Acta Geotechnica
https://doi.org/10.1007/s11440-021-01326-2

RESEARCH PAPER

A maximum entropy method using fractional moments and deep learning for geotechnical reliability analysis

Ze Zhou Wang[1,2] · Siang Huat Goh[1,2]

Received: 23 April 2021 / Accepted: 1 August 2021

Abstract
The spatial variability of the properties of natural soils is one of the major sources of uncertainties that can influence the performance of geotechnical structures. The direct Monte-Carlo simulation (MCS) method, although robust and versatile, may incur prohibitively high computational burdens, especially for cases involving low failure probability levels. In this paper, a hybrid strategy that fuses convolutional neural networks (CNNs) and maximum entropy distribution with fractional moments (MaxEnt-FM) is proposed. MaxEnt-FM is a post-processing technique that fits a probability distribution to a set of sample data. The proposed hybrid strategy starts by training a CNN as the metamodel of the time-demanding random field finite element model. The trained CNN is then used to generate sample data, which is subsequently processed using the MaxEnt-FM technique to obtain failure probability. The use of a CNN allows MaxEnt-FM to be carried out without explicit calls to the finite element model. Therefore, the proposed hybrid strategy has the potential to provide a computationally efficient technique to calculate failure probability. For the illustrative slope stability example that has a failure probability of 3×10^{-4}, the proposed hybrid strategy yields a less than 10% error in the predicted failure probability with only hundreds of finite element simulations.

Keywords Convolutional neural networks · Fractional moments · Maximum entropy distribution · Metamodel · Small probability of failure · Spatial variability

1 Introduction

In recent years, the use of the random field finite element method (RF-FEM) has become more widespread in the geotechnical engineering community. This is the result of an increasing recognition and consensus that the spatial variability of the physical and mechanical properties of natural soils is one of the major sources of uncertainties that can influence the performance of geotechnical structures [12, 14, 41, 66]. Some reported works involving the probabilistic treatment of geotechnical structures in spatially variable soils include those by [6, 8, 11, 12, 15, 17, 43, 50–54], to name but a few. The RF-FEM approach broadly consists of two building blocks. The random field theory [49] allows the spatial variability of soil properties to be modelled explicitly while the finite element method can be used to analyse the geotechnical system with the given realization of the random field mapped to the finite element mesh. The advanced capabilities of the finite element method to analyse complex geotechnical systems have made RF-FEM more attractive than other schemes, such as those which combine the limit equilibrium method (LEM) with random fields. With the spatial variability properly represented and the geosystem behaviour properly characterized and simulated, a wealth of reliability analysis techniques, e.g. Monte-Carlo simulation (MCS), can be adopted to study the safety margin of the system against some predefined mode of failure, which is typically characterized by the term "probability of failure (P_{failure})".

✉ Ze Zhou Wang
e0054291@u.nus.edu

Siang Huat Goh
gohsianghuat@nus.edu.sg

[1] Department of Civil & Environmental Engineering, National University of Singapore, Block E1A, #07-03, No.1 Engineering Drive 2, Singapore 117576, Singapore

[2] Centre for Protective Technology, Faculty of Engineering, National University of Singapore, No. 12 Kent Ridge Road, Singapore 119221, Singapore

Published online: 15 August 2021

Monte-Carlo simulation (MCS) is a robust and useful technique that has been adopted in many reliability studies of geotechnical structures [6, 12, 14, 15, 52]. This technique starts by carrying out the geotechnical analyses for X simulated samples. The number of failure samples X_{failure} is obtained by comparing the sample results against some pre-defined performance function, following which the value of P_{failure} can be calculated as X_{failure}/X. The coefficient of variation (CoV) of P_{failure} [3] can then be calculated to assess the uncertainty associated with the calculated value of P_{failure}. The method of Monte-Carlo simulation, although versatile and conceptually simple, suffers from a lack of efficiency because a large number of samples is usually required to ensure that the calculated P_{failure} converges and a reasonably low value of CoV of P_{failure} is attained, especially for cases involving low failure probability levels. For example, in the slope stability example presented in [32], where the slope failure probability P_{failure} is 1.8×10^{-4}, 60,000 Monte-Carlo samples were simulated to achieve a CoV of 0.3. It is therefore very computationally challenging to combine MCS with the time-consuming RF-FEM analyses for problems involving low failure probability levels.

In this regard, other techniques, such as first-order reliability analysis (FORM) [25, 30], importance sampling [7, 36] and metamodel-based techniques [18, 33, 34, 50, 53, 54, 62, 64], have been proposed in the literature to improve the computational efficiency of probabilistic studies. Moreover, studies that adopt advanced techniques for problems involving low probability levels have also been reported in the literature [24, 32, 37, 51]. For example, an advanced MCS technique called "subset simulation (SS)" [2] has been applied and demonstrated to be effective by Li et al. [32] on a slope stability problem that involves spatially variable soils. The calculated P_{failure} of this example is 1.8×10^{-4}, and SS successfully yields an error of approximately 6% in the predicted P_{failure} with 2300 simulations.

Alternatively, hybrid strategies have also been proposed to study problems involving low probability levels [10, 24, 60]. This class of strategy usually combines the use of a metamodel with an advanced technique, such as the subset simulation (SS). The metamodel is first built in lieu of the original time-demanding simulation model. The analysis using the advanced probabilistic technique is then carried out using the metamodel that consumes little computational cost. For example, Jiang and Huang [24] combined the use of a polynomial chaos expansion (PCE) with simulation subset (SS). For a slope stability problem involving a P_{failure} of 1×10^{-4}, 7400 SS samples were used to calculate the failure probability and compared against the results obtained using the MCS method with 200,000 samples. The simulation time taken by the 7400 SS samples, however, is almost negligible because they were performed using a PCE constructed in lieu of the simulation model. The true computational cost associated with this hybrid strategy is the 1713 limit equilibrium simulations used to construct the PCE. In this way, Jiang and Huang [24] demonstrate that a hybrid strategy can be more computationally efficient because fewer simulations are needed for the construction of the metamodel than for the analysis using the subset simulation.

While the aforementioned studies have successfully demonstrated the effectiveness of the techniques for problems involving low failure probability levels, the need to perform approximately 2000 simulations, in order to construct the metamodel, may still incur significant computational costs from a practical point of view. To further alleviate the computational burden, a hybrid metamodel-based post-processing method is proposed in this paper. Similar to the hybrid strategy described in [10, 24, 60], the current proposed strategy consists of two building blocks. The first building block pertains to the construction of a metamodel to replace the computationally expensive finite element model. A novel metamodel using convolutional neural networks (CNNs) is adopted in the present study. In the second building block, the metamodel is used to generate a sufficiently large number of samples, which is then processed using the method of maximum entropy distribution with fractional moments (MaxEnt-FM).

This technique belongs to an important class of post-processing techniques used for reliability analysis. To implement such techniques, a set of sample data needs to be generated first. As will be explained later, the data in this study refer to a set of factor of safety (FoS) values obtained from random field finite element analyses. Using the statistical properties (e.g. statistical moments) deduced from the sample dataset, a post-processing technique is then applied to fit a probability distribution to the sample dataset. There are many post-processing techniques proposed in the literature. Other than the MaxEnt-FM technique used in the present study, the generalized lambda distribution is also commonly reported in the literature. Depending on the choice of post-processing technique, it is possible to recover different statistical properties from the fitted probability distribution. The MaxEnt-FM technique used in the present study aims to recover the probability density function based on the first-n fractional moments of the sample data. More details will be provided in Sect. 3. Using the MaxEnt-FM derived distribution, the failure probability can then be calculated. The use of fractional moments has been demonstrated in the literature to be more accurate and versatile than the use of integer moments [19, 35, 57–59].

Although MaxEnt-FM has also been shown to be effective in modelling distribution tails [55, 58, 61], it may

end up requiring thousands of model simulations in order to arrive at an accurate value of P_{failure} [61]. Such a large number of model simulations, if performed using random field finite element analyses, may incur excessively high computational costs that render the method impractical. To overcome the computational constraints, the use of surrogate models to replace the finite element analyses is a possible option that has been proposed by several authors [18, 24, 33, 34, 45]. The use of CNNs as a surrogate for random field finite element analyses has been demonstrated by Wang and Goh [54] to be a computationally efficient technique for problems involving spatially variable soil properties. In this regard, the combination of these two techniques, CNN and MaxEnt-FM, can potentially provide a more computationally efficient strategy for reliability analysis of geotechnical systems in spatially variable soils involving low failure probability levels.

The proposed hybrid strategy, referred to as CNNs-MaxEnt-FM in the subsequent parts of this study, is illustrated using a multi-layered soil slope example modified from a real slope case history. The direct MCS yields a P_{failure} of 3×10^{-4} with 100,000 simulations. In addition to the proposed CNNs-MaxEnt-FM, three other schemes are also presented. This paper starts by introducing the use of CNNs as metamodels in Sect. 2. Section 3 presents the mathematical formulation and implementation procedures of the proposed hybrid strategy. Details pertaining to the other three schemes are explained in Sect. 4. The computational performance of the proposed CNN-based hybrid strategy is then illustrated and compared against the other three schemes in Sects. 5 and 6. Discussions and conclusions are presented in Sects. 7 and 8.

2 Convolutional neural networks (CNNs)

In the field of computer science, CNNs have been commonly adopted as image and pattern recognition methods [27, 29, 46]. This method has been demonstrated to be effective in solving problems such as face recognition [28, 31], speech recognition [4], human action recognition [23] and remote sensing [56]. The use of deep learning algorithms in civil and geotechnical engineering has also been reported in the literature [5, 9, 16, 63, 65].

The idea of using CNNs as surrogates of random field finite element models was inspired by a comparison of a digital colour image with a random field. Figure 1a illustrates the conventional use of CNNs for image processing. A digital colour image is made of pixels. Each pixel has three channels that correspond to the three primary colours, i.e. red, green and blue. As a result, each pixel of the image is described by three values that correspond to the intensities of the three channels. The input layer of CNNs then contains the pixel information of the image to be processed.

With reference to Fig. 1b, in finite element analysis, the random field of a variable property of interest corresponds to the "channel" in a digital colour image. The Gauss integration points within each finite element mesh are then equivalent to the pixels of a digital colour image, while the values of the variable property of interest mapped to the finite element mesh are equivalent to the intensities of the pixels of the digital colour image. In this regard, a random field can essentially be considered as an "image" that describes the spatial distribution and intensity of the variable property of interest. For example, a soil layer characterized by a spatially varying undrained shear strength S_u involves one random field for each finite element realization. Accordingly, the input layer of a CNN is supplied with information from a one-channel "image", namely the "S_u" channel, the values of which correspond to the values of the spatially varying undrained shear strength S_u. CNNs are then trained to learn the high-level features that contain the information pertaining to the spatial variabilities, following which they will output the FEM-predicted quantities, for example, the factor of safety (FoS) of the geotechnical structure, in lieu of performing out the time-demanding finite element analyses.

In addition, it is possible to have geotechnical problems involving more than one random field. Taking a c–φ slope as an example, it is possible to have one channel containing information about the cohesion c and the other channel containing information about the friction angle φ. These two channels can then be stacked, in a similar manner as a digital colour image composed of the R, G and B channels, to form a multi-channel "image" that contains information about both c and φ. In this way, CNN can simultaneously process the two random fields [54].

Other than the input layer, a complete CNN also contains other layers, such as convolutional layer, pooling layer, activation layer and output layer. Details pertaining to these layers can be found in [9, 53]. The configuration of these layers to form a complete CNN is referred to as the architecture of the CNN. The effective use of CNNs as metamodels of the random field finite element model has been demonstrated by Wang and Goh [54].

3 The proposed hybrid strategy (CNNs-MaxEnt-FM)

3.1 Mathematical formulation

The concept of information-theoretic entropy was first proposed by Shannon [44], with the entropy calculation as follows:

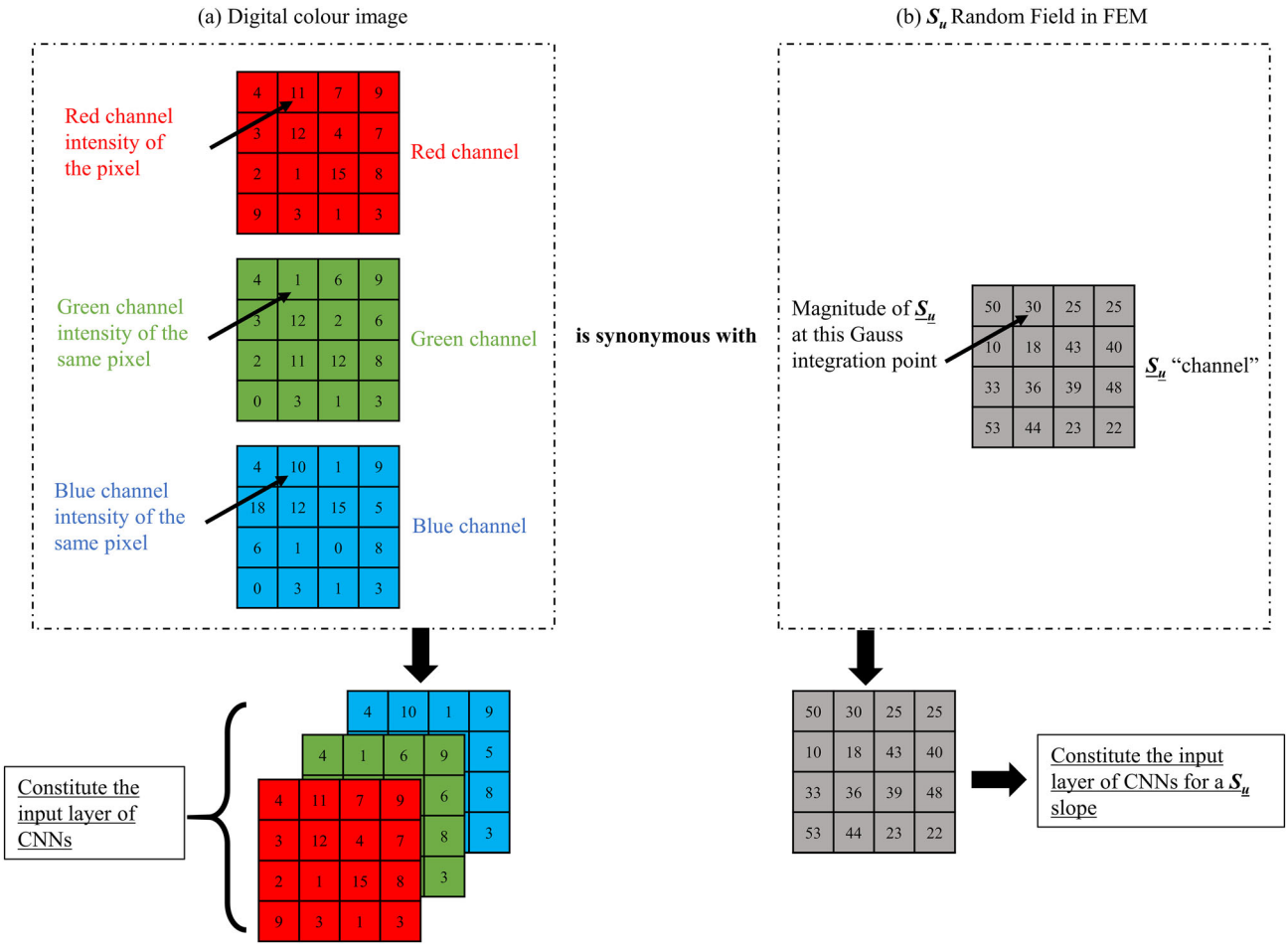

Fig. 1 Representation and processing of channel information in CNNs for **a** digital colour images and **b** random fields in finite element models

$$H[f_X] = -\int_X f_X(x)\ln f_X(x)\mathrm{d}x \tag{1}$$

where $f_X(x)$ is the PDF of a continuous random variable X. The integration domain corresponds to the feasible range of the random variable X.

Based on this concept of information-theoretic entropy, Jaynes [22] proposed the maximum entropy principle, which states that one can make the most unbiased inference by solving the following optimization problem:

$$\begin{cases} \text{Solve for}: f_X(x) & 2a \\ \text{Maximize}: H[f_X] = -\int_X f_X(x)\ln f_X(x)\mathrm{d}x & 2b \\ \text{Subject to}: E(X^{\alpha_m}) = \int_X f_X(x)x^{\alpha_m}\mathrm{d}x (\text{for } m = 1,2,3,\ldots,M) & 2c \end{cases}$$

where $E(X^{\alpha_m})$ is the mth moment and M refers to the number of moment constraints, which takes the form of an integer. The $E(X^{\alpha_m})$ in Eq. 2c is written in a more generic form. For integer moments, α_m takes the form of an integer. For fractional moments, α_m takes a value in the domain of real numbers. The integration domain in Eq. 2c also follows the feasible range of variable X.

The maximum entropy PDF used in the current study is defined as follows:

$$f_X(x) = \exp\left(-\lambda_0 - \sum_{m=1}^{M} \lambda_m (x-b)^{\alpha_m}\right), \quad x > b \tag{3}$$

where b is a location parameter, M is the number of fractional moment constraints, α_m is the mth order of the fractional moment, and λ_m is the mth Lagrange's multiplier. The normalization constant λ_0 is given as follows:

$$\lambda_0 = \ln\left[\int_b^{\infty} \exp\left[-\sum_{m=1}^{M} \lambda_m (x-b)^{\alpha_m}\right] \mathrm{d}x\right] \tag{4}$$

Without further mathematical formalities, the constrained optimization task specified in Eq. 2 can be transformed to an unconstrained optimization task with the objective function specified by Zhang et al. [61] as follows:

$$Q(\alpha, \lambda, b, M) = \lambda_0 + \sum_{m=1}^{M} \lambda_m E((X-b)^{\alpha_m}) + \frac{2M+1}{N} \tag{5}$$

where N is the size of the sample data to be post-processed and X is the sample data.

Some features in this version of MaxEnt-FM are highlighted. The parameter b in Eq. 2 is a location parameter. As specified by Zhang et al. [61], b refers to the smallest possible value of the random variable. In many engineering applications, the random variable is usually bounded. For example, the factor of safety FoS, which is the quantity studied in the subsequent illustrative example, is a positive real number. However, the exact lower bound that the problem can attain can be unknown. In this regard, Zhang et al. [61] recommend including b as a parameter determined in the optimization task as indicated in Eq. 5. This has also resulted in the integration domain in Eqs. 1, 2 and 4 being located between b and ∞.

In addition, M controls the quality of the fitted distribution. If M is too small, significant bias errors can be expected. If M is too large, the fitted distribution may become overcomplex, resulting in reduced predictive capability over unseen data. While some studies used a predetermined M based on engineering heuristic, this dilemma can be better resolved by including M in the optimization task as indicated in Eq. 5.

Furthermore, the effect of M is accommodated in the optimization task by introducing the term $\frac{2M+1}{N}$ that is derived based on the Akaike information criterion (AIC) [1]. Zhang et al. [61] have demonstrated that including M in this form leads to a better performance compared to the original form reported by Novi Inverardi and Tagliani [21]. In the present study, M, being the only integer parameter, is optimized between 1 and 6.

To simplify the optimization task, a system of linear equations is solved before the optimization process [13]. This system of linear equations is as follows:

$$\lambda = \beta^{-1}\mu \tag{6}$$

where $\mu = [E((X-b)^{\alpha_1}), E((X-b)^{\alpha_2}), \ldots, E((X-b)^{\alpha_M})]$ and the entries of β are $\beta_{i,j} = \frac{\alpha_j}{\alpha_i+1} E(X^{\alpha_i+\alpha_j})$, where i and $j = [1, 2, \ldots, M]$. The benefit of solving Eq. 6 is that λ can be correlated with a given set of X and α and therefore, the optimization task in Eq. 5 is effectively simplified to a problem with α, b and M as the parameters.

The optimization task is solved in the present study using an open-source surrogate-model-based optimization toolbox in Matlab [38]. A cubic polynomial regression is chosen as the surrogate model in the optimization process, and the dynamically dimensioned search (DDS) algorithm [47] is used to find the minimum site of the surrogate model. The parameter M, being the only integer in the optimization task, is searched between 1 and 6, which are typical values for most engineering applications [61].

3.2 Implementation procedure

In the remaining parts of this paper, the proposed CNN-based hybrid strategy is referred to as Scheme I, as shown in Table 1. The other three schemes considered in the present study are also shown in the same table. The implementation procedures of Schemes I to IV are illustrated in Fig. 2. With reference to Fig. 2a, the proposed hybrid strategy (Scheme I) can be implemented in 6 steps:

Step 1 After constructing the finite element model of the geotechnical system to be investigated and assigning the statistics of the variable properties of interest, X samples of random fields are generated. RF-FEM analyses of the X samples of random fields are then carried out to obtain the associated predicted quantities of interest, which is the factor of safety FoS in this study.

Step 2 The X pairs of random fields and predicted quantities are then divided into a training set and a validation set. In the current study, 75% of the X pairs is used as the training set while the remaining 25% of the X pairs is used as the validation set. Such a configuration is commonly adopted to avoid overfitting.

Step 3 After designing the architecture of the CNNs, the training and validation data are fed to the CNNs for training. The accuracy of the training and validation exercise is assessed by evaluating the root-mean-square errors of the CNN-predicted FoS values against the corresponding values from RF-FEM analyses.

Step 4 After sufficient training iterations, the trained CNNs are used in lieu of the time-demanding finite element analyses for making predictions of a sufficiently large size of Monte-Carlo samples.

Step 5 The CNN-predicted FoS values associated with the Monte-Carlo samples are then post-processed by

Table 1 Descriptions of the four schemes presented in the study

Scheme	Description
I	CNNs + MaxEnt-FM (CNNs-MaxEnt-FM)
II	Maximum entropy distribution with fractional moments (MaxEnt-FM)
III	Convolutional neural networks (CNNs) + non-adaptive training + Monte-Carlo simulation (MCS)
IV	CNNs + adaptive training + MCS

using the maximum entropy distribution with fractional moments (MaxEnt-FM) technique to fit a probability distribution to the FoS data set.

Step 6 Based on the fitted distribution (Eq. 3) obtained in Step 5, the value of P_{failure} can be calculated in MATLAB using the *integral* function. In the current study, the stability of the multi-layer slope is determined using the FoS values. Therefore, the P_{failure} is defined as the probability of obtaining FoS values that fall below 1.0.

4 Other schemes and implementation procedures

4.1 Scheme II: MaxEnt-FM

With reference to Fig. 2b, Scheme II is similar to Scheme I in that MaxEnt-FM is also used to fit a distribution to the sample FoS values. However, CNNs are not used in Scheme II and therefore, the sample FoS values to be processed by MaxEnt-FM in Scheme II are directly generated and calculated via finite element analysis. In this regard, Schemes I and II differ in the way the sample FoS values are generated. Scheme II can then provide benchmark results (involving the use of MaxEnt-FM) for comparing against the results obtained using Scheme I in which metamodels are used. The comparison provides a means to assess the accuracy and the effectiveness of CNNs as the metamodel of choice to generate sample FoS values for post-processing using the MaxEnt-FM technique.

4.2 Scheme III: CNNs + MCS

This scheme describes the conventional use of a metamodel for a reliability analysis. With reference to Fig. 2c,

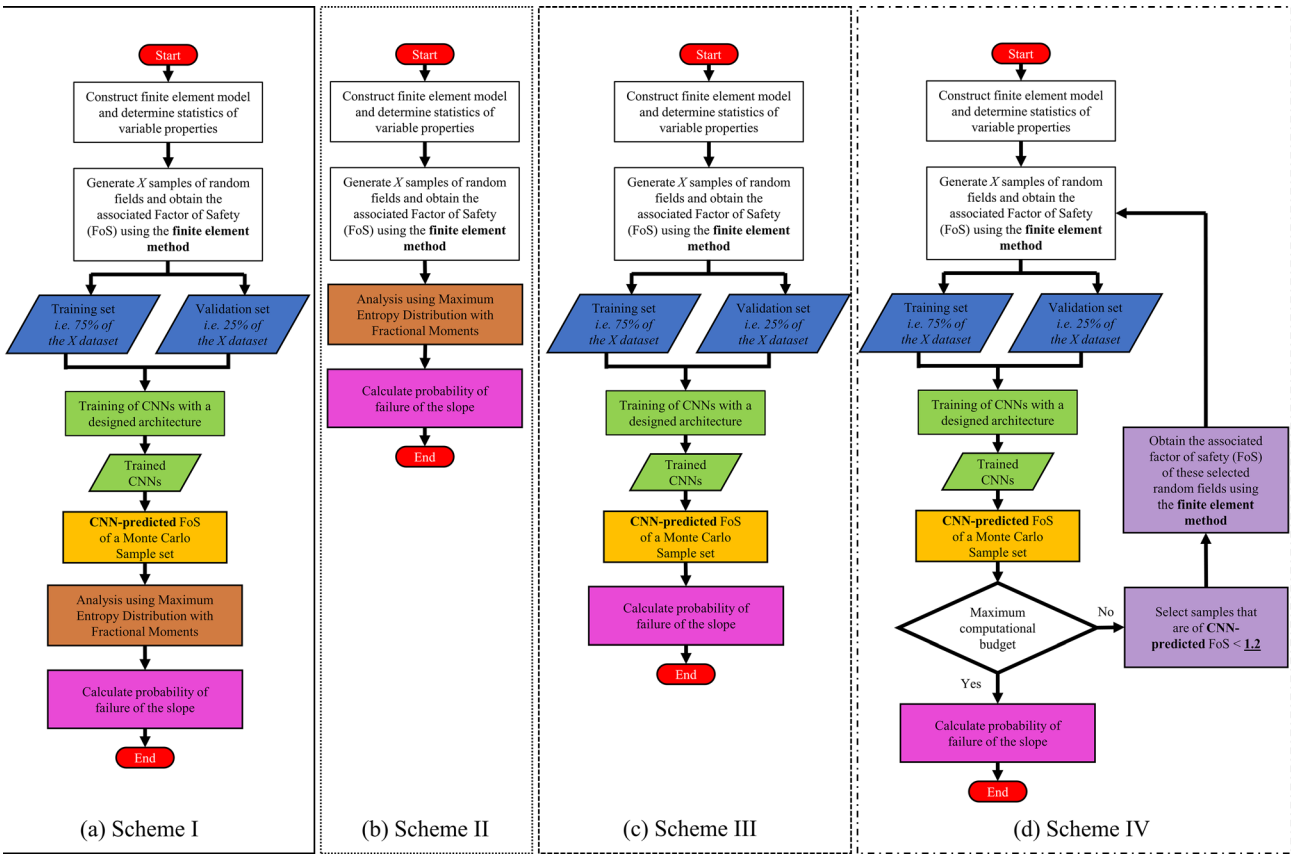

Fig. 2 Implementation procedures for Schemes I to IV

the steps that involve training of a CNN are identical to those of Scheme I. However, instead of using MaxEnt-FM to fit a distribution to the CNN-predicted FoS values, as has been proposed in Scheme I, Scheme III directly calculates the failure probability following the Monte-Carlo simulation technique with 100,000 samples, that is, $P_{\text{failure}} = \frac{\text{Number of realizations with FoS} < 1}{100{,}000}$. This scheme is compared against Scheme I to illustrate the improved prediction of the P_{failure} values that may be achieved using the MaxEnt-FM technique.

4.3 Scheme IV: CNNs + adaptive training + MCS

This scheme provides an improved version of Scheme III. In the literature, the adaptive infill strategy has been commonly adopted to construct a metamodel [36, 62]. This strategy often starts from a coarse sample set and then selects additional samples based on a pre-defined criterion using some selection algorithm. In this way, the metamodel is refined after including the selected additional samples. For example, Liu et al. [36] adopts an importance sampling strategy to select samples that belong to the failure region to refine the predictive capability of the metamodel over the failure region. A similar strategy is adopted in Scheme IV in the present study, and the implementation procedure is illustrated in Fig. 2d. The comparison of Schemes IV and I, which will be discussed in the later parts of this paper, aims to highlight the advantage associated with the hybrid strategy of Scheme I (CNN-MaxEnt-FM) over the adaptive training strategy of Scheme IV commonly used in the literature, although the latter strategy can also yield reasonable results.

The initialization steps of Scheme IV are identical to those of Schemes I and III. In the next step, in contrast to Scheme III, which directly calculates the failure probability following the Monte-Carlo simulation technique, Scheme IV first identifies potential 'failure' samples corresponding to CNN-predicted FoS values lower than 1.2. The threshold value of 1.2 is selected as it is close to the slope failure condition of FoS = 1, and there may exist CNN-predicted FoS values that are larger than 1.0 which actually correspond to failure cases with FoS values below 1.0 had the analyses been performed using the finite element method. Such errors may occur given that CNN is used as a metamodel for the more rigorous finite element analysis, and hence, some prediction discrepancies are inevitable. The relaxation of the FoS threshold to 1.2 reduces the possibility that failure samples are falsely rejected during the adaptive sampling process involving the CNN-predicted FoS values.

These identified "failure" samples with FoS < 1.2 are then re-run using the finite element model to obtain the "true" FoS values. Subsequently, these additional samples are combined with the initial X samples to form a new dataset to re-train the CNN. Such an adaptive sampling and training loop is then repeated till the prescribed computational budget is used up. This strategy helps to identify samples in the failure domain, and therefore, the training dataset is expected to contain more failure samples after every loop. The training using the new dataset can then improve the predictive capability of the trained CNN over the failure domain. After the CNN is sufficiently trained, it can be used to predict the FoS values of a Monte-Carlo sample set, based on which the value of P_{failure} can be calculated.

5 Illustrative example

The four schemes listed in Table 1 are implemented on a slope stability problem, whose geometry is based on the Congress Street cut in Chicago [20, 40]. With reference to Fig. 3, this slope comprises four soil layers. The spatial variabilities of the undrained shear strength S_u(kPa) in the three clay layers are considered, while the properties of the top sand layer are taken to be deterministic. The statistics of the material parameters are listed in Table 2. It should be noted that these statistics are modified from the real values of the in-situ soils so that a low value of P_{failure}, i.e. 3×10^{-4}, can be obtained for the purpose of the present study.

The commercial finite element software Optum G2 [26] is used to perform the strength-reduction calculations to obtain the factor of safety against slope failure. This software has a built-in engine to generate random fields using the Karhunen–Loève expansion and a single exponential autocorrelation function. The random fields are generated with 1000 K-L terms to ensure sufficient accuracy. Details pertaining to the mathematical formulation of K-L expansion can be found in [42].

To validate the results of the four schemes listed in Table 1, a direct MCS run with 100,000 samples is performed to calculate the benchmark P_{failure}. With a threshold FoS = 1.0, the value of P_{failure} is 3×10^{-4} and the associated value of CoV is approximately 0.183. Each simulation takes approximately 1 min on a desktop computer with 8 GB RAM and one Intel Core i5-6400 CPU @ 2.70 GHz.

Figure 4a to b shows two selected S_u random fields, while Fig. 4c to d shows the failure mechanisms corresponding to these two stochastic realizations. In general, a deep-seated failure mechanism with an almost circular failure surface is observed. This is in line with the typical failure pattern observed in an undrained slope.

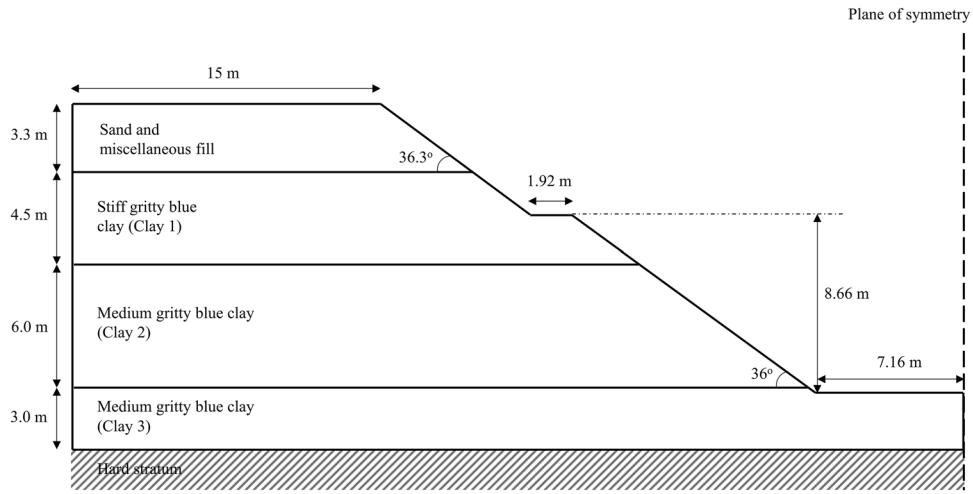

Fig. 3 Schematic illustration of the slope

Figure 5 shows the architecture of the CNNs used in the current study. The first layer is the input layer with a random field "image" of 30 × 100. The outputs produced by a convolutional layer that consists of 20 filters of size 25 × 90, a padding of 0 and a stride of 1, are fed into the ReLU layer [39] before being further processed by a max pooling layer with a size of 2, a padding of 0 and a stride of 1. Finally, a regression is constructed between the FoS values and the outputs of a fully connected layer. A dropout layer, which randomly disconnects the neurons in the layer with a 50% chance, is used to prevent the network from overfitting.

The architecture in the current study is determined by a trial-and-error approach. In this regard, the CNNs reported in the literature [9, 27] could offer insights into the architecture design. The stochastic gradient descent algorithm is adopted to tune the filter weights and the biases in the training of the CNNs. The learning rate is 0.001 and the batch size is 256, which is a typical value. Validation is carried out after every training epoch. The training process is terminated when the validation accuracy does not improve for 500 consecutive epochs or when 15,000 epochs are reached, whichever is earlier. These steps are taken to reduce overfitting. The accuracy of the training and validation is assessed by calculating the root-mean-square errors (RMSE) of the predicted FoS values. Table 3 summarizes the configurable hyperparameters used in the training of the CNNs in Schemes I, III and IV.

6 Results of reliability analysis

One of the important aspects of the results in the present study pertains to the computational cost of each scheme. In the present study, the computational cost involves two parts. The first part pertains to the number of finite element (FE) simulations needed to train a CNN. In the remaining parts of this paper, these numbers are denoted as "size of training samples". The second part of the computational cost pertains to the number of FE simulations needed to generate sample FoS values for the post-processing using MaxEnt-FM. In the remaining parts of this paper, these numbers are denoted as "size of post-processed samples".

The computational cost of Scheme I is equal to the size of training samples for the CNN, as the sample FoS values to be post-processed are generated using the trained CNN.

Table 2 Statistics of the material parameters considered in the illustrative example

Layers	Unit weight (kN/m^3)	Friction angle ϕ (°)	Undrained shear strength S_u (kPa)			Scale of fluctuation (SoF)	
			Mean	CoV	Distribution	SoF_h (m)	SoF_v (m)
Sand	21	30	–	–	–	–	–
Clay 1	19.5	–	85	0.25	Lognormal	35	10
Clay 2	19.5	–	70	0.19	Lognormal	35	10
Clay 3	20	–	80	0.20	Lognormal	35	10

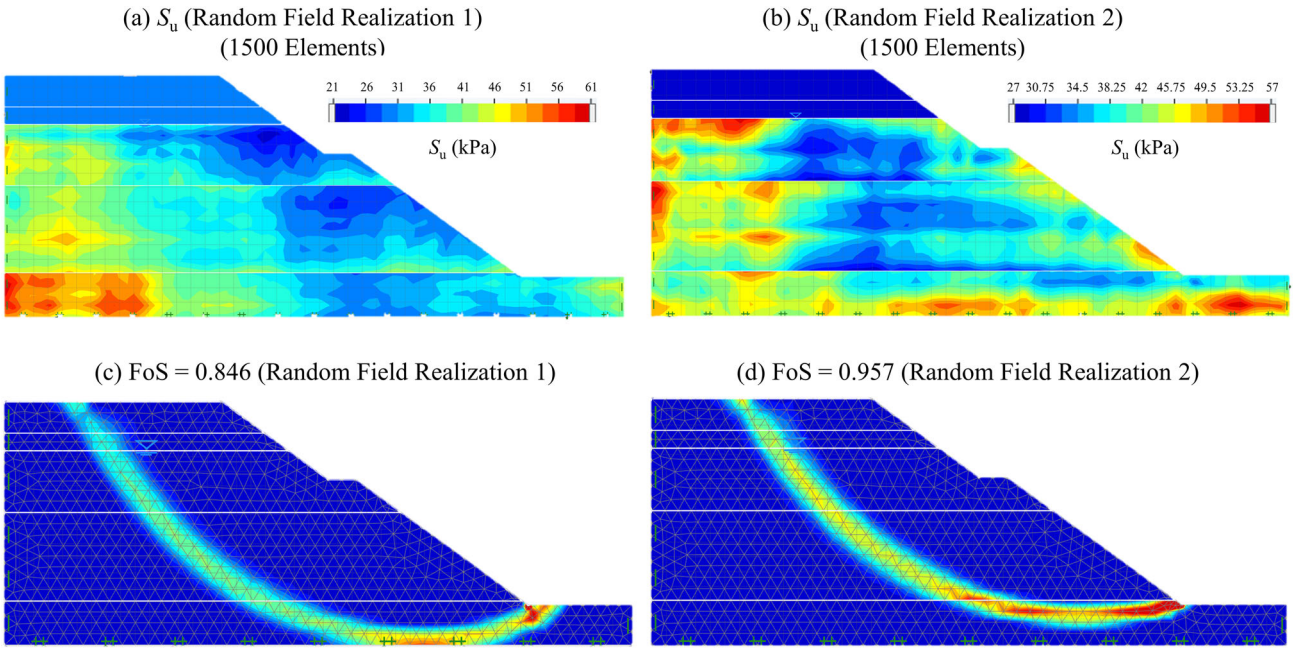

Fig. 4 Examples of random field realizations and the associated slope failure mechanisms deduced from the computed shear strains. Subplots **a** and **b** show two selected Su random field realizations. Subplots **c** to **d** show the mechanisms of RF-FEM simulations

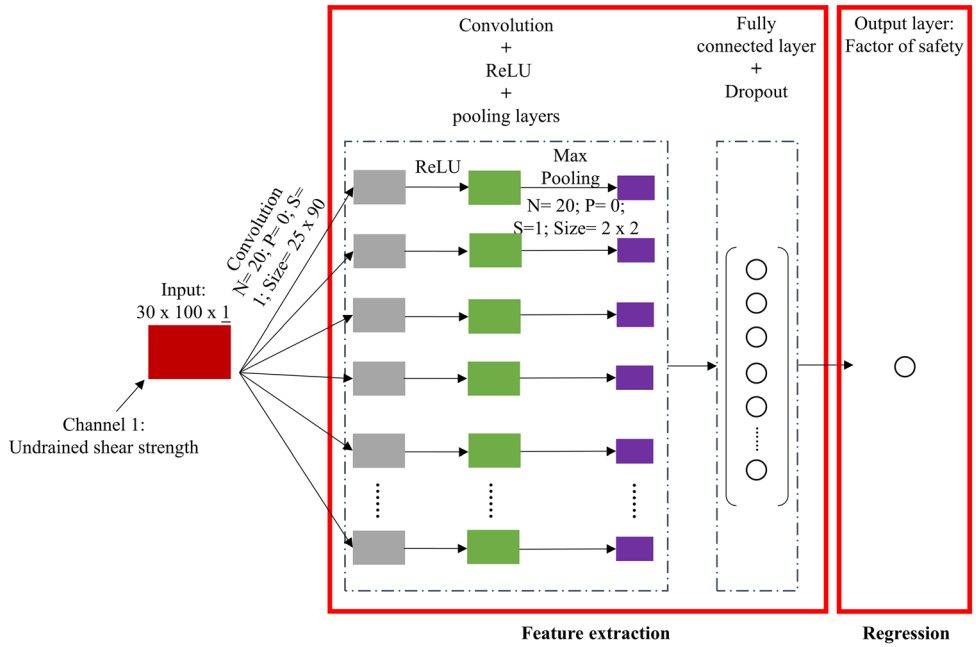

Fig. 5 Architecture of the CNNs used in Schemes I, III and IV. (N: number of filters; P: padding; S: stride)

In the present study, the performance of Scheme I was examined using sample sizes of 400, 500, 600 and 700 for training the CNN and 2000, 3000, 4000 and 5000 for post-processing using MaxEnt-FM. For example, the CNNs trained with 400 samples are used to generate four sets of CNN-predicted FoS values, with sizes ranging from 2000 to 5000, for the post-processing exercise using MaxEnt-FM. The results of the reliability analyses will then be separately tabulated for the four cases. As a result, there are in total 16 sets of results associated with Scheme I.

The computational cost of Scheme II is equal to the size of the post-processed sample FoS values, as all the sample values used in the MaxEnt-FM exercise are directly generated using the finite element analyses. The performance

Table 3 The configurable hyperparameters used in the training of the CNNs in Schemes I, III and IV

Learning rate	Batch size	Training epoch	Validation frequency
0.001	256	15,000 (early stop criterion applied)	1

of Scheme II is examined by performing the MaxEnt-FM fitting exercise using the results obtained from 2000, 3000, 4000 and 5000 finite element analyses. The results of reliability analyses are tabulated for the four cases and compared against the 16 sets of results of Scheme I.

The computational costs of both Scheme III and IV are equal to the size of samples for training the CNN. In Scheme III, the number of finite element analyses performed is 400, 500, 600, 700 and 5000 calls, while the numbers are 400, 500, 600 and 700 for Scheme IV.

Furthermore, as highlighted in many studies [32, 34, 61], it is important to examine the consistency of the proposed methods. This consistency is usually measured by the variations in the calculated results in repeated analyses with different datasets. Therefore, for each of the cases mentioned in the preceding paragraphs, the current study repeats the analysis 100 times, each involving a CNN trained using a different training dataset. For example, the analysis in Scheme IV with 400 training samples is repeated 100 times using CNNs trained with 100 different sets of 400 training samples. The trained CNNs are then used to predict the same set of 100,000 MCS samples to provide the FoS datasets for the reliability analyses. As a result, 100 values of P_{failure} can be obtained. The CoV of these 100 values of P_{failure} can then be calculated and used as the performance metric to assess the consistency of the method.

6.1 Results of Scheme I

With reference to Fig. 2, after the CNN is trained, it can be used to generate the FoS predictions for a Monte-Carlo sample set. Based on the CNN-predicted FoS values, the analysis using MaxEnt-FM can be carried out to calculate the value of P_{failure}. Figure 6a and b plots, respectively, the CNN-predicted FoS against the FEM-predicted FoS for the cases with training sample sizes of 400 and 700. Each subplot contains 100,000 scatters that correspond to a 100,000 Monte-Carlo sample set. It is worth highlighting that the 100,000 samples and the training samples belong to two different datasets. This means that the 100,000 samples do not contain the samples used to train the CNN. Figure 6c shows a typical training progress of a CNN. For both the training and validation results, the root-mean-square errors of the FoS values decrease rapidly to reach relatively constant values after about 6000 epochs. The minor fluctuations observed after 6000 epochs are due to the presence of a dropout layer that randomly disconnects the connections between neurons. The training stops at 9810 epochs when the validation root-mean-square error does not improve over 500 epochs.

The root-mean-square error (RMSE) and the coefficient of determination (r^2) between the CNN-predicted FoS and the FEM-predicted FoS are separately calculated for the two subplots. Globally, the predictions made by the trained CNN agree reasonably well with the FEM predictions. The RMSE values are less than 0.05, and the r^2 values are all greater than 0.95, suggesting that there is a strong correlation between the CNN-predicted FoS and the FEM-predicted FoS. These observations indicate that the trained CNN can be considered as a good metamodel to make predictions in lieu of the finite element analyses.

The next step involves the use of MaxEnt-FM to post-process the sample FoS values generated using the trained CNN. Section 3.1 highlights that the analysis using MaxEnt-FM is essentially an optimization task, which can be performed using an optimization toolbox in Matlab [38]. A cubic polynomial regression is chosen as the surrogate model, and the dynamically dimensioned search (DDS) algorithm [47] is used to find the minimum site of the surrogate model. In this regard, before presenting the results of Scheme I, it is necessary to first assess the robustness of the optimization algorithm. This is achieved by repeating the optimization for the same set of post-processed sample FoS values with different initial starting points.

A case with 5000 post-processed sample FoS values and a pre-determined M (Eq. 5) of 2 is set up. It is necessary to highlight that M is pre-determined to simplify this auxiliary analysis for the robustness check. In the actual analyses carried out for Schemes I and II, M is considered as a parameter in the optimization task. Five different sets of initial starting points are generated using the method of Latin hypercube sampling, and five optimization runs are carried out for the same set of 5000 sample FoS values. The results of the five optimization runs are presented in Fig. 7.

Each optimization run consists of 3000 iterations. The results indicate that all five optimization runs attain convergence within the prescribed number of iterations. A

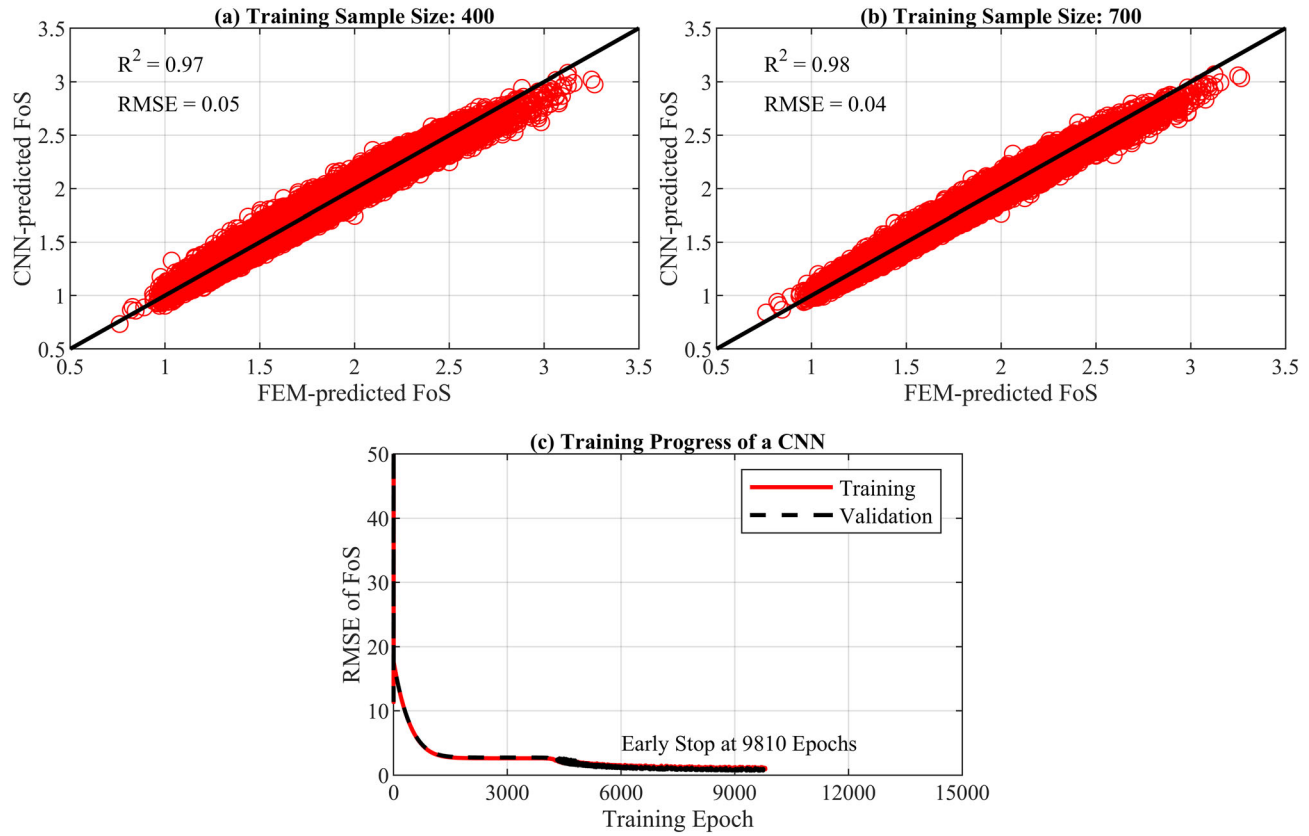

Fig. 6 Training progress of a CNN and comparisons of FEM-predicted FoS and CNN-predicted FoS. **a** With 400 training samples, **b** with 700 training samples, and **c** a typical training progress

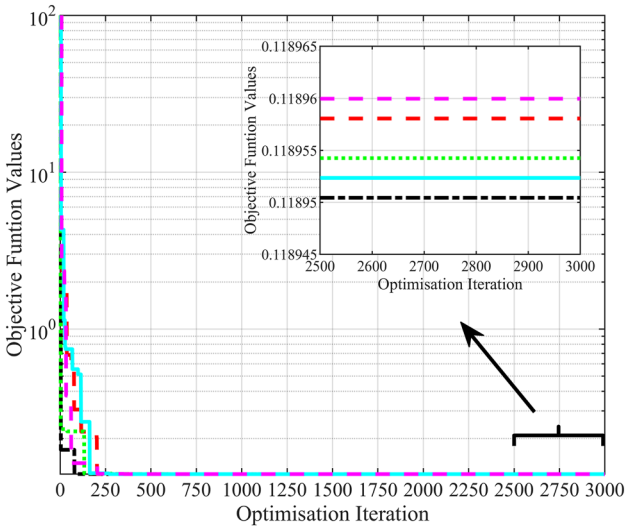

Fig. 7 Robustness study of the optimization algorithm used with MaxEnt-FM in Schemes I and II

magnified view of the results after 2500 iterations is also shown in Fig. 7, which indicates that no further improvements in the objective function values are achieved beyond 2500 iterations. The results also show that although all five runs start from different initial points, they converge to largely similar objective function values at the end of the 3000 iterations. The maximum and minimum objective function values among the five independent runs are 0.11896 and 0.11895, respectively. This observation supports the conclusion that the optimization algorithm is robust and is insensitive to initial starting points.

As stated previously, the performance of Scheme I was examined using sample sizes of 400, 500, 600 and 700 for training the CNN and 2000, 3000, 4000 and 5000 for post-processing using the MaxEnt-FM approach. This gives rise to a total of 16 sets of results associated with Scheme I. Figure 8 presents the results of P_{failure} of Scheme I. The four subplots show the results associated with the use of different post-processing sample sizes. In each subplot, the results of Scheme I are represented by the second to fifth bars that correspond to the four sample sizes used in training the CNN. For example, the four bars corresponding to Scheme I in Fig. 8a show, respectively, the results of using CNNs trained with 400 to 700 samples to predict 2000 post-processed sample FoS values for the post-processing using MaxEnt-FM.

In each subplot, the y-axis on the left indicates the value of P_{failure}, while the y-axis on the right indicates the percentage error in P_{failure} evaluated with respect to the value

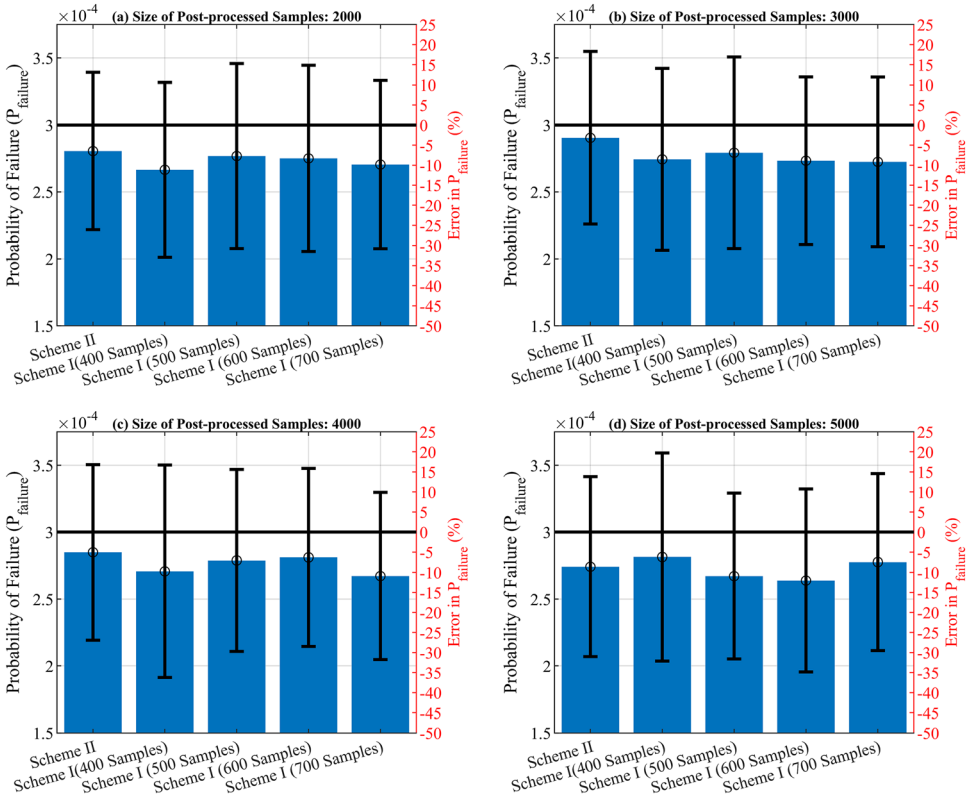

Fig. 8 Results of failure probabilities of Schemes I and II for four different post-processed sample sizes, calculated based on the 100 independent runs

of P_{failure} (= 3 × 10^{-4}) obtained using direct MCS with 100,000 samples. The result of the direct MCS is represented by a horizontal line that corresponds to a P_{failure} of 3 × 10^{-4} and an error of 0%. The height of a bar indicates the average value of P_{failure} calculated from the 100 independent runs, and the 1-standard-deviation bounds of the P_{failure} over the 100 runs are also indicated on each bar. The results of the reliability analyses of Scheme I are also summarized in Table 4.

Overall, the performance of Scheme I is promising in all 16 cases considered. The errors in P_{failure} are within 10% in most cases, with some exceptions showing errors slightly higher than 10%. The CoV values of Scheme I are approximately 25% for all 16 cases, indicating that the proposed hybrid strategy (CNNs-MaxEnt-FM) produces reasonably consistent results. Moreover, no noticeable differences are observed among the 16 sets of results. This observation implies that, for the slope considered in the study, (i) a CNN trained with 400 samples should be sufficiently accurate for the purpose of generating sample FoS values which will be fitted to a distribution using the MaxEnt-FM technique, and (ii) a FoS sample size of between 2000 and 5000, when used in conjunction with the MaxEnt-FM technique, is sufficient to yield accurate results of P_{failure}. Therefore, it can be concluded that the proposed hybrid strategy can successfully predict a P_{failure} of 3 × 10^{-4} using a CNN trained with the FoS results from 400 finite element simulations, which would otherwise require 100,000 finite element simulations by the direct MCS technique.

Figure 9 plots the results of the fitted PDFs and CDFs obtained from one of the 100 independent runs. The optimized parameter values that characterise the distributions shown in this figure are summarized in Table 5. Equations 3 and 4 can then be used to generate the distributions. Overall, the fitted distributions in Fig. 9a, c, e and g show good agreement with the histogram that contains the 100,000 Monte-Carlo samples. The left end of the CDFs close to FoS values of 1.0 are also highlighted and magnified in Fig. 9b, d, f and h, which indicate that the left tails of the distributions are also well captured using the CNN-MaxEnt-FM approach. This further demonstrates that the MaxEnt-FM technique is effective in modelling the distribution tail based on only a small set of sample FoS values.

6.2 Results of Scheme II

As explained in Sect. 4.1, CNNs are not used in this scheme. All the sample FoS values to be processed in this

scheme are generated and calculated solely using the finite element model. This scheme provides benchmark results for comparison against the results obtained using Scheme I. Such a comparison provides a means to assess the accuracy and the effectiveness of CNNs as metamodels to generate sample FoS values for post-processing using MaxEnt-FM.

The first bar in each subplot of Fig. 8 shows the results of Scheme II. The results are also presented in Table 6. Overall, Scheme II performs well for all four cases considered. The errors in the predicted P_{failure} are consistently smaller than 10%, which indicates that MaxEnt-FM is effective in modelling the distribution tail and is accurate in predicting the level of P_{failure} in this illustrative example. The CoVs are between 0.21 and 0.25, which suggests that the method produces fairly consistent results for different post-processed FoS values.

Figure 10 shows, in a similar manner as Fig. 9, results of the fitted PDFs and CDFs for the cases with 4000 and 5000 post-processed sample FoS values, obtained from one of the 100 independent runs. The optimized parameter values that characterise the distributions shown in this figure are summarized in Table 7. The cases with 4000 and 5000 post-processed sample FoS values are best characterized with M values of 2 and 3, respectively. It is worth highlighting that the values of M are entirely determined by the optimization algorithm. Globally, the fitted distributions agree well with the histogram that consists of the 100,000 Monte-Carlo samples in both selected cases.

The comparison of the performance of Schemes I and II shows that the percentage errors in the P_{failure} values are only slightly higher in Scheme I. As Scheme I involves the use of a surrogate for predicting the 2000 to 5000 FoS values, the accuracy of the P_{failure} values obtained using Scheme I is not expected to be as good as the P_{failure} values obtained using Scheme II (based on the same 2000 to 5000 random fields analysed using the finite element method). Nevertheless, the discrepancies may be deemed quite marginal, with the percentage errors in the P_{failure} values of Scheme I falling within 10% in many of these cases. The CoV values of Scheme I are also only slightly larger than the values of Scheme II. These observations indicate that the use of CNN as an approximate surrogate of the finite element analyses does not lead to any significant reduction in the accuracy and consistency of the results obtained using the proposed hybrid strategy of Scheme I.

While the overall performance of Schemes I and II is largely comparable, the computational costs associated with the two schemes are significantly different. For example, the results of Scheme II shown in Fig. 8a are obtained using 2000 finite element simulations; in contrast, the results of Scheme I shown in the same subplot require only between 400 and 700 finite element analyses for training the CNN. This demonstrates that a comparable

Table 4 Results of failure probabilities of Scheme I

Size of training sample	Size of post-processed sample	p_f	$\varepsilon_r(\%)$	CoV
400	2000	0.000267	− 11.16	0.24
	3000	0.000274	− 8.55	0.25
	4000	0.000271	− 9.75	0.29
	5000	0.000281	− 6.20	0.28
500	2000	0.000277	− 7.73	0.25
	3000	0.000279	− 6.91	0.26
	4000	0.000279	− 7.06	0.24
	5000	0.000267	− 10.94	0.23
600	2000	0.000275	− 8.31	0.25
	3000	0.000273	− 8.89	0.23
	4000	0.000281	− 6.30	0.24
	5000	0.000264	− 12.05	0.26
700	2000	0.000270	− 9.85	0.23
	3000	0.000272	− 9.18	0.23
	4000	0.000267	− 10.93	0.23
	5000	0.000278	− 7.47	0.24

p_f: probability of failure (averaged value of 100 sets of analyses)

ε_r: error (%) in CNNs-MaxEnt-FM-predicted p_f

CoV: coefficient of variation of predicted p_f over 100 sets of analyses

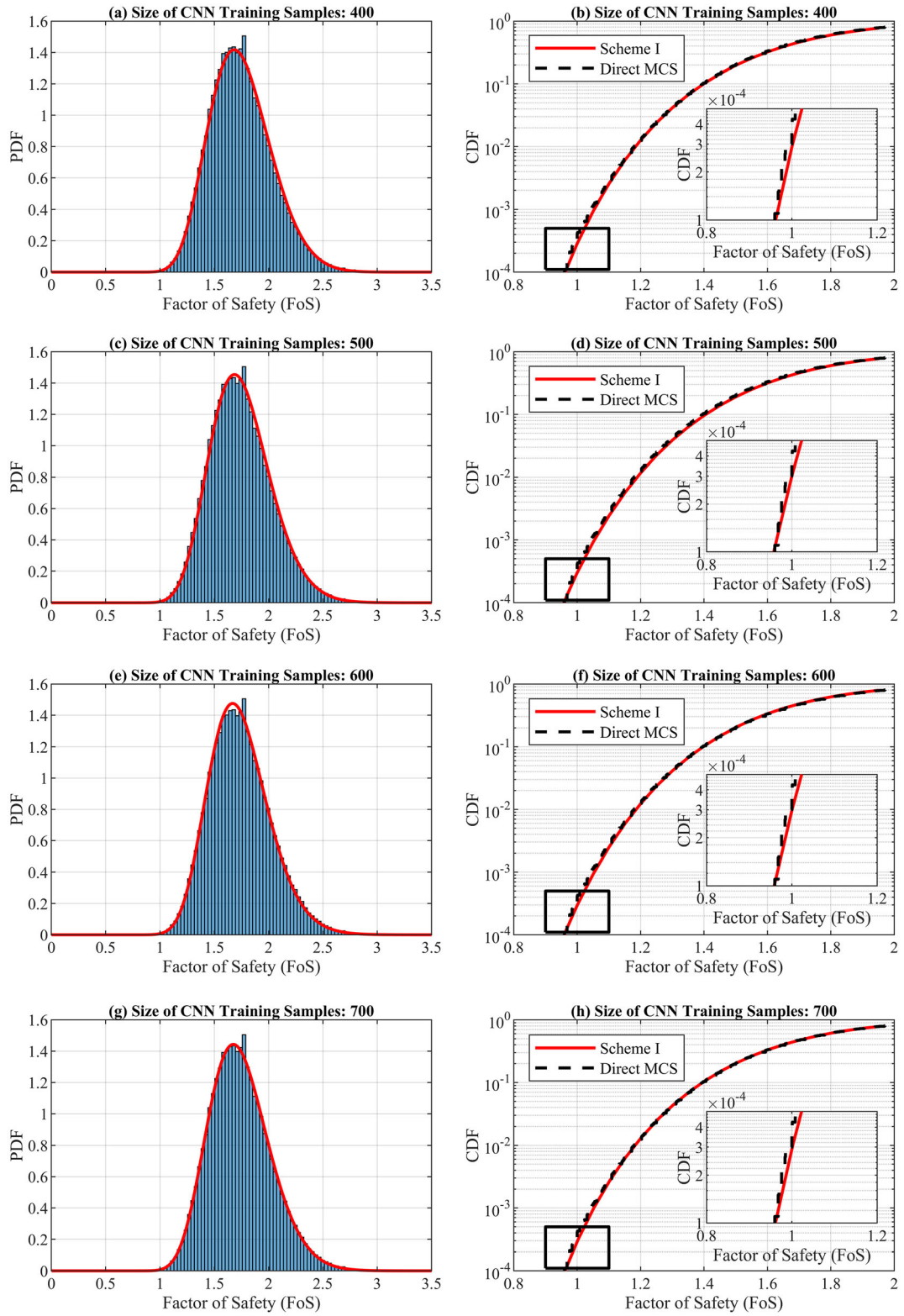

Fig. 9 Results of fitted PDFs and CDFs of Scheme I obtained from one of the 100 independent runs. (All four sets are based on the analysis with 5000 post-processed sample FoS values, but these four cases differ in the sizes of training samples.)

performance using Scheme I can be achieved with a significantly lower computational budget, thus making Scheme I an attractive alternative option for performing such reliability analyses.

6.3 Results of Schemes III and IV

Figure 11a and b shows, in a similar manner as Fig. 8, the P_{failure} results obtained using Schemes III and IV, respectively. The data used to produce Fig. 11 are summarized in Table 8. The average values of P_{failure} yielded by Scheme III, as can be seen in Fig. 11, are consistently larger than the benchmark P_{failure} obtained using the direct MCS method, with percentage errors higher than 30% for all five training sample sizes. Moreover, the variations in P_{failure}, as reflected in the bounds for $\pm$ one standard deviation, are also much more significant compared to Schemes I and II. Table 8 highlights that the CoVs are consistently larger than 0.35 for the cases with 400 to 700 samples used in training the CNNs and that it takes about 5000 training samples to achieve a CoV of about 0.22. The large errors in P_{failure} across all five cases can be explained by the fact that the training datasets contain very few failure samples. For example, a P_{failure} of 3×10^{-4} suggests that a Monte-Carlo sample set of 5000 contains, on average, fewer than two failure samples. Hence, it may be expected that the other four cases, with much smaller training sample sizes, may not even contain any failure samples. As a result, the CNNs are inadequately trained for the failure region and therefore, the accuracy of the predictions made with respect to this region depends almost entirely on the extrapolation capability of the model, which could then lead to the compromised accuracy of Scheme III. Such a limitation associated with the use of metamodels has been a common criticism, especially for cases involving low failure probability levels [36, 45].

In contrast, the average values of P_{failure} obtained using Scheme IV in Fig. 11b suggest that this scheme yields promising results. Scheme IV starts from the same initial training sample set of 400 samples as in Scheme III. The trained CNNs, following the procedure described in Sect. 4.3, are then used to identify 100 samples within or near to the failure region, for which the finite element analyses are then performed. After these 100 failure samples are used to re-train the CNNs, a reduced error in P_{failure} is observed. This is represented by the second bar in Fig. 11b. However, the 30% error in P_{failure} indicates that the 100 additional failure samples are still not sufficient for training a CNN whose predictions can lead to a good modelling of the distribution tail. Hence, more failure samples are needed. As more adaptive training iterations are carried out, further reductions in both the errors in P_{failure} and the CoV are achieved. After 600 samples, which contains 200 failure samples, are used for training, the error in P_{failure} reduces to 24.5%. Finally, when 700 samples are considered, of which 300 samples are failure cases, the error in P_{failure} is successfully brought down to be within 10%. At the same time, the CoV of the P_{failure} values (based on 100 independent runs) calculated using the final 700 samples is 0.25, which contrasts with the CoV of 0.38 predicted by the CNN trained with the initial 400 samples. This reduction attests to an improvement in the consistency of the results with the adaptive training strategy.

Table 6 Results of failure probabilities of Scheme II

Size of post-processed sample	p_f	ε_r(%)	CoV
2000	0.000281	− 6.49	0.21
3000	0.000290	− 3.19	0.22
4000	0.000285	− 5.06	0.23
5000	0.000274	− 8.60	0.25

p_f: probability of failure (averaged value of 100 sets of analyses)

ε_r: error (%) in MaxEnt-FM-predicted p_f

CoV: coefficient of variation of predicted p_f over 100 sets of analyses

7 Discussions

7.1 Auxiliary results

Figure 12 compiles and compares the P_{failure} results of Schemes I, III and IV based on the computational costs.

Table 5 Parameters of the MaxEnt-FM distributions shown in Fig. 9

Training sample size	Post-processed sample size	Objective function value	λ_0	λ_1	α_1	λ_2	α_2	λ_3	α_3	b
400	5000	0.141	− 10.480	12.978	− 2.005	0.003	6.195	2.402	1.588	0.001
500	5000	0.131	− 26.312	4.747	− 2.160	21.526	− 0.830	4.960	1.427	0.001
600	5000	0.121	− 27.918	24.963	− 1.223	7.866	1.151	–	–	0.001
700	5000	0.131	− 10.193	0.184	− 6.324	1.683	2.003	13.163	− 1.842	0.003

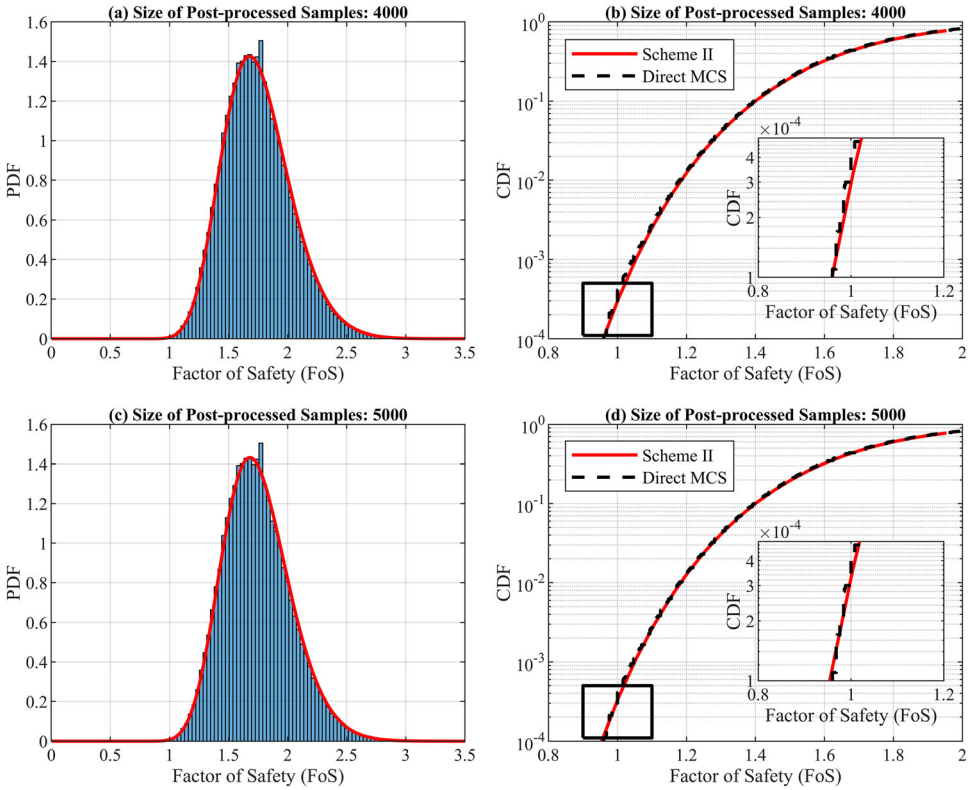

Fig. 10 Results of fitted PDFs and CDFs of Scheme II obtained from one of the 100 independent runs

Table 7 Parameters of the MaxEnt-FM distributions shown in Fig. 10

Post-processed sample size	Objective function value	λ_0	λ_1	α_1	λ_2	α_2	λ_3	α_3	b
4000	0.151	− 14.487	3.261	1.556	16.053	− 1.649	–	–	0.004
5000	0.145	− 16.898	0.164	− 5.232	17.894	− 1.425	3.592	1.539	0.004

The results contained within the same subplot incur the same computational costs. In the case where 400 finite element simulations were carried out, Scheme I performs better than both Schemes III and IV. While Scheme III shows limited improvement in term of accuracy as additional training samples are considered, Scheme IV can achieve, in the cases with 600 and 700 training samples, comparable performance as Scheme I. This observation suggests that Scheme IV can be used as an alternative strategy for reliability analysis in spatially variable soils at low probability levels, with only a marginal increase in the computational cost incurred compared to the best-performing CNNs-MaxEnt-FM strategy (Scheme I).

Table 9 summarizes the computational times of all schemes and analyses in hours. The brute-force MCS approach requires 1680 h (∼ 70 days) to complete the 100,000 finite element simulations. Scheme I, which adopts the combined CNN-MaxEnt-FM approach, can complete the reliability analysis with acceptable accuracy in 13 h (∼ 2 working days). Although Scheme II offers marginally better accuracy, the generation of the 2000 to 5000 FoS samples using the finite element analyses takes about 81 h. Scheme IV, being an alternative option to Scheme I in terms of accuracy performance, completes the analysis in 17 h. The extra 4 h is consumed by the adaptive sampling phase. In future work, it is recommended to investigate the use of an advanced adaptive sampling strategy, e.g. importance sampling, to further improve the efficiency of Scheme IV.

7.2 On the use of CNNs

The use of CNNs as metamodels has several advantages. First, CNNs have been shown to be highly efficient for problems involving spatially variable soils. For example, the extreme gradient boosting method reported by Wang et al. [50] requires the number of samples to be no fewer than the number of gauss integration points in the finite

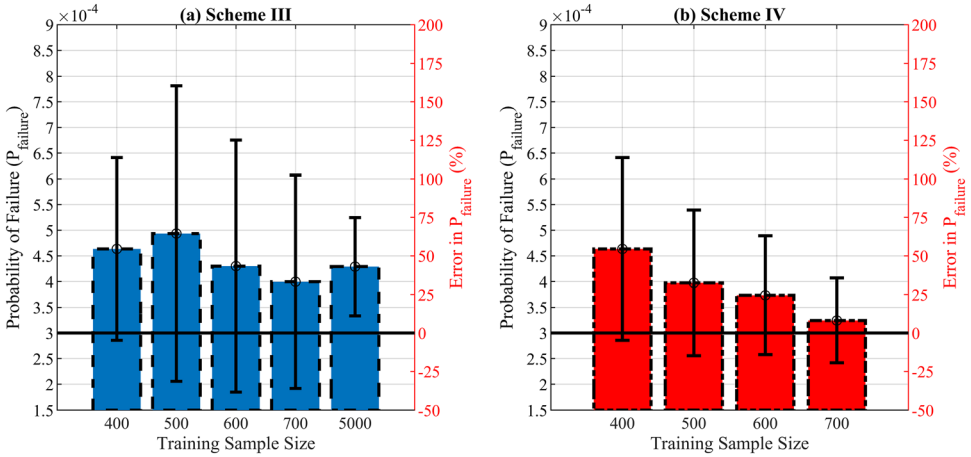

Fig. 11 Results of failure probabilities of Schemes III and IV calculated based on the 100 independent runs

Table 8 Results of failure probabilities of Schemes III and IV

Schemes	Scheme III			Scheme IV		
Training sample size	p_f	$\varepsilon_r(\%)$	CoV	p_f	$\varepsilon_r(\%)$	CoV
400	0.000464	54.5	0.38	0.000464	54.5	0.38
500	0.000494	64.5	0.58	0.000398	32.5	0.36
600	0.000430	43.3	0.57	0.000374	24.5	0.31
700	0.000400	33.2	0.52	0.000324	8.13	0.25
5000	0.000429	43.0	0.22	–	–	–

p_f: probability of failure (averaged value of 100 sets of analyses)

ε_r: error (%) in CNN-predicted p_f

CoV: coefficient of variation of predicted p_f over 100 sets of analyses

element mesh. This could then limit the use of this technique for large problems. In addition, the computational costs of a polynomial chaos expansion (PCE), which is a commonly used metamodel for such problems, can be sensitive to the method that discretizes the random field and the level of statistical heterogeneity of the random field. For example, Jiang and Huang [24] report that 1713 simulations are needed to construct a PCE for a problem with a similar level of P_{failure} and statistical heterogeneity as the illustrative example in the current study. In comparison, only 400 model simulations are needed for the training of CNNs in the present study.

Moreover, the use of CNNs, which read the random fields directly as inputs, makes this technique independent of the method used to discretize and generate the random fields [53]. On the other hand, the efficiency of a PCE, being a commonly used surrogate model, can be heavily dependent on the random field discretization strategy and the choice of autocorrelation function. In the literature, the random field discretization strategy is typically based on the truncated K-L expansion, while the autocorrelation function is commonly defined using the squared exponential form. The computational costs associated with the construction of a PCE can be quite high should other random field generation techniques and autocorrelation functions be used. In this regard, not only does the proposed CNN-based strategy provide a highly efficient computational tool, but it also offers engineers the flexibility to choose the random field generation technique and autocorrelation function that are deemed appropriate for the geotechnical engineering problem under study.

In addition, the ability of the trained CNNs in Scheme I to generate a large number of sample FoS values for the processing using the MaxEnt-FM technique, thus obviating the need for performing the time-consuming finite element simulations, makes repeated analysis viable. As highlighted in Figs. 8 and 11, the predicted P_{failure} can vary for repeated analyses with different datasets. The use of Scheme I makes it feasible to study the variability in the P_{failure} values with different datasets, as such datasets can be generated with reduced computational cost by the trained CNNs for the distribution fitting exercise using the MaxEnt-FM approach. The ability to examine possible variations in the predicted results with different datasets can help engineers make more informed engineering decisions.

7.3 Limitations

The implementation of a CNN model involves the design of an architecture and the selection of appropriate values for the configurable hyperparameters. Variations in architecture and values of hyperparameters can affect the performance of a CNN. In the present study, the architecture and hyperparameter values are determined using a trial-

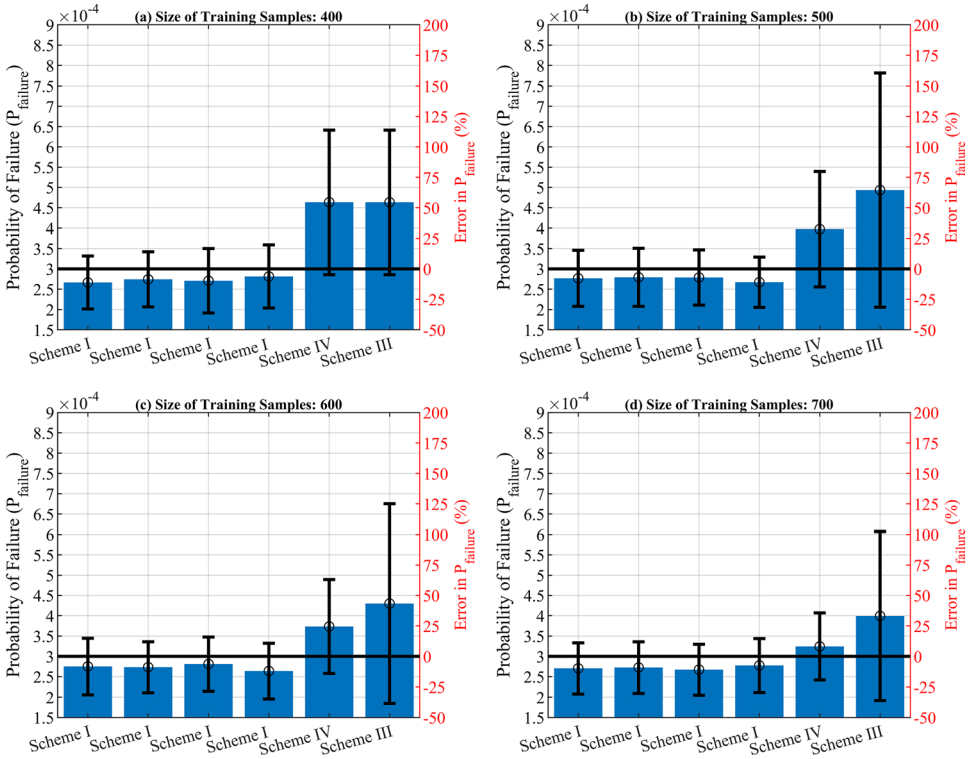

Fig. 12 Comparisons of failure probabilities of Schemes I, III, and IV calculated based on the 100 independent runs

and-error process. The information contained in existing published works can also provide useful insights and guidelines. The development of a "unified" architecture and guidelines on choosing appropriate values of hyper-parameters deserve further study.

In addition, the present study adopts a commercial software, Optum G2, to generate the random field and perform the strength-reduction calculations. The mapping of the generated random fields to the discretized finite element mesh is carried out by the software. However, in situations where the mapping has to be performed manually by the user, the optimal ratio between the discretized finite element mesh and the random field grid should be properly considered and addressed. Guidelines on this issue can be found in [48].

8 Conclusions

In summary, this paper presents a hybrid strategy (CNNs-MaxEnt-FM) for reliability analysis in spatially variable soils at low probability levels. The improved performance of the proposed strategy in terms of both accuracy and computational efficiency is demonstrated using a slope stability example. The specific conclusions of the current study are listed as follows:

(i) The conventional use of CNNs (Scheme III) is not sufficiently accurate for problems involving low probability levels.
(ii) An adaptive training strategy (Scheme IV) is effective in improving the performance of CNNs

Table 9 Computational time of all schemes and analyses

Methods	Number of FEM simulations	Computational time				
		FEM simulation (hours)	Training of CNN (hours)	Adaptive sampling (hours)	Optimization (hours)	Total (hours)
MCS	100,000	1680	–	–	–	1680
Scheme I	400–700	12	0.5	–	0.1	13
Scheme II	2000–5000	80	–	–	0.1	81
Scheme III	400–700	12	0.5	–	–	13
Scheme IV	400–700	12	0.5	4	–	17

in modelling the distribution tail and predicting the level of failure probability.

(iii) The MaxEnt-FM approach allows a good modelling of the distribution tail from a limited sample FoS values of the slope stability problem.

(iv) The proposed CNNs-MaxEnt-FM strategy accurately predicts the level of failure probability for the slope examined in the current study, this being achieved at a significant reduction in the computational costs compared to the other three schemes considered.

Acknowledgements This research was supported and conducted at the Centre for Protective Technology at the National University of Singapore.

Data availability Some or all data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Akaike H (1998) Information theory and an extension of the maximum likelihood principle. Springer. https://doi.org/10.1007/978-1-4612-1694-0_15
2. Au SK, Beck JL (2001) Estimation of small failure probabilities in high dimensions by subset simulation. Probab Eng Mech. https://doi.org/10.1016/S0266-8920(01)00019-4
3. Ang AHS, Tang WH (2007) Probability concepts in engineering planning and design: emphasis on application to civil and environmental engineering. Wiley
4. Abdel-Hamid O, Mohamed AR, Jiang H, Deng L, Penn G, Yu D (2014) Convolutional neural networks for speech recognition. IEEE Trans Audio Speech Lang Process. https://doi.org/10.1109/TASLP.2014.2339736
5. Abdeljaber O, Avci O, Kiranyaz S, Gabbouj M, Inman DJ (2017) Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. J Sound Vib. https://doi.org/10.1016/j.jsv.2016.10.043
6. Cho SE (2007) Effects of spatial variability of soil properties on slope stability. Eng Geol. https://doi.org/10.1016/j.enggeo.2007.03.006
7. Ching J, Phoon KK, Hu YG (2009) Efficient evaluation of reliability for slopes with circular slip surfaces using importance sampling. J Geotech Geoenviron Eng. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000035
8. Cho SE (2010) Probabilistic assessment of slope stability that considers the spatial variability of soil properties. J Geotech Geoenviron Eng. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000309
9. Cha YJ, Choi W, Büyüköztürk O (2017) Deep learning-based crack damage detection using convolutional neural networks. Computer-Aided Civ Infrastruct Eng. https://doi.org/10.1111/mice.12263
10. Cui F, Ghosn M (2019) Implementation of machine learning techniques into the subset simulation method. Struct Saf. https://doi.org/10.1016/j.strusafe.2019.02.002
11. Chen F, Wang L, Zhang W (2019) Reliability assessment on stability of tunnelling perpendicularly beneath an existing tunnel considering spatial variabilities of rock mass properties. Tunn Undergr Space Technol. https://doi.org/10.1016/j.tust.2019.03.013
12. El-Ramly H, Morgenstern NR, Cruden DM (2002) Probabilistic slope stability analysis for practice. Can Geotech J. https://doi.org/10.1139/t02-034
13. Erdogmus D, Hild KE, Rao YN, Príncipe JC (2004) Minimax mutual information approach for independent component analysis. Neural Comput. https://doi.org/10.1162/089976604773717595
14. Fenton GA, Griffiths DV (2002) Probabilistic foundation settlement on spatially random soil. J Geotech Geoenviron Eng. https://doi.org/10.1061/(ASCE)1090-0241(2002)128:5(381)
15. Griffiths DV, Huang J, Fenton GA (2009) Influence of spatial variability on slope reliability using 2-D random fields. J Geotech Geoenviron Eng. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000099
16. Gopalakrishnan K, Khaitan SK, Choudhary A, Agrawal A (2017) Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. Constr Build Mater. https://doi.org/10.1016/j.conbuildmat.2017.09.110
17. Goh ATC, Zhang WG, Wong KS (2019) Deterministic and reliability analysis of basal heave stability for excavation in spatial variable soils. Comput Geotech. https://doi.org/10.1016/j.compgeo.2018.12.015
18. Huang S, Mahadevan S, Rebba R (2007) Collocation-based stochastic finite element analysis for random field problems. Probab Eng Mech. https://doi.org/10.1016/j.probengmech.2006.11.004
19. He W, Li G, Hao P, Zeng Y (2019) Maximum entropy method-based reliability analysis with correlated input variables via hybrid dimension-reduction method. J Mech Des. https://doi.org/10.1115/1.4043734
20. Ireland HO (1954) Stability analysis of the congress street open cut in Chicago. Geotechnique. https://doi.org/10.1680/geot.1954.4.4.163
21. Inverardi PLN, Tagliani A (2003) Maximum entropy density estimation from fractional moments. Commun Stat Theory Methods. https://doi.org/10.1081/sta-120018189
22. Jaynes ET (1957) Information theory and statistical mechanics. Phys Rev. https://doi.org/10.1103/PhysRev.106.620
23. Ji S, Xu W, Yang M, Yu K (2013) 3D Convolutional neural networks for human action recognition. IEEE Trans Pattern Anal Mach Intell. https://doi.org/10.1109/TPAMI.2012.59
24. Jiang SH, Huang JS (2016) Efficient slope reliability analysis at low-probability levels in spatially variable soils. Comput Geotech. https://doi.org/10.1016/j.compgeo.2016.01.016
25. Ji J, Zhang C, Gao Y, Kodikara J (2018) Effect of 2D spatial variability on slope reliability: a simplified FORM analysis. Geosci Front. https://doi.org/10.1016/j.gsf.2017.08.004
26. Krabbenhøft K, Lyamin AV (2014) Optum G2. Optum Computational Engineering
27. Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Commun ACM. https://doi.org/10.1145/3065386
28. Lawrence S, Giles CL, Tsoi AC, Back AD (1997) Face recognition: a convolutional neural-network approach. IEEE Trans Neural Netw 10:1115. https://doi.org/10.1109/72.554195
29. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 10:159. https://doi.org/10.1109/5.726791
30. Low BK, Tang WH (2007) Efficient spreadsheet algorithm for first-order reliability method. J Eng Mech. https://doi.org/10.1061/(ASCE)0733-9399(2007)133:12(1378)
31. Li H, Lin Z, Shen X, Brandt J, Hua G (2015) A convolutional neural network cascade for face detection. Proc IEEE Comput

Soc Conf Comput Vis Pattern Recognit. https://doi.org/10.1109/CVPR.2015.7299170
32. Li DQ, Xiao T, Cao ZJ, Zhou CB, Zhang LM (2016) Enhancement of random finite element method in reliability analysis and risk assessment of soil slopes using subset simulation. Landslides. https://doi.org/10.1007/s10346-015-0569-2
33. Li DQ, Zheng D, Cao ZJ, Tang XS, Phoon KK (2016) Response surface methods for slope reliability analysis: review and comparison. Eng Geol. https://doi.org/10.1016/j.enggeo.2015.09.003
34. Li DQ, Zheng D, Cao ZJ, Tang XS, Qi XH (2019) Two-stage dimension reduction method for meta-model based slope reliability analysis in spatially variable soils. Struct Saf. https://doi.org/10.1016/j.strusafe.2019.101872
35. Li G, He W, Zeng Y (2019) An improved maximum entropy method via fractional moments with Laplace transform for reliability analysis. Struct Multidiscip Optim. https://doi.org/10.1007/s00158-018-2129-6
36. Liu WS, Cheung SH, Cao WJ (2019) An efficient surrogate-aided importance sampling framework for reliability analysis. Adv Eng Softw. https://doi.org/10.1016/j.advengsoft.2019.102687
37. Liu X, Li DQ, Cao ZJ, Wang Y (2020) Adaptive Monte Carlo simulation method for system reliability analysis of slope stability based on limit equilibrium methods. Eng Geol. https://doi.org/10.1016/j.enggeo.2019.105384
38. Müller J (2012) User guide for modularized surrogate model toolbox
39. Nair V, Hinton GE (2010) Rectified linear units improve restricted Boltzmann machines. In: ICML 2010 - proceedings, 27th international conference on machine learning
40. Oka Y, Wu TH (1990) System reliability of slope stability. J Geotech Eng. https://doi.org/10.1061/(ASCE)0733-9410(1990)116:8(1185)
41. Phoon KK, Kulhawy FH (1999) Characterization of geotechnical variability. Can Geotech J. https://doi.org/10.1139/t99-038
42. Phoon KK, Huang SP, Quek ST (2002) Simulation of second-order processes using Karhunen–Loeve expansion. Comput Struct. https://doi.org/10.1016/S0045-7949(02)00064-0
43. Pan Y, Shi G, Liu Y, Lee FH (2018) Effect of spatial variability on performance of cement-treated soil slab during deep excavation. Constr Build Mater. https://doi.org/10.1016/j.conbuildmat.2018.08.112
44. Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
45. Sudret B, Der Kiureghian A (2002) Comparison of finite element reliability methods. Probab Eng Mech. https://doi.org/10.1016/S0266-8920(02)00031-0
46. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. https://doi.org/10.1109/CVPR.2015.7298594
47. Tolson BA, Shoemaker CA (2007) Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resour Res. https://doi.org/10.1029/2005WR004723
48. Tabarroki M, Ching J (2019) Discretization error in the random finite element method for spatially variable undrained shear strength. Comput Geotech. https://doi.org/10.1016/j.compgeo.2018.10.001
49. Vanmarcke E (2010) Random fields analysis and synthesis. World Scientific
50. Wang L, Wu C, Tang L, Zhang W, Lacasse S, Liu H, Gao L (2020) Efficient reliability analysis of earth dam slope stability using extreme gradient boosting method. Acta Geotech. https://doi.org/10.1007/s11440-020-00962-4
51. Wang MX, Tang XS, Li DQ, Qi XH (2020) Subset simulation for efficient slope reliability analysis involving copula-based cross-correlated random fields. Comput Geotech. https://doi.org/10.1016/j.compgeo.2019.103326
52. Wang MY, Liu Y, Ding YN, Yi BL (2020) Probabilistic stability analyses of multi-stage soil slopes by bivariate random fields and finite element methods. Comput Geotech. https://doi.org/10.1016/j.compgeo.2020.103529
53. Wang ZZ, Xiao CL, Goh SH, Deng MX (2021) Meta-model based reliability analysis in spatially variable soils using convolutional neural networks. J Geotech Geoenviron Eng. https://doi.org/10.1061/(asce)gt.1943-5606.0002486
54. Wang ZZ, Goh SH (2021) Novel approach to efficient slope reliability analysis in spatially variable soils. Eng Geol. https://doi.org/10.1016/j.enggeo.2020.105989
55. Xu J (2016) A new method for reliability assessment of structural dynamic systems with random parameters. Struct Saf. https://doi.org/10.1016/j.strusafe.2016.02.005
56. Xiao C, Qin R, Huang X, Li J (2018) A study of using fully convolutional network for treetop detection on remote sensing dataA study of using fully convolutional network for treetop detection on remote sensing data.ISPRS Ann Photogramm Remote Sens Spat Inf Sci. https://doi.org/10.5194/isprs-annals-IV-1-163-2018
57. Xu J, Dang C (2019) A novel fractional moments-based maximum entropy method for high-dimensional reliability analysis. Appl Math Model. https://doi.org/10.1016/j.apm.2019.06.037
58. Zhang X, Pandey MD (2013) Structural reliability analysis based on the concepts of entropy, fractional moment and dimensional reduction method. Struct Saf. https://doi.org/10.1016/j.strusafe.2013.03.001
59. Zhang X, He W, Zhang Y, Pandey MD (2017) An effective approach for probabilistic lifetime modelling based on the principle of maximum entropy with fractional moments. Appl Math Model. https://doi.org/10.1016/j.apm.2017.07.036
60. Zhang J, Xiao M, Gao L (2019) An active learning reliability method combining Kriging constructed with exploration and exploitation of failure region and subset simulation. Reliab Eng Syst Saf. https://doi.org/10.1016/j.ress.2019.03.002
61. Zhang X, Low YM, Koh CG (2020) Maximum entropy distribution with fractional moments for reliability analysis. Struct Saf. https://doi.org/10.1016/j.strusafe.2019.101904
62. Zhu B, Hiraishi T, Pei H, Yang Q (2020) Efficient reliability analysis of slopes integrating the random field method and a Gaussian process regression-based surrogate model. Int J Numer Anal Meth Geomech. https://doi.org/10.1002/nag.3169
63. Zhang W, Zhang R, Wu C, Goh ATC, Lacasse S, Liu Z, Liu H (2020) State-of-the-art review of soft computing applications in underground excavations. Geosci Front 11(4):1095–1106. https://doi.org/10.1016/j.gsf.2019.12.003
64. Zhou Z, Li D-Q, Xiao T, Cao Z-J, Du W (2021) Response surface guided adaptive slope reliability analysis in spatially varying soils. Comput Geotech. https://doi.org/10.1016/j.compgeo.2020.103966
65. Zhang W, Li H, Li Y, Liu H, Chen Y, Ding X (2021) Application of deep learning algorithms in geotechnical engineering: a short critical review. Artif Intell Rev. https://doi.org/10.1007/s10462-021-09967-1
66. Zhang WG, Meng FS, Chen FY, Liu HL (2021) Effects of spatial variability of weak layer and seismic randomness on rock slope stability and reliability analysis. Soil Dyn Earthq Eng 146:106735. https://doi.org/10.1016/j.soildyn.2021.106735