.219, 0.232, 0.299, 0.299 and 0.047, for Penzien [21] no-slip, Penzien [21] full-slip, Anderson [22], Zhang and Liu [23], and MFO-XGBoost model, respectively. Fig. 20 presents the deviation of the ratios from different models. As shown in the figure, the ratio results of the existing models are more scattered than that of the MFO-XGBoost model.

Fig. 21 (a-e) presents the comparison of the predicted racking ratio of exiting models and numerical results. The dashed line (i.e. the 1:1 line) represents the target, while the solid line indicates the linear regression line of the scatters. The closer scattering to the dashed line, the higher accuracy of the predicted results. As can be seen from the figure, the results of the MFO-XGBoost model are the closest to the numerical results, followed by Penzien [21] no-slip, Penzien [21] full-slip, Anderson [22], and Zhang and Liu [23]. The difference in the existing model can be attributed to the number of input parameters for calculating the racking ratio. In particular, Anderson [22], and Zhang and Liu [23] only considered the effect of the flexibility ratio on the racking ratio, while they are the flexibility ratio and Poisson's ratio in the study of Penzien [21]. Significantly, twelve input parameters mentioned in section 7, are examed in the MFO-XGBoost model.

Overall, the MFO-XGBoost model provides a higher precision than the existing models. Among four considered empirical formulas, the Penzien [21] no-slip presents the best performance for calculating the racking ratio.

Web application

This section presents the development of a web application for predicting the racking ratio based on the proposed MFO-XGBoost model. The WA requires twelve input parameters: *F*, PGA, *B/H* ratio, *h*, S_a (T_1), PGV, S_v (T_1), T_p, S_d (T_1), PGV/PGA, I_a, and CAV to obtain the racking ratio. It is worth mentioning that ground motions IMs are automatically calculated using SeismoSigmal program (Seismosoft, 2012). This web application is available by using the following link: https://tvl-rackingratio.herokuapp.com. Notably, the results obtained from the proposed MFO-XGBoost model and the web application are identical. However, the web based model is easier to use than the proposed MFO-XGBoost model. Therefore, the web application is recommended for practical engineering applications to predict the racking ratio. However, it should be noted that they are only applicable to determine the racking ratio for input values within the ranges specified in Table 7.

Conclusions

This study presents the development of a novel hybrid MFO-XGBoost model to predict the racking ratio of rectangular tunnels subjected to seismic loading, which is the primary demand for performing a pseudo-static frame analysis. A total of 2040 numerical simulations were generated to train and test the MFO-XGBoost model. Based on the feature selection, 12 variables, including *F*, PGA, *B/H* ratio, *h*, S_a (T_1), PGV, S_v (T_1), T_p, S_d (T_1), PGV/PGA, I_a, CAV, were considered as input variables of the proposed MFO-XGBoost model. The performance of the proposed MFO-XGBoost model was examined and evaluated against four empirical models. The following findings are yielded from this study.

(1) The proposed MFO-XGBoost model is capable of favorably predicting the racking ratio for rectangular tunnels. The coefficients of determination (R^2) were 0.998, 0.992, and 0.998 for the training, testing, and all data, respectively.
(2) Compared with four empirical models, the proposed MFO-XGBoost model is revealed to provide the most accurate predictions of the racking ratio. The performance of the model was verified in terms of statistical properties (i.e., R^2, *RMSE*, *MAE*, SD, mean value, and COV). Among existing models, the Penzien [21] no-slip case provided the most agreeable prediction of the racking ratio.
(3) The parameters *F*, PGA, *B/H* ratio, *h*, and S_a (T_1) have primary influence on the calculated racking ratio among 12 input variables, whereas I_a and CAV have secondary effects.
(4) A web application is developed for possible application in routine practice for a more convenient estimate of the racking ratio.

CRediT authorship contribution statement

Van-Quang Nguyen: Conceptualization, Software, Writing – original draft. **Viet-Linh Tran:** Software, Data curation. **Duy-Duan Nguyen:** Conceptualization, Software. **Shamsher Sadiq:** Data curation, Software. **Duhee Park:** Conceptualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

This work was partly supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2020-0-01373, Artificial Intelligence Graduate School Program (Hanyang University)), the research fund of Hanyang University (HY-202000000790002), and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2022R1A2C3003245).

References

[1] Arango I. Theme paper: earthquake engineering for tunnels and underground structures. a case history. Geotech Earthquake Eng and Soil Dynamics IV2008. p. 1-34.
[2] Hashash YM, Hook JJ, Schmidt B, John I, Yao C. Seismic design and analysis of underground structures. Tunn Undergr Space Technol 2001;16(4):247–93.
[3] Dowding CH, Rozan A. Damage to rock tunnels from earthquake shaking. J Geotech Eng Div 1978;104(2):175–91.
[4] Iida H, Hiroto T, Yoshida N, Iwafuji M. Damage to Daikai subway station. Soils Found 1996;36(Special):283–300.
[5] Wang W, Wang T, Su J, Lin C, Seng C, Huang T. Assessment of damage in mountain tunnels due to the Taiwan Chi-Chi earthquake. Tunn Undergr Space Technol 2001;16(3):133–50.
[6] Ghasemi H, Cooper JD, Imbsen R, Piskin H, Inal F, Tiras A. The November 1999 Duzce Earthquake: post-earthquake investigation of the structures on the TEM. Rep No FHWA-RD-00 2000;146(00–146).
[7] Tsinidis G. Response characteristics of rectangular tunnels in soft soil subjected to transversal ground shaking. Tunn Undergr Space Technol 2017;62:1–22.
[8] Abate G, Massimino MR, Maugeri M. Numerical modelling of centrifuge tests on tunnel–soil systems. Bull Earthq Eng 2015;13(7):1927–51.
[9] Xu H, Li T, Xia L, Zhao JX, Wang D. Shaking table tests on seismic measures of a model mountain tunnel. Tunn Undergr Space Technol 2016;60:197–209.
[10] Wang G, Yuan M, Miao Yu, Wu J, Wang Y. Experimental study on seismic response of underground tunnel-soil-surface structure interaction system. Tunn Undergr Space Technol 2018;76:145–59.
[11] Rabeti Moghadam M, Baziar MH. Seismic ground motion amplification pattern induced by a subway tunnel: shaking table testing and numerical simulation. Soil Dyn Earthq Eng 2016;83:81–97.
[12] Sayed MA, Kwon O-S, Park D, Van Nguyen Q. Multi-platform soil-structure interaction simulation of Daikai subway tunnel during the 1995 Kobe earthquake. Soil Dyn Earthq Eng 2019;125:105643.
[13] Wang YX, Shan SB, Zhang C, Guo PP. Seismic response of tunnel lining structure in a thick expansive soil stratum. Tunn Undergr Space Technol 2019;88:250–9.
[14] Nguyen V-Q, Nizamani ZA, Park D, Kwon O-S. Numerical simulation of damage evolution of Daikai station during the 1995 Kobe earthquake. Eng Struct 2020;206:110180.
[15] Nguyen D-D, Lee T-H, Nguyen V-Q, Park D. Seismic damage analysis of box metro tunnels accounting for aspect ratio and shear failure. Appl Sci 2019;9(16):3207.
[16] Nguyen D-D, Park D, Shamsher S, Nguyen V-Q, Lee T-H. Seismic vulnerability assessment of rectangular cut-and-cover subway tunnels. Tunn Undergr Space Technol 2019;86:247–61.
[17] Huo H, Bobet A, Fernández G, Ramírez J. Analytical solution for deep rectangular structures subjected to far-field shear stresses. Tunn Undergr Space Technol 2006;21(6):613–25.
[18] Bobet A, Fernandez G, Huo H, Ramirez J. A practical iterative procedure to estimate seismic-induced deformations of shallow rectangular structures. Can Geotech J 2008;45(7):923–38.
[19] Bobet A. Drained and undrained response of deep tunnels subjected to far-field shear loading. Tunn Undergr Space Technol 2010;25(1):21–31.
[20] Wang J-N. Seismic design of tunnels: a simple state-of-the-art design approach: parsons brinckerhoff, 1993.
[21] Penzien J. Seismically induced racking of tunnel linings. Earthquake Eng 2000;29(5):683–91.
[22] Anderson DG. Seismic analysis and design of retaining walls, buried structures, slopes, and embankments. Transp Res Board 2008:105–29.
[23] Zhang L, Liu Y, Lantsoght EOL. Seismic responses of rectangular subway tunnels in a clayey ground. PLoS ONE 2018;13(10):e0204672.
[24] Cilingir U, Madabhushi SG. A model study on the effects of input motion on the seismic behaviour of tunnels. Soil Dyn Earthq Eng 2011;31(3):452–62.
[25] Cilingir U, Madabhushi SG. Effect of depth on the seismic response of square tunnels. Soils Found 2011;51(3):449–57.
[26] Tsinidis G, Pitilakis K, Madabhushi G, Heron C. Dynamic response of flexible square tunnels: centrifuge testing and validation of existing design methodologies. Geotechnique 2015;65(5):401–17.
[27] Abuhajar O, El Naggar H, Newson T. Experimental and numerical investigations of the effect of buried box culverts on earthquake excitation. Soil Dyn Earthq Eng 2015;79:130–48.
[28] Tsinidis G, Rovithis E, Pitilakis K, Chazelas JL. Seismic response of box-type tunnels in soft soil: experimental and numerical investigation. Tunn Undergr Space Technol 2016;59:199–214.
[29] Tsinidis G, Pitilakis K. Improved R-F relations for the transversal seismic analysis of rectangular tunnels. Soil Dyn Earthq Eng 2018;107:48–65.
[30] Zhang W, Gu X, Tang L, Yin Y, Liu D, Zhang Y. Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: comprehensive review and future challenge. Gondwana Res 2022;109:1–17.
[31] Jong S, Ong D, Oh E. State-of-the-art review of geotechnical-driven artificial intelligence techniques in underground soil-structure interaction. Tunn Undergr Space Technol 2021;113(103946).
[32] Phoon K-K, Zhang W. Future of machine learning in geotechnics. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards 2022:1–16.
[33] Zhang W, Phoon K-K. Editorial for Advances and applications of deep learning and soft computing in geotechnical underground engineering. J Rock Mech Geotech Eng 2022.
[34] Novellino A, Cesarano M, Cappelletti P, Di Martire D, Di Napoli M, Ramondini M, et al. Slow-moving landslide risk assessment combining Machine Learning and InSAR techniques. CATENA 2021;203(105317).
[35] Wang H, Zhang L, Yin K, Luo H, Li J. Landslide identification using machine learning. Geosci Front 2021;12(1):351–64.
[36] Stanley T, Kirschbaum D, Sobieszczyk S, Jasinski M, Borak J, Slaughter S. Building a landslide hazard indicator with machine learning and land surface models. Environ Modell Software 2020;129(104692).
[37] Pu Y, Apel DB, Hall R. Using machine learning approach for microseismic events recognition in underground excavations: Comparison of ten frequently-used models. Eng Geol 2020;268(105519).
[38] Zhang R, Wu C, Goh AT, Böhlke T, Zhang W. Estimation of diaphragm wall deflections for deep braced excavation in anisotropic clays using ensemble learning. Geosci Front 2021;12(1):365–73.
[39] Zhang W, Zhang R, Wu C, Goh AT, Wang L. Assessment of basal heave stability for braced excavations in anisotropic clay using extreme gradient boosting and random forest regression. Underground Space 2020.
[40] Goh AT, Zhang Y, Zhang R, Zhang W, Xiao Y. Evaluating stability of underground entry-type excavations using multivariate adaptive regression splines and logistic regression. Tunn Undergr Space Technol 2017;70(148–54).
[41] Zhang W, Zhang R, Wang W, Zhang F, Goh ATC. A Multivariate Adaptive Regression Splines model for determining horizontal wall deflection envelope for braced excavations in clays. Tunn Undergr Space Technol 2019;84(461–71).
[42] He X, Xu H, Sabetamal H, Sheng D. Machine learning aided stochastic reliability analysis of spatially variable slopes. Comput Geotech 2020;126(103711).
[43] Pham K, Kim D, Park S, Choi H. Ensemble learning-based classification models for slope stability analysis. Catena 2021;196(104886).
[44] Weidner L, Walton G, Kromer R. Classification methods for point clouds in rock slope monitoring: A novel machine learning approach and comparative analysis. Eng Geol 2019;263(105326).
[45] Hoang N-D, Pham A-D. Hybrid artificial intelligence approach based on metaheuristic and machine learning for slope stability assessment: A multinational data analysis. Expert Syst Appl 2016;46(60–8).
[46] Chen S, Gu C, Lin C, Wang Y, Hariri-Ardebili MA. Prediction, monitoring, and interpretation of dam leakage flow via adaptative kernel extreme learning machine. Measurement 2020;166(108161).
[47] Pourghasemi HR, Yousefi S, Sadhasivam N, Eskandari S. Assessing, mapping, and optimizing the locations of sediment control check dams construction. Sci Total Environ 2020;739(139954).
[48] Zhou Y, Zhang Y, Pang R, Xu B. Seismic fragility analysis of high concrete faced rockfill dams based on plastic failure with support vector machine. Soil Dyn Earthq Eng 2021;144(106587).
[49] Hariri-Ardebili MA, Barak S. A series of forecasting models for seismic evaluation of dams based on ground motion meta-features. Eng Struct 2020;203(109657).
[50] Khandelwal M, Marto A, Fatemi SA, Ghoroqi M, Armaghani DJ, Singh T, et al. Implementing an ANN model optimized by genetic algorithm for estimating cohesion of limestone samples. Eng Comput 2018;34(2):307–17.
[51] Debnath P, Dey AK. Prediction of laboratory peak shear stress along the cohesive soil–geosynthetic interface using artificial neural network. Geotech Geol Eng 2017;35(1):445–61.
[52] Bejarbaneh BY, Bejarbaneh EY, Fahimifar A, Armaghani DJ, Abd Majid MZ. Intelligent modelling of sandstone deformation behaviour using fuzzy logic and neural network systems. Bull Eng Geol Environ 2018;77(1):345–61.
[53] Armaghani DJ, Mohamad ET, Hajihassani M, Yagiz S, Motaghedi H. Application of several non-linear prediction tools for estimating uniaxial compressive strength of granitic rocks and comparison of their performances. Eng Comput 2016;32(2):189–206.
[54] Wang Y, Zhao T. Bayesian assessment of site-specific performance of geotechnical design charts with unknown model uncertainty. Int J Numer Anal Methods Geomech 2017;41(5):781–800.
[55] Shi C, Wang Y. Non-parametric machine learning methods for interpolation of spatially varying non-stationary and non-Gaussian geotechnical properties. Geosci Front 2021;12(1):339–50.
[56] Wang Y, Hu Y, Zhao T. Cone penetration test (CPT)-based subsurface soil classification and zonation in two-dimensional vertical cross section using Bayesian compressive sampling. Can Geotech J 2020;57(7):947–58.
[57] Ghaderi A, Shahri AA, Larsson S. An artificial neural network based model to predict spatial soil type distribution using piezocone penetration test data (CPTu). Bull Eng Geol Environ 2019;78(6):4579–88.
[58] Shen J, Jimenez R. Predicting the shear strength parameters of sandstone using genetic programming. Bull Eng Geol Environ 2018;77(4):1647–62.

[59] Soleimani S, Jiao P, Rajaei S, Forsati R. A new approach for prediction of collapse settlement of sandy gravel soils. Eng Comput 2018;34(1):15–24.
[60] Harandizadeh H, Toufigh MM, Toufigh V. Application of improved ANFIS approaches to estimate bearing capacity of piles. Soft Comput 2019;23(19):9537–49.
[61] Ebrahimian B, Movahed V. Application of an evolutionary-based approach in evaluating pile bearing capacity using CPT results. Ships Offshore Structures 2017;12(7):937–53.
[62] Singh T, Pal M, Arora V. Modeling oblique load carrying capacity of batter pile groups using neural network, random forest regression and M5 model tree. Front Struct Civil Eng 2019;13(3):674–85.
[63] Moayedi H, Hayati S. Artificial intelligence design charts for predicting friction capacity of driven pile in clay. Neural Comput Appl 2019;31(11):7429–45.
[64] Goh A, Zhang W, Zhang Y, Xiao Y, Xiang Y. Determination of earth pressure balance tunnel-related maximum surface settlement: a multivariate adaptive regression splines approach. Bull Eng Geol Environ 2018;77(2):489–500.
[65] Zhou J, Shi X, Du K, Qiu X, Li X, Mitri HS. Feasibility of random-forest approach for prediction of ground settlements induced by the construction of a shield-driven tunnel. Int J Geomech 2017;17(6):04016129.
[66] Zhang W, Li H, Wu C, Li Y, Liu Z, Liu H. Soft computing approach for prediction of surface settlement induced by earth pressure balance shield tunneling. Undergr Space 2020.
[67] Zheng G, Yang P, Zhou H, Zeng C, Yang X, He X, et al. Evaluation of the earthquake induced uplift displacement of tunnels using multivariate adaptive regression splines. Comput Geotech 2019;113(103099).
[68] Wang P, Wang S, Jierula A. Automatic identification and location of tunnel lining cracks. Adv Civ Eng Mater 2021;2021.
[69] Zhang W, Li Y, Wu C, Li H, Goh A, Liu H. Prediction of lining response for twin tunnels constructed in anisotropic clay using machine learning techniques. Underground Space 2020.
[70] Wang R, Li D, Chen EJ, Liu Y. Dynamic prediction of mechanized shield tunneling performance. Autom Constr 2021;132(103958).
[71] Luat N-V, Han SW, Lee K. Genetic algorithm hybridized with eXtreme gradient boosting to predict axial compressive capacity of CCFST columns. Compos Struct 2021;278(114733).
[72] Duan J, Asteris PG, Nguyen H, Bui X-N, Moayedi H. A novel artificial intelligence technique to predict compressive strength of recycled aggregate concrete using ICA-XGBoost model. Eng Comput 2021;37(4):3329–46.
[73] Zhang W, Wu C, Zhong H, Li Y, Wang L. Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. Geosci Front 2021;12(1):469–77.
[74] Nayak J, Naik B, Dash PB, Souri A, Shanmuganathan V. Hyper-parameter tuned light gradient boosting machine using memetic firefly algorithm for hand gesture recognition. Appl Soft Comput 2021;107(107478).
[75] Tran V-L, Nguyen D-D. Novel hybrid WOA-GBM model for patch loading resistance prediction of longitudinally stiffened steel plate girders. Thin Wall Struct 2022;177(109424).
[76] Tsinidis G, de Silva F, Anastasopoulos I, Bilotta E, Bobet A, Hashash YM, et al. Seismic behaviour of tunnels: From experiments to analysis. Tunn Undergr Space Technol 2020;99(103334).
[77] Zhu T, Wang R, Zhang J-M. Evaluation of various seismic response analysis methods for underground structures in saturated sand. Tunn Undergr Space Technol 2021;110(103803).
[78] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. San Francisco California USA2016. p. 785-94.
[79] Nguyen HD, Dao ND, Shin M. Machine learning-based prediction for maximum displacement of seismic isolation systems. J Build Eng 2022;51(104251).
[80] Nguyen HD, LaFave JM, Lee Y-J, Shin M. Rapid seismic damage-state assessment of steel moment frames using machine learning. Eng Struct 2022;252(113737).
[81] Nguyen HD, Dao ND, Shin M. Prediction of seismic drift responses of planar steel moment frames using artificial neural network and extreme gradient boosting. Eng Struct 2021;242(112518).
[82] Nguyen HD, Truong GT, Shin M. Development of extreme gradient boosting model for prediction of punching shear resistance of r/c interior slabs. Eng Struct 2021;235(112067).
[83] Nguyen NV, Nguyen HD, Dao ND. Machine learning models for predicting maximum displacement of triple pendulum isolation systems. Structures: Elsevier; 2022. p. 404–15.
[84] Mirjalili S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowl-Based Syst 2015;89(228–49).
[85] Itasca Consulting Group. FLAC - Fast Lagrange Analysis of Continua; Version7.0. In: User Manual; 2011.
[86] Gillis KM. Seismic Response of Shallow Underground Structures in Dense Urban Environments. University of Colorado at Boulder; 2015 [Ph.D. dissertation]:.
[87] Kuhlemeyer RL, Lysmer J. Finite element method accuracy for wave propagation problems. J Soil Mech Found Div. 1973;99(Tech Rpt).
[88] Lee J-H, Ahn J-K, Park D. Prediction of seismic displacement of dry mountain slopes composed of a soft thin uniform layer. Soil Dyn Earthq Eng 2015;79(5–16).
[89] Lu C-C, Hwang J-H. Implementation of the modified cross-section racking deformation method using explicit FDM program: a critical assessment. Tunn Undergr Space Technol 2017;68(58–73).
[90] Callisto L, Ricci C. Interpretation and back-analysis of the damage observed in a deep tunnel after the 2016 Norcia earthquake in Italy. Tunn Undergr Space Technol 2019;89(238–48).
[91] Sanderson ST. Non-Linear Deformation Analysis of an Embankment Dam. Portland State University; 2018 [Master thesis]:.
[92] Park K-C, Nguyen V-Q, Kim J-H, Park D, Choi B-H. Estimation of seismically-induced crest settlement of earth core rockfill dams. Appl Sci 2019;9(20):4343.
[93] Darendeli MB. Development of a new family of normalized modulus reduction and material damping curves [Ph.D. Dissertation]. Austin, TX, USA: University of Texas at Austin, 2001.
[94] Chopra A. Structural dynamics: Theory and applications to earthquake engineering. 2nd ed. Prentice Hall; 2001. p. 455–8.
[95] Kwok AO, Stewart JP, Hashash YM, Matasovic N, Pyke R, Wang Z, et al. Use of exact solutions of wave propagation problems to guide implementation of nonlinear seismic ground response analysis procedures. J Geotech Geoenviron Eng 2007;133(11):1385–98.
[96] Hashash YM, Park D, John I, Yao C. Ovaling deformations of circular tunnels under seismic loading, an update on seismic design and analysis of underground structures. Tunn Undergr Space Technol 2005;20(5):435–41.
[97] Baziar MH, Moghadam MR, Kim D-S, Choo YW. Effect of underground tunnel on the ground surface acceleration. Tunn Undergr Space Technol 2014;44(10–22).
[98] Richart Jr FE. Some effects of dynamic soil properties on soil-structure interaction. J Geotech Geoenviron Eng 1975;101(ASCE# 11764 Proceeding).
[99] Gillis K, Dashti S, Jones C. Test-1: Seismic response of an isolated cut and cover tunnel in dry sand. DesignSafe-CI [publisher] 2012. doi: 104231/D3JQ0SW10.
[100] Sadiq S, Nguyen VQ, Jung H, Park D. Effect of flexibility ratio on seismic response of cut-and-cover box tunnel. Adv Civ Eng Mater 2019;2019.
[101] Kramer SL. Geotechnical earthquake engineering. Upper Saddle River, NJ, USA: Prentice-Hall; 1996.
[102] Housner G, Jennings PC. Generation of artificial earthquakes. J Eng Mech Division 1964;90(1):113–50.
[103] Arias A. A measure of earthquake intensity. Massachusetts Institute of Technology: Seismic Design for Nuclear Power Plants; 1970.
[104] Park Y-J, Ang A-H-S, Wen YK. Seismic damage analysis of reinforced concrete buildings. J Struct Eng (N Y N Y) 1985;111(4):740–57.
[105] Housner GW. Spectrum intensities of strong-motion earthquakes; 1952.
[106] Nuttli OW. The relation of sustained maximum ground acceleration and velocity to earthquake intensity and magnitude. US Army Engineer Waterways Experiment Station 1979.
[107] Benjamin JR. A criterion for determining exceedances of the operating basis earthquake, Epri report np-5930. Palo Alto: Electric Power Research Institute; 1988.
[108] Shome N, Cornell CA, Bazzurro P, Carballo JE. Earthquakes, records, and nonlinear responses. Earthq Spectra 1998;14(3):469–500.
[109] Sarma S, Yang K. An evaluation of strong motion records and a new parameter A95. Earthq Eng Struct Dyn 1987;15(1):119–32.
[110] Rathje EM, Abrahamson NA, Bray JD. Simplified frequency content estimates of earthquake ground motions. J Geotech Geoenviron Eng 1998;124(2):150–9.
[111] Nishioka T, Unjoh S. A simplified seismic design method for underground structures based on the shear strain transmitting characteristics. Minamihara Tsukuba: Public works Research Institute; 2000.
[112] Zhang C, Zhao M, Zhong Z, Du X. Seismic intensity measures and fragility analysis for subway stations subjected to near-fault ground motions with velocity pulses. J Earthq Eng 2021:1–27.
[113] Nguyen D-D, Thusa B, Azad MS, Tran V-L, Lee T-H. Optimal earthquake intensity measures for probabilistic seismic demand models of ARP1400 reactor containment building. Nucl Eng Technol 2021;53(12):4179–88.
[114] Du W, Wang G. Fully probabilistic seismic displacement analysis of spatially distributed slopes using spatially correlated vector intensity measures. Earthq Eng Struct Dyn 2014;43(5):661–79.
[115] Zhang W, Shokrabadi M, Bozorgnia Y, Taciroglu E. A methodology for fragility analysis of buried water pipes considering coupled horizontal and vertical ground motions. Comput Geotech 2020;126(103709).