Comput Geosci (2013) 17:739–755
DOI 10.1007/s10596-013-9353-3

ORIGINAL PAPER

Probabilistic slope stability analysis by a copula-based sampling method

Xing Zheng Wu

Received: 26 August 2012 / Accepted: 17 April 2013 / Published online: 23 May 2013

Abstract In probabilistic slope stability analysis, the influence of cross correlation of the soil strength parameters, cohesion and internal friction angle, on the reliability index has not been investigated fully. In this paper, an expedient technique is presented for probabilistic slope stability analysis that involves sampling a series of combinations of soil strength parameters through a copula as input to an existing conventional deterministic slope stability program. The approach organises the individual marginal probability density distributions of componential shear strength as a bivariate joint distribution by the copula function to characterise the dependence between shear strengths. The technique can be used to generate an arbitrarily large sample of soil strength parameters. Examples are provided to illustrate the use of the copula-based sampling method to estimate the reliability index of given slopes, and the computed results are compared with the first-order reliability method, considering the correlated random variables. A sensitivity study was conducted to assess the influence of correlational measurements on the reliability index. The approach is simple and can be applied in practice with little effort beyond what is necessary in a conventional analysis.

Keywords Probabilistic analysis · Slope stability · Monte Carlo simulation · Copula · Cross correlation · Cohesion · Friction angle

X. Z. Wu (✉)
Department of Applied Mathematics, School of Applied Science, University of Science and Technology Beijing,
30 Xueyuan Road, Haidan District, Beijing
100083, People's Republic of China
e-mail: xingzhengwu@gmail.com

1 Introduction

Slope stability analysis is a traditional problem in geotechnical engineering that is highly amenable to probabilistic treatment and that has received considerable attention recently [19, 52]. Uncertainties in soil properties, environmental conditions and theoretical models are the most important sources of lack of confidence in deterministic analysis [1, 5]. There have been numerous attempts to use a probabilistic approach complementary to the conventional approach to analyse the safety of slopes and especially to explore the effect of variabilities in soil shear strengths. A common approach to determine the reliability of a slope is based on calculating the reliability index corresponding to a surface with the minimum factor of safety (referred to as the critical deterministic surface, defined by a limit equilibrium approach of slices), as described by [12] and [15]. However, critical slip surfaces may not necessarily be those with the lowest conventional factors of safety [30] but rather are determined by a combination of the mean factor of safety and uncertainty [6, 14].

It is therefore imperative that greater use is made of probabilistic assessments of slope stability and that capabilities for considering the statistical variation of input properties are enhanced [19]. These reliability model approaches do provide a better basis for making engineering judgments in a more transparent way. However, correlations between the cohesion c and the internal friction angle φ (referred to as the friction angle hereinafter) are commonly ignored in probabilistic slope stability analysis [15, 33, 38, 46, 58, 60]. A number of recent studies have been oriented toward careful consideration of the complicated nature of these correlations [10, 21, 23, 28, 39]. Most of these investigators believe that the cross correlation is negative, with a value between −0.24 and −0.70. However, several researchers have reported a positive correlation [39, 62].

Dependencies among the uncertainties in the estimates of these parameters can be critical to obtaining correct numerical results from reliability analyses in geotechnical engineering (see, e.g., [48]). Cho and Park [11] reported their findings on stochastic behaviour in a bearing capacity problem. The assumption of independence between cohesion and friction angle gives conservative results if the actual correlation is negative, but slightly unconservative results are obtained if the actual correlation is positive. Lü and Low [40] investigated the probability of failure with respect to the plastic zone criterion of underground rock excavations, and they concluded that assuming uncorrelated friction angle and cohesion will generate a higher probability of failure than assuming that these shear strength parameters are negatively correlated.

The influence of the correlation between strength parameters on slope stability analyses is often not well understood. Some researchers have shown that the probability of failure in slope stability analysis is insensitive to the correlation coefficient between the strength parameters [31]. However, the influence of cross correlation between the strength parameters on the reliability index of slope stability has been reported by some others [13, 62]. Interestingly, an accurate and reliable statistical description should be required to reproduce the multivariate joint characteristics of all the relevant marginal laws (the joint probability distribution of c and φ), considering the dependent relationships (their cross correlation) effectively in slope stability analysis. Recent advances in mathematics show how copulas [34, 45, 54] may be very useful in modelling dependence between correlated random variables. The detailed theoretical background and descriptions of copulas can be found in the literature.

Copulas represent an efficient tool for investigating the statistical behaviour of dependent variables. Specifically, copulas are operators on the family of one-dimensional probability distributions of random variables that yield multivariate laws with well-defined properties [55]. Their efficiency lies in the possibility of studying marginal behaviours and global dependence separately. In fact, it is precisely the copula that captures many of the features of a joint distribution: it is possible to prescribe the properties of a multivariate law simply by working on the structure of the corresponding copula. The flexibility offered by copulas for constructing joint distributions is evident from related studies in civil engineering (for a thorough review, see [54]) and in finance [20].

In addition, the concept of a copula is relatively simple; the construction does not constrain the choice of marginal distributions, and it provides a good way to impose a dependence structure on predetermined marginal distributions [16, 37]. Particularly when the normality assumption for data usually does not provide an adequate approximation to data sets with heavy-tail, non-normal multivariate distributions are used in practice (see [36]). Thus, a non-normal multivariate distribution is particularly useful when a geotechnical engineering problem involves the dependence properties of the random variables.

To obtain accurate quantitative predictions of the probability of failure of a slope system, the joint probability characteristics of multivariate random soil parameters, incorporating the dependence structure among parameters through a copula, should be implemented in a conventional slope stability approach. Then, a parametric study on the calculated reliability index should be carried out for a range of dependence properties to explore the influence of correlation extremes on reliability assessment. To achieve this goal, a methodology was developed within a probabilistic framework for analysing slope stability using random samplings to represent the various cross correlations of soil strength properties. A joint probability distribution of the strength parameters is derived through a copula for the probabilistic slope stability analysis to obtain the desired reliability index. The reliability indices obtained by this copula-based sampling technique are compared with the results obtained by the first-order reliability method (FORM).

This paper presents a description of cross correlation between cohesion and friction angle as determined by shear strength tests and definitions of their correlation measurements in Section 2. The copula theory, including the construction of the joint description of cross-correlated shear strength parameters and the forecasting of dependent random variates through copula, is presented in Section 3. Full details of the methodologies for calculating the reliability index of slope stability by copula-based random sampling are discussed in Section 4, and several examples are presented to demonstrate the effects of correlations between shear strength parameters on reliability indices by parametric sensitivity analysis. Discussion and conclusions are presented in Sections 5 and 6, respectively.

2 Marginal distributions and cross correlation characteristics of soil shear strength parameters

2.1 Marginal distributions of soil strength variables

An issue in the applicability of measured values of soil shear strength parameters is the consideration of whether soil properties follow normal distributions. The applicability of the normal distribution to soil properties is supported by [39, 60], and [5]. Brejda et al. [9] and Fenton and Griffiths [21] found it difficult to fit a normal distribution to sampled soil properties, but a log-normal distribution showed a better fit to their data. Other distributions, such as the triangular, the versatile beta, and the generalised gamma distributions, are gaining popularity [5]. The best-fit criteria for marginal distributions are identified by the Anderson–Darling (AD) [3] test initially with p_m (AD statistic). However, because it does not account for the estimated number of parameters, the Akaike information criterion (AIC) [2] values should be considered. The smaller the AIC value, the better the fit is. The AIC is defined as follows:

$$\begin{aligned}\mathrm{AIC} = &-2 \times \mathrm{Ln}\,(\text{maximized likelihood for the model})\\ &+2 \times \text{number of fitted parameters}\end{aligned} \quad (1)$$

2.2 Dependency measures

Soil shear strength pairs based on the Mohr–Coulomb criterion are associated with a single observation, so they are not independent. The dependence between random variables is best determined using Pearson's linear correlation coefficient ρ_{p}, as reported by some investigators [10, 23, 39]. More extensive discussion of this important subject requires more data that are realistic, and the development of techniques for reproducing or establishing the correlations while maintaining the desired accuracy is crucial to probabilistic assessment.

Let 'observed' pairs $(z_{1i}, z_{1j}), \ldots, (z_{ni}, z_{nj})$ be drawn from a multivariate population of (Z_i, Z_j), where n is the number of observations. Pearson's product–moment correlation coefficient ρ_{p} between two random variables Z_i (cohesion) and Z_j (friction angle) is usually written as follows:

$$\rho_{\mathrm{p}}\left(Z_i, Z_j\right) = \frac{\mathrm{CoV}\left(Z_i, Z_j\right)}{\sigma\left(Z_i\right)\sigma\left(Z_j\right)} \quad (2)$$

where $\mathrm{CoV}(Z_i, Z_j)$ is the covariance between Z_i and Z_j, $\mathrm{CoV}(Z_i, Z_j) = \mu(Z_i, Z_j) - \mu(Z_i)\mu(Z_j)$. $\mu(Z_i)$ and $\sigma(Z_j)$ denote the mean and standard deviation of Z_i, respectively. ρ_{p} is restricted to the interval from -1 to 1. As stated by [20] and [8], it is not necessarily informative for non-normal distributions.

Kendall's tau, which uses concordant or discordant values, is simply the probability of concordance minus the probability of discordance for the bivariate random pairs (Z_i, Z_j):

$$\begin{aligned}\tau\left(Z_i, Z_j\right) = &\Pr\left(\left(Z_i - \widetilde{Z}_i\right)\left(Z_j - \widetilde{Z}_j\right) > 0\right)\\ &- \Pr\left(\left(Z_i - \widetilde{Z}_i\right)\left(Z_j - \widetilde{Z}_j\right) < 0\right)\end{aligned} \quad (3)$$

Obviously, Kendall's tau is calculated by looking at the ordering of the sample for each variable of interest rather than the actual numerical values. Having defined the indicator variable $A_{ij} = sign\left(Z_{ti} - \widetilde{Z}_{si}\right)\left(Z_{tj} - \widetilde{Z}_{sj}\right)$, as in [42], one notices that an unbiased empirical estimator of Kendall's coefficient τ can be written as follows:

$$\tau\left(Z_i, Z_j\right) = \frac{\sum\limits_{1 \le t \le s \le n} A_{ij}\,(s, t)}{\frac{1}{2} n\,(n-1)} \quad (4)$$

where $sign$ is expressed by sign $= \begin{cases} 1, & \left(Z_{ti} - \widetilde{Z}_{si}\right)\left(Z_{tj} - \widetilde{Z}_{sj}\right) \ge 0, \text{ concordance} \\ -1, & \left(Z_{ti} - \widetilde{Z}_{si}\right)\left(Z_{tj} - \widetilde{Z}_{sj}\right) < 0, \text{ discordance} \end{cases}$, and $\left(\widetilde{Z}_i, \widetilde{Z}_j\right)$ is an independent copy of the vector (Z_i, Z_j). Equation 4 is the empirical approximation of the theoretical Kendall's tau in Eq. 3. The range of values of Kendall's correlation coefficient is -1 to $+1$.

3 Understanding the relationships between shear strength parameters using a copula

As Sklar's theorem [56] states that for any joint bivariate distribution function $H_{zi\,zj}\,(z_i, z_j)$, say, with marginal distribution functions $F_{Z_i}\,(z_i)$ and $F_{Z_j}\left(z_j\right)$, there exists at least one copula C such that, for all $z_i, z_j \in \mathrm{R}$, $H_{z_i z_j}\left(z_i, z_j\right) = C\left(F_{zi}\left(z_i\right), F_{zj}\left(z_j\right)\right)$. If $F_{Z_i}\,(z_i)$ and $F_{Z_j}\left(z_j\right)$ are continuous, then $C(u_i, u_j)$ is unique; otherwise, $C(u_i, u_j)$ is uniquely determined for the range of $F_{Z_i}\,(z_i)$, which is multiplied by the range of $F_{Z_j}\left(z_j\right)$. Thus, the joint bivariate distribution of (Z_i, Z_j) is connected with their one-dimensional marginal probability distributions $F_{Z_i}\,(z_i)$ and $F_{Z_j}\left(z_j\right)$ through copula [45]. In applying probability transforms $u_i = F_{Z_i}\,(z_i)$ and $u_j = F_{Z_j}\left(z_j\right)$ to Z_i and Z_j, there exists a bivariate joint distribution function with standard uniform marginals $C(u_i, u_j)$ [18, 56], such that

$$C\left(u_i, u_j\right) = H_{Z_i, Z_j}\left(F_{Z_i}^{-1}\,(u_i), F_{Z_j}^{-1}\left(u_j\right)\right) \quad (5)$$

where $0 \le u_i \le 1$ and $0 \le u_j \le 1$. If F is strictly increasing, F^{-1} is a quasi-inverse (or quantile) of F. Equation 5 gives an expression for copulas in terms of a joint distribution function H and the 'inverse' of the two margins. Moreover, Eq. 5 shows how copulas express dependence on a quantile scale, which provides a means of generating pseudo-random samples from general classes of multivariate probability distributions. That is, given a procedure to generate a sample (u_i, u_j) from the copula distribution, the required sample can be constructed as $\left(z_i, z_j\right) = \left(F_{Z_i}^{-1}\,(u_i), F_{Z_j}^{-1}\left(u_j\right)\right)$ (which we will return to later).

Copulas are consulted on the assumption that marginal distributions are known or can be estimated from the data. The procedure for constructing the joint distribution is flexible because no restrictions are placed on the marginal distributions [16, 42]. In other words, marginal distributions of any form can be knitted together to obtain their joint distribution, which is the main reason for the popularity of the copula theory in many areas of research [20, 37, 67]. Most importantly, this approach can handle arbitrarily complicated dependence between the input variables. This makes the approach more general than methods implemented in some risk analysis software packages that incorporating correlations but not dependence [22].

There are many different copulas to choose from, varying in correlation properties such as symmetry, tail dependence and range of dependence [34, 45]. Considering the correlation characteristics between soil strength parameters, we can choose the normal copula and Student copula from the elliptical class of copulas, the Clayton, Frank and Gumbel copula from the Archimedean class, and the Plackett copula in a class of its own. These copulas are listed in Table 1, along with their parameter ranges. Some of these copulas may not allow negative correlation, but negating the values of one variable can achieve a positive value for the correlation. For some general comments on the choice and further details of copulas, the interested reader should consult [34, 45] and [42]. The following is a brief summary of the theory behind these popular copulas, limited to two-dimensional copulas for the sake of brevity.

3.1 Elliptical class of copulas

The bivariate normal copula is defined as follows:

$$C_\rho^G\left(u_i,u_j;\rho_{\mathrm{p}}\right)=\Phi_\rho\left(\Phi^{-1}\left(u_i\right),\Phi^{-1}\left(u_j\right);\rho_{\mathrm{p}}\right)$$
$$=\int_{-\infty}^{\Phi^{-1}(u_i)}\int_{-\infty}^{\Phi^{-1}(u_j)}\frac{1}{2\pi|\Sigma_2|^{1/2}}\exp\left(-\frac{1}{2}\boldsymbol{W}^T\left(\Sigma_2\right)^{-1}\boldsymbol{W}\right)\mathrm{d}\boldsymbol{W} \tag{6}$$

where $\Phi_\rho(\cdot)$ is a joint distribution function of a bivariate normal distribution with zero mean and variance–covariance matrix Σ_2. $\Phi(t)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}\mathrm{d}t$ is the normal distribution, and $\Phi^{-1}(t)$ is the quantile function of the univariate standard normal distribution. The integral variable $\boldsymbol{W}=\begin{Bmatrix}t_i\\t_j\end{Bmatrix}$, and $\Sigma_2=\begin{bmatrix}1 & \rho_{\mathrm{p}}^{ij}\\ \rho_{\mathrm{p}}^{ij} & 1\end{bmatrix}$ is a symmetrical covariance matrix with the linear Pearson's correlation coefficient ρ_{p}. ρ_{p}^{ij} represents the correlation coefficient between Z_i and Z_j.

The Student copula has two parameters, one corresponding to the dependence parameter and the other to the number of degrees of freedom λ. The number of degrees of freedom controls the heaviness of the tails, and as it increases, the copula approaches the normal copula. Both the normal and Student copulas are symmetric, and the normal copula is a limiting case of the Student copula when λ becomes infinity (λ is set to 9 in this study). The advantage of the Student copula is that it can capture lower- and upper-tail dependence in the data (i.e., joint non-exceedance and exceedance probabilities for rare events; see [42] for details).

3.2 Archimedean class of copulas

The widely used copulas in the Archimedean class [45] are constructed in a completely different way from the normal copula. An important component of constructing an Archimedean copula is an explicit generator function ϕ_θ. An Archimedean copula is usually written as follows:

$$C_\phi\left(u_i,u_j;\theta\right)=\phi_\theta^{-1}\left(\phi_\theta\left(u_i\right),\phi_\theta\left(u_j\right);\theta\right) \tag{7}$$

where ϕ_θ is a convex decreasing function with $\phi_\theta(1)=0$, $\phi_\theta^{-1}(\cdot)$ is the pseudo-inverse of $\phi_\theta(\cdot)$, and θ is a copula dependence parameter or associated parameter. The definitions of the generator function for this family

Table 1 Summary of the adopted bivariate copula functions and their dependence parameters

Family	Copula function	Generator function	Relationship between θ and τ	Range of θ
Normal	$N_\theta\left(\Phi^{-1}\left(u_i\right),\Phi^{-1}\left(u_2\right)\right)$	/	$\theta=\sin(\pi\tau 2)$	[-1,1]
Student	$T_{\theta,\lambda}\left(T_\lambda^{-1}\left(u_1\right),T_\lambda^{-1}\left(u_2\right)\right)$	/	$\theta=\sin(\pi\tau/2)$	[-1,1]
Clayton	$\left(u_1^{-\theta}+u_2^{-\theta}-1\right)^{-1/\theta}$	$(t^{-\theta}-1)/\theta$	$\theta=2\,\tau\,/\,(1-\tau)$	[-1,∞]–{0}
Frank	$-\frac{1}{\theta}\mathrm{Ln}\left(1+\frac{(\exp(-\theta u_1)-1)(\exp(-\theta u_2)-1)}{\exp(-\theta)-1}\right)$	$-\mathrm{Ln}\frac{\exp(-\theta t)-1}{\exp(-\theta)-1}$	$\tau=1-\frac{4}{\theta}+\frac{4}{\theta^2}\int_0^\theta\frac{t}{\exp(t)-1}\mathrm{d}t$	(-∞,∞)-{0}
Gumbel	$\exp\left(-\left[(-\mathrm{Ln}u_1)^\theta+(-\mathrm{Ln}u_2)^\theta\right]^{1/\theta}\right)$	$(-\ln t)^\theta$	$\theta=1/(1-\tau)$	[1,∞)
Plackett	$\left[1+(\theta-1)(u_1+u_2)-\mathrm{TermA}^{1/2}\right]/\{2(\theta-1)\}$ TermA$=\{1+(\theta-1)(u_1+u_2)\}^2-4u_1u_2\theta(\theta-1)$	/	Obtained numerically	(0,∞)–{1}

$^{\Phi}$cumulative distribution function of the standard normal distribution;

$^{N_\theta}$cumulative distribution function of the standard bivariate normal distribution with Pearson correlation θ;

$^{T_\lambda}$cumulative distribution function of the Student with λ degrees of freedom; T_θ,

$^{\lambda}$cumulative distribution function of the bivariate Student distribution with λ degrees of freedom of copulas are given in Table 1. The Frank copula is a symmetric copula; the Clayton and Gumbel copulas are asymmetric Archimedean copulas. The Clayton copula exhibits greater dependence in the negative tail than in the positive, but the Gumbel copula (also known as the Gumbel–Hougard copula) exhibits greater dependence in the positive tail than in the negative tail.

3.3 Plackett copula

The Plackett copula is the best known example of an algebraically constructed copula. The association θ is determined by the odds ratio, based on observed frequencies in the four quadrants, rather than on the correlation of random variables [45].

3.4 Relationship between Kendall's Tau and the copula's parameter

For the elliptical class of copulas, there is a relationship between the linear correlation ρ_{p} and the rank correlation τ [24]

$$\tau\left(Z_i, Z_j\right) = \frac{2}{\pi}\arcsin\left(\rho_{\mathrm{p}}^{ij}\right) \tag{8}$$

where $\arcsin(t)$ is an inverse trigonometric function such that $\sin(\arcsin(t)) = t$. This expression prompts the alternative estimation of ρ_{p}. The use of Eq. 8 may be more advantageous because τ is rank-dependent and invariant with respect to strictly monotonic nonlinear transformations.

For the Archimedean class, [25] have shown that τ depends on the generator $\phi_\theta(\cdot)$ and its derivative, according to the following simple form:

$$\begin{aligned}\tau\left(Z_i, Z_j\right) &= 4\int_0^1\int_0^1 C\left(u_i, u_j; \theta\right)\mathrm{d}C\left(u_i, u_j; \theta\right) - 1 \\ &= 1 + 4\int_0^1 \frac{\phi_\theta(t)}{\phi_\theta^{'}(t)}\mathrm{d}t\end{aligned} \tag{9}$$

The explicit function of this expression for the copulas is given in Table 1. If Kendall's tau is known, the correlation parameter of the copula θ can be estimated using this expression. An illustration of the correlation between ρ_{p} or θ and Kendall's τ is shown in Fig. 1.

3.5 Identification of the best-fitting copula

The goodness-of-fit for the alternative copulas is usually assessed using the Cramér–von Mises statistic [27]. The Cramér–von Mises statistic is based on the empirical process of comparing the empirical copula with a parametric estimate of the copula derived under the null hypothesis

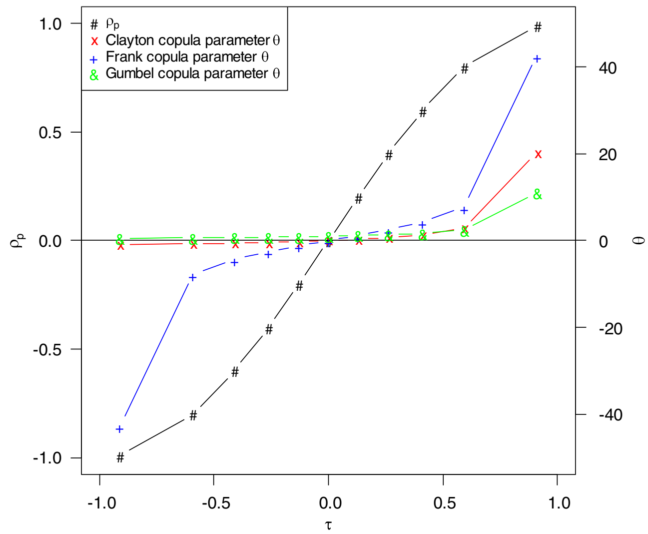

Fig. 1 Relationships of Pearson's rho and copula parameter theta with Kendall's tau

H_0. The Cramér–von Mises function represents a type of distance between the true and the observed copula:

$$S^n = \sum_{i=1}^{n}\left\{C^n\left(U_1^{i,n}, U_2^{i,n}\right) - C^{\theta_n}\left(U_1^{i,n}, U_2^{i,n}\right)\right\}^2 \tag{10}$$

where C^n is the empirical copula used as the most objective benchmark and C^{θ_n} is an estimator of C under the hypothesis that H_0: $C \in \{C^\theta\}$ holds. Here, θ^n is an estimator of θ computed from the ranked pseudo-observations $\left(U_1^{1,n}, U_2^{1,n}\right), \ldots, \left(U_1^{n,n}, U_2^{n,n}\right)$ and could be estimated via the inversion of Kendall's τ. Large values of S^n lead to the rejection of H_0. Approximate p_c values for the test function S^n are obtained using a parametric bootstrapping approach [35]. The p_c value represents the level at which the copula is not rejected, meaning that models with higher p_c values are better in terms of not being rejected.

The best-fitting copula from among the candidate copulas for the set of shear strengths is assessed in terms of the AIC [2]. The copula associated with the smallest AIC value is considered to be the best-fitting copula.

3.6 Copula-based sampling with correlation

The copula provides a convenient way to fit each variable to a distribution separately and then joins the marginal distributions together through their dependence [47]. Thus, copula-based sampling makes it possible to reconstruct the dependence structures of these observed data sets by random draws from the above copula functions. In particular, if (u_i, u_j) is a random draw from a copula, then $\left(z_i, z_j\right) = \left(F_{Z_i}^{-1}(u_i), F_{Z_j}^{-1}(u_j)\right)$ is a random draw from the joint distribution $H_{Z_iZ_j}\left(z_i, z_j\right) = C\left(F_{Z_i}(u_i), F_{Z_j}(u_j)\right)$.

Generating random samples from the distributions that correspond to those copulas are associated with a variety of algorithms called copula-based sampling methods (CBSM).

The simulation of copulas can, in principle, be based on the conditional distribution approach, which is appealing because only univariate simulations are required. The main steps of this technique are the following [42]:

1. Generate two independent uniform (0,1) variates u_i and x.
2. Set $u_j = C_{u_i}^{-1}(x)$, where $C_{u_i}(u_j) = \frac{\partial}{\partial u_i} C(u_i, u_j)$ is a conditional copula, and $C_{u_i}^{-1}$ denotes a quasi-inverse of C_{u_i}.
3. The desired pair of cumulative distribution functions is (u_i, u_j).
4. The desired variates or realisations are $z_i = F_i^{-1}(u_i)$ and $z_j = F_j^{-1}(u_j)$, where F_i and F_j are cumulative distribution functions. (z_i, z_j) is a quantile pair of random vectors, i.e., the cohesion and friction angle.

Unfortunately, for most copulas, the function $C_{u_i}^{-1}$ does not exist in closed form. In this case, after sampling u_i, one has to use a root-finding routine to obtain u_j. The common way of proceeding is thus based on specific techniques for various classes of copulas.

For elliptical copulas, the Choleski decomposition provides an easy solution in the normal and Student cases [42, 64]. For Archimedean copulas, the Laplace transformation of the inverse of the generator exists in closed form. A general simulation procedure exists that uses an approach [42, 64] based on the first derivation by [41]. This approach requires generating random numbers from a positive random variable K, often called frailty: in particular, for the simulation of Clayton, Frank and Gumbel copulas, K is the gamma, log-series and positive stable [64]. For the remaining copulas, essentially no method is available except the conditional distribution approach.

3.7 Application of the CBSM to the soil shear strength pairs by [39]

Taking data obtained for soils in Class BL-2 by [39] as an example to illustrate the above procedures, the values of the cohesion and friction angle obtained from 45 core samples are shown in Fig. 2. The surface soils in the decomposed granite area, named clayey coarse sands, were collected as samples to carry out consolidated undrained triaxial tests. The strength pairs are dependent variables, as shown in Table 2, with a correlation coefficient τ of −0.236 (the corresponding ρ_{p} is −0.382; however, a value of −0.43 was reported in Lumb's investigation, which may result from digital interpretation of the data set in the figures).

Among various possible candidate marginal distributions for the cohesion and friction angle, the following functions are generally used for goodness-of-fit: the normal, log-normal, Gumbel, Weibull and gamma distributions. No detailed explanation of these distributions is given here

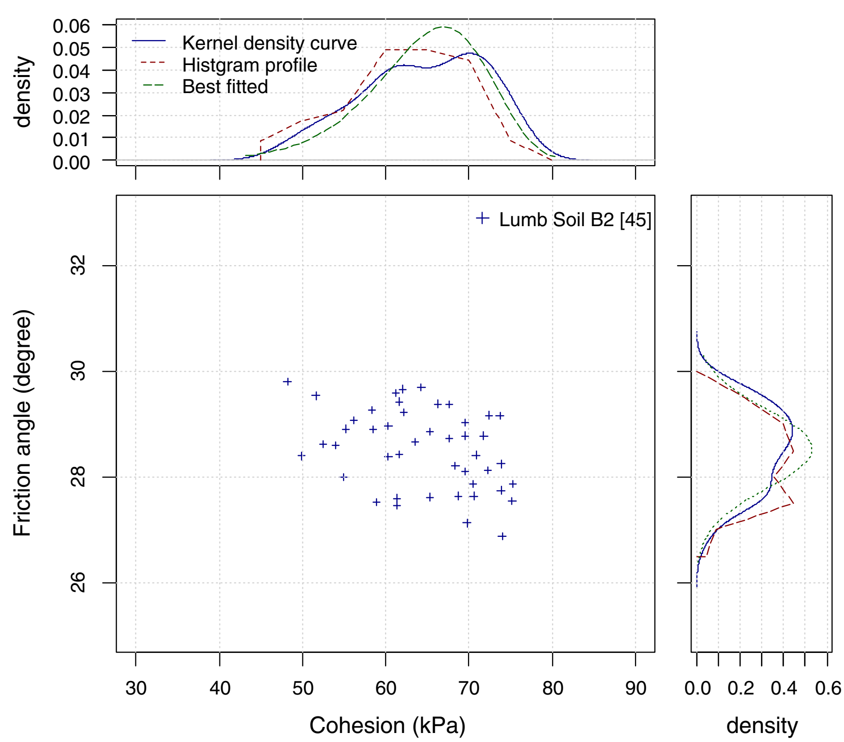

Fig. 2 Paired data for cohesion and internal friction angle (obtained from [39])

Table 2 Mean, standard deviation and correlation coefficient for soil BL-2

Cohesion (kPa)			Friction angle (deg)			Correlation	
Mean	Standard deviation	Best fitting	Mean	Standard deviation	Best fitting	ρ_p	τ
64.443	7.358	Weibull (shape = 10.76, scale = 67.62)	28.531	0.757	Normal	−0.382	−0.236

because they are readily available in many standard textbooks [43]. The R package [51] routine 'fitdistrplus', which gathers tools for choosing and fitting a parametric univariate distribution to a given data set [50], was utilised here to compute the AD statistic p_m and AIC values listed in Table 3. The p_m value is greater than the significance levels usually mentioned in the statistical literature [4] for most of the candidate distributions. The values of AIC provide an objective way of determining which model among a set of models is most parsimonious. To obtain further intuitive knowledge on the distribution of the strengths, quantile–quantile plots can be developed to compare two distributions by plotting their quantiles (or percentiles) against each other. The quantiles of observed distributions of cohesion and friction angle are plotted against the quantiles of the standard normal distribution (i.e., the normal distribution with a mean of 0 and a standard deviation of 1) in Fig. 3a, b, respectively. If the observed data have a standard normal distribution, the points on the plot will fall approximately along the reference line $Y = X$. The greater the departure from the reference line, the greater the evidence for the conclusion that the data set have come from a population with a different distribution. Overall, the fit of data to a normal distribution is good, although the distribution struggles slightly with the extreme tail of the distributions.

By combining the individual marginal models of soil shear strengths with the rank correlation estimated from the observed pairs, any copula can be used to build a multivariate model that is consistent with the available information. The R package routine 'copula' helps to build and study multivariate modelling for fitting copulas [64, 65]. After a 'mvdc' class designed to construct multivariate distributions with given margins and their dependence using copulas is imposed, the package easily allows the generation of random variables through 'rmvdc' function or 'rcopula'. The command 'gofCopula', where, by default, the approximate p_c values for the test statistics are obtained using the parametric bootstrap, makes the goodness-of-fit test procedure easier to compute. These R packages are freely available at the Comprehensive R Archive Network (cran.r-project.org).

The Cramér-von Mises statistic p_c values and the AIC are listed in Table 4. Usually, two or more copulas are not rejected if their p_c values are greater than 0.05. However, the Clayton copula gives a slightly lower AIC value (−10.99) than the one (−6.17) given by the normal copula, which indicates that the Clayton copula is better suited to this set of observations. Visual scatter plots of realisations from the best-fitting copula are shown in Fig. 4. Only 200 random samples are selected for legibility. The confidence region (CR) is defined in the original physical space of two random variables to characterise the spread of the sampled data in different directions. At the 95 % confidence level, the confidence curves for both the observed (enclosed area I_o) and predicted data (enclosed area I_p), determined using a 2D kernel density estimator ('kde2d' of MASS package in R, see [61]) using 300 grid points in each direction are illustrated in this graph. To quantify the differences of these confidence regions, the percentage form of relative change d_{area} between the simulated and measured regions can be expressed by the ratio of the absolute change and divided by the measured region, i.e., $d_{area} = \frac{\mathrm{abs}(I_P - I_o)}{I_o} \times 100$. Here, the I_o associated with the measured region is taken as a reference value. If the relative percentage difference d_{area} is large, the predictions are less valuable than the observations. The relative percentage area difference d_{area} of predictions is calculated as 3.18 %. This graphical technique can provide an alternative tool for understanding the performance

Table 3 AD statistic and AIC of marginal distributions for soil BL-2

Type	AD statistic p_m					AIC				
	Norm	Lognorm	Gumbel	Weibull	Gamma	Norm	Lognorm	Gumbel	Weibull	Gamma
c (kPa)	0.6	0.75	1.13	0.49	0.69	310.32	312.77	319.68	307.54	311.81
φ (degree)	0.39	0.46	0.83	0.44	0.45	105.7	106.03	112.69	105.75	105.93

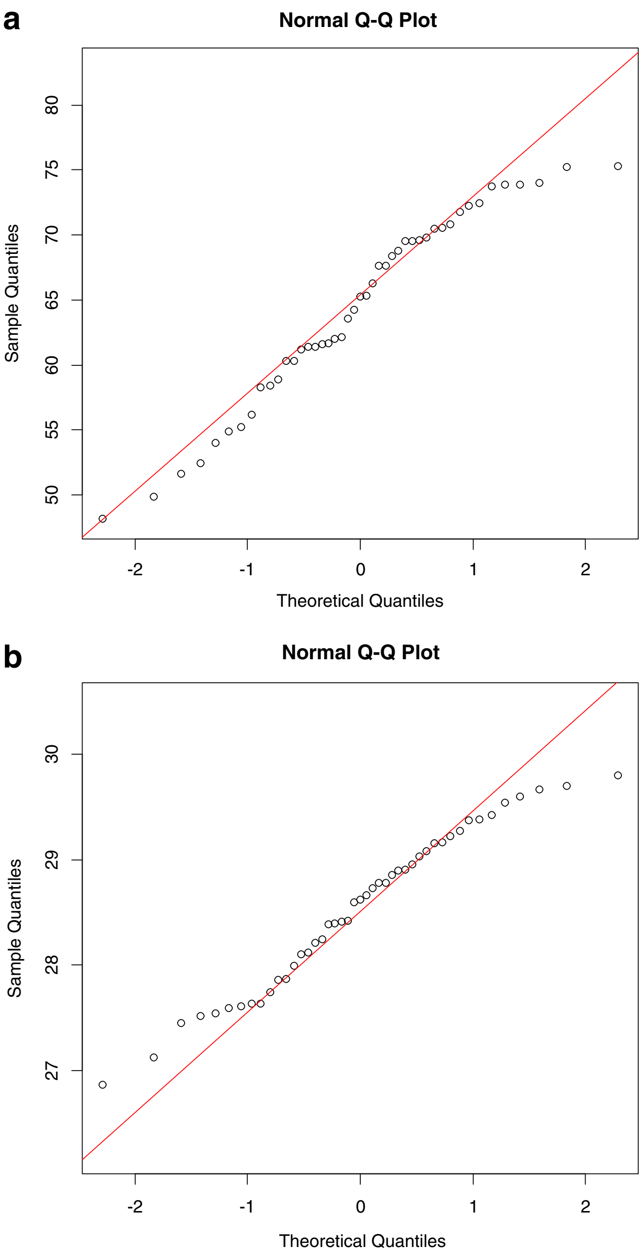

Fig. 3 **a** Observed and theoretical quantiles for quantile–quantile plots comparing observations of the cohesion of soil BL-2 by [39] under a normal distribution. **b** Observed and theoretical quantiles for quantile–quantile plots comparing observations of the friction angle of soil BL-2 by [39] under a normal distribution

of a simulation and preselecting appropriate copulas. Visual examination suggests that the copula model does an adequate job of mimicking the true distribution and maintaining the correlation relationships of these observed data.

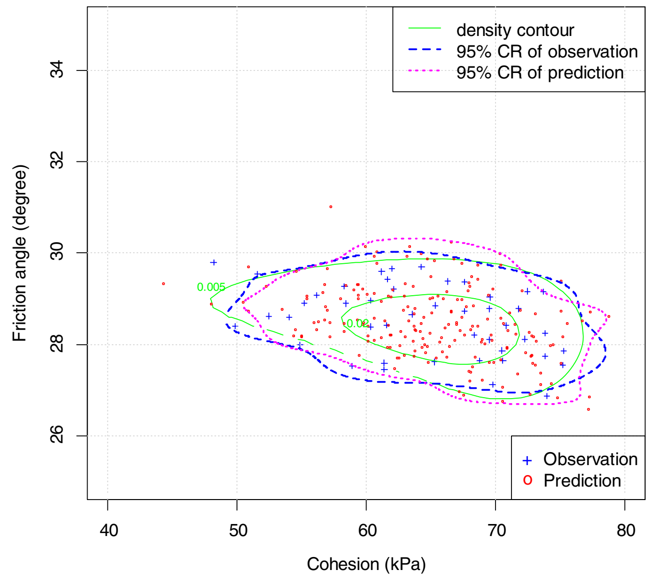

Fig. 4 Contour of the best-fitting copula (Clayton) confidence regions of the simulated and observed data for soil BL-2 [39]

For the BL-2 soil studied by [39], Fig. 5 illustrates a further comparison of the density contours for the bivariate pair of (c, φ) for different models. As this figure shows, the level curves of the empirical density for a bivariate normal distribution model (the values of the mean and standard deviation are taken from Table 2) are elliptical, whereas the level curves of the density through copulas with the best-fitting marginal distributions (listed in Table 2) take a different shape. The observed data are superimposed on the contour plots. The Clayton copula with the best-fitting margins provides a much better fit to the bivariate shear strength pairs than the traditional bivariate normal distribution model. A distinct advantage of normal copulas is their ability to capture dependence behaviours often observed in geotechnical engineering. The normal copula with the best-fitting margins provides a distribution that is quite similar, although not identical, to the one provided by the bivariate Clayton copula, and this distribution more reasonably represents the observed data than does the traditional model.

The information obtained from the results of a limited number of tests can only reflect a small part of the entire truth. For instance, the 45 observations in this example are very meagre multivariate data. Nevertheless, in engineering practice, including geotechnical design and analysis, it is often necessary to assume that the engineering behaviour indicated by limited data is true of an entire engineering

Table 4 AIC and p_c values of various copulas for soil BL-2

Type	Normal	Student ($df = 9$)	Clayton	Frank	Gumbel	Plackett
p_c	0.553	0.158	0.059	0.589	0.224	0.179
AIC	−6.17	−3.46	−10.99	−5.58	3.71	−4.9

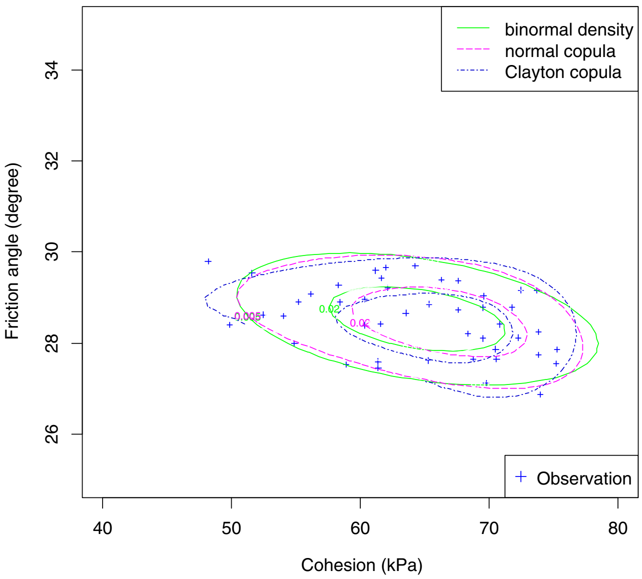

Fig. 5 Density contours of bivariate models for (c, φ) for soil BL-2 by [39] with (1) a bivariate normal density distribution, (2) a normal copula using the best-fitting margins and (3) a Clayton copula using the best-fitting margins

system. The method provides a complete approximation to observed data sets. The differences between the confidence regions of the observed data and the confidence regions of the simulated data are not pronounced, which suggest that the copula model can provide a good description of the given experimental data. Typically, these models could then be used in a Monte Carlo risk analysis. Because the CBSM is obviously prone to model risk, it should be seen as a form of sensitivity analysis. Varying dependencies can be chosen to represent the variability of the correlated soil strength properties and to assess the performance of the existing slope stability analysis program within a probabilistic framework, as illustrated below.

4 Probabilistic slope stability analyses

4.1 Deterministic analysis

Bishop's deterministic simplified method [7] is the most widely used limit equilibrium method and is based on the effective stress approach. The soil mass is divided into a number of vertical slices of equal width. The forces between the slices are neglected; each slice is considered to be an independent column of soil of unit thickness. Considering the entire slip surface, the factor of safety against sliding, F_s, is expressed by the resisting moments against the driving moments

$$F_s = \frac{1}{\sum W \sin \alpha} \sum \left(\left\{ c'b + \tan \phi' \left(W - ub \right) \right\} \frac{1}{m_\alpha} \right) \quad (11)$$

where $m_\alpha = \cos \alpha + \frac{\tan \phi' \sin \alpha}{F_s}$, c' and ϕ' are the effective shear strength parameters, W is the weight of a slice, u is the pore pressure and b is the width of the slice. Taking one slice as an example, the weight of slice W is calculated to be equal to $\gamma h_a b$, where γ is the bulk unit weight of the soil, h_a is the average height of the slice and b is its width. This is called Bishop's simplified method. Equation 11 includes the factor of safety F_s on both sides of the equation; therefore, the equation has to be solved by an iterative process. A trial value of F_s is first assumed, and the factor of safety is computed by iteration until the assumed and computed values of F_s coincide.

4.2 Reliability index determined by the correlated shear strength parameters using the CBSM

The reliability index is often used to express the degree of uncertainty in the calculated factor of safety for an input set of basic random variables. This type of reliability-based analysis provides quantification of the safety of a system by examining the variability of the relevant parameters as well as their interdependence. Well-established reliability methods, such as the FORM, the first-order second-moment (FOSM) method and Monte Carlo simulation, are useful in determining the reliability of geotechnical designs where the random variables are correlated [4, 5]. The FOSM approach provides a computationally efficient way of estimating the probability of failure [4], but the reliability index estimated using this approach is not 'invariant' and gives several expressions of the performance function [17, 44]. The FORM yields an invariant definition of the reliability index [46] by transforming basic input variables from the physical space to the standard normal space. To address correlated normal distributions, two techniques may be used with the FORM to pursue the expression for independent variables—one based on the Cholesky decomposition of the correlation matrix ([5], pp. 393–398) and the other based on orthogonal transformation by solving the eigenvectors of the covariance matrix ([4], pp. 353–359). For the non-normal correlated variables, a Rosenblatt transformation should be adopted. In this study, a transformation algorithm derived by [66] is used, and this algorithm does not require the user to leave the original space of the correlated variables. A brief description of the FORM is provided in Appendix 2.

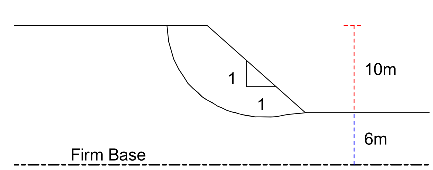

Fig. 6 Homogeneous slope

Table 5 Statistical properties of soil parameters for the homogeneous slope

Random variable	Unit	mean	Standard deviation	Distribution
c	kPa	18.0	3.6	Normal
tan ϕ		0.577	0.058	Normal

The applicability of the commonly used Monte Carlo simulation method for correlated variables to geotechnical problems has been described in detail in reference [46]. A number of algorithms have been developed in the literature to generate correlated random numbers [57]. Alternatively, the technique based on the copula sampling scheme imposed on the best-fitting marginal distributions and rank correlation matrices provides useful reconstructions of the joint behaviour of shear strengths, and the mean and standard deviation of the factor of safety can be obtained through these reconstructions by running the conventional definition of F_s repeatedly. Therefore, the reliability index β_{cb} determined by the CBSM can be calculated as follows [53]:

$$\beta_{\mathrm{cb}} = \frac{\mu_{F_s} - 1}{\sigma_{F_s}} \tag{12}$$

Finally, the failure of probability P_{f} can be estimated by the ratio of the running sum of the failed cases $(F_s < 1)$ m to the running sum of the total samples n_{sim}, i.e., $P_{\mathrm{f}} = {}^{m}\!/n_{\mathrm{sim}}$.

This leads to the following computational procedure:

1. Establish the number of realisations to be used, as discussed in Section 3.6;
2. For each point k, generate a paired value of $(c, \tan\varphi)$, with consideration of the dependence;
3. Calculate the factor of safety $F_s(c, \tan\varphi)$ and count the number to be added to a running sum m if $F_s(c, \tan\varphi) \le 1$;
4. After all points have been evaluated, evaluate the estimate of P_{f} from the running sum n_{sim}, i.e., $P_{\mathrm{f}} = m/n_{\mathrm{sim}}$, and calculate the reliability index β_{cb}.

When the number of simulations is sufficiently large, the standard deviation of the estimated values F_s can be obtained by simulating sample inverses with the square root of the simulating number. Thus, the accuracy increases as the number of simulations increases. In general, when the number of simulations is greater than $n_{\mathrm{sim}} \ge 100/P_{\mathrm{f}}$, the accuracy may be satisfactory [33, 60], and the probability of failure P_{f} can be calculated to represent a deterministic solution.

4.3 Illustrative numerical example

To illustrate the influence of the correlation on the reliability index β_{cb}, a series of analyses by the CBSM and the FORM are demonstrated in the following typical slope cases.

Example 1: application to a homogeneous slope

Slope stability analyses were performed using the simplified Bishop method, assuming circular slip surfaces. For instance, a homogeneous slope is shown in Fig. 6 and analysed by the proposed methods (the FORM and CBSM). The parameters considered as random variables

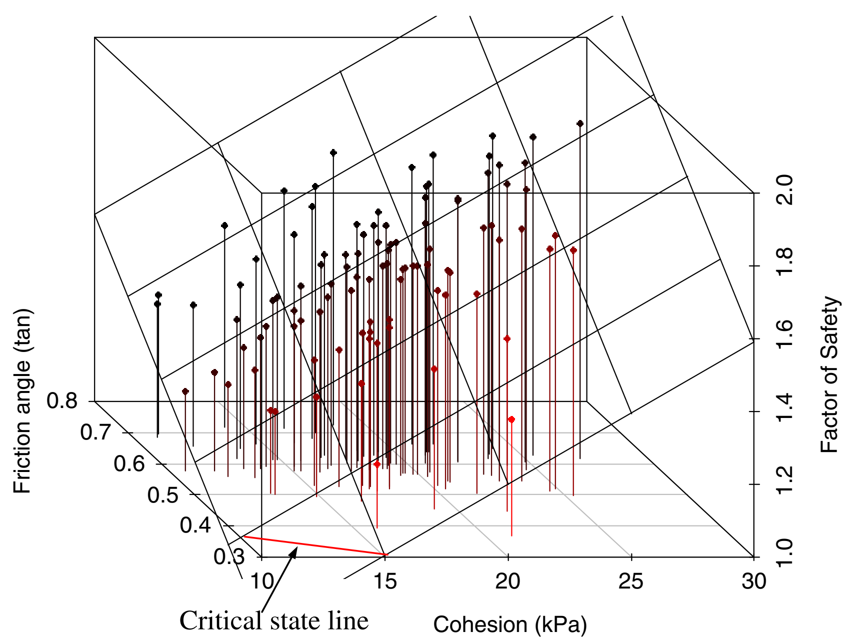

Fig. 7 Scatter plot of the factor of safety against cohesion and friction angle

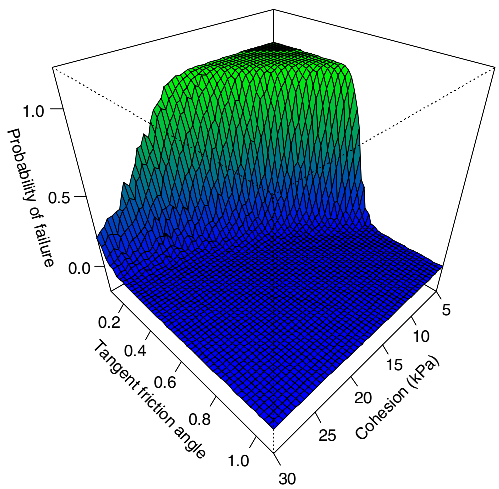

Fig. 8 Relation of probability of failure against cohesion and friction angle

for c, tanφ, as described previously by [38], are listed in Table 5. The mean value of the unit weight is assumed to be a constant 18.0 kN/m^3 (the same is true below, unless otherwise mentioned). The critical slip circle is shown in Fig. 6, according to a deterministic analysis based on the mean values of the soil parameters, similar to that reported by [30].

The τ of shear strengths is taken as -0.43, with a corresponding $\rho_p = -0.61$ [10]. Their marginals are listed in Table 2. Some of the sampled data from the CBSM (100 points) are shown in Fig. 7. The computed results for the factor of safety relative to the above 100 combined pairs are summarised in Fig. 7. The graph shows the factor of safety versus the cohesion and friction angle, using

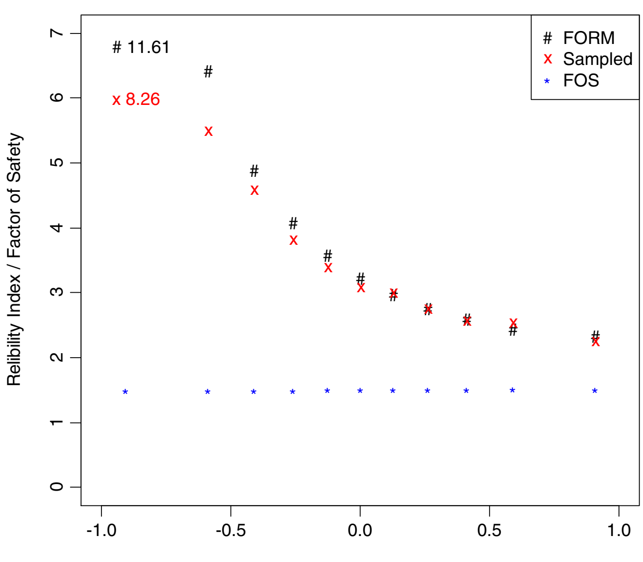

Fig. 9 Reliability index versus ranked correlations of soil properties for the homogeneous slope

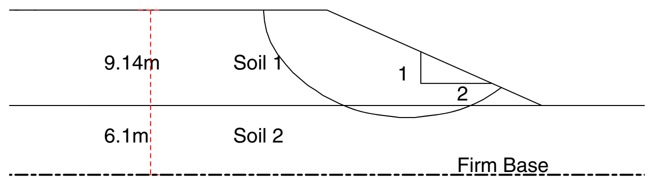

Fig. 10 Cross section of two-layer slope

a three-dimensional cube. A regression plane is added to the plot to support the visual impression. The factor of safety increases dramatically with the cohesion and friction angles, although it can be less than 1 for some small values of the cohesion and friction angle. A critical state line is defined as the projection line of the regression plane on the horizontal plane $F_s = 1$, as illustrated in Fig. 7. Pairs of (c, tanφ) values, i.e., (15, 0.3) and (10, 0.36), follow this line. If pairs of (c, tanφ) values are sampled with a correlation coefficient close to perfect ($\tau \approx 1$), these values will approximately follow a straight line on the c and tan φ plane (in this case, the variance of the shear strength is reduced in some degree); thus, the regression problem is reduced to a projection line rather than a plane.

The probability of failure can be calculated for specific combinations of the cohesion and friction angle. These combinations are obtained by sequentially setting one parameter with the remaining parameter set at its mean value. The standard deviation of cohesion is assumed to be 20 % of the mean, and the standard deviation of friction angle is assumed to be 10 % of the mean. The computed probability of failure is plotted against cohesion and friction angle in Fig. 8. The probability of failure increases considerably as the cohesion and friction angle decrease.

To demonstrate the influence of correlation extremes on reliability indices, the cross correlation τ is varied from -0.91 to 0.91 (such extreme values cannot be expected in reality), and the same statistics (including the means and standard deviations) for the cohesion and friction angle are fed into the FORM and CBSM

Table 6 Statistical properties of soil parameters for the stratified slope

Random variable	Unit	mean	Standard deviation	Distribution
c_1	kPa	38.31	7.662	Normal
tan ϕ_1		0	0	Normal
c_2	kPa	23.94	4.788	Normal
tan ϕ_2		0.209	0.021	Normal
c_1^*	kPa	38.31	7.662	Normal
tan ϕ_1^*		0.349	0.035	Normal

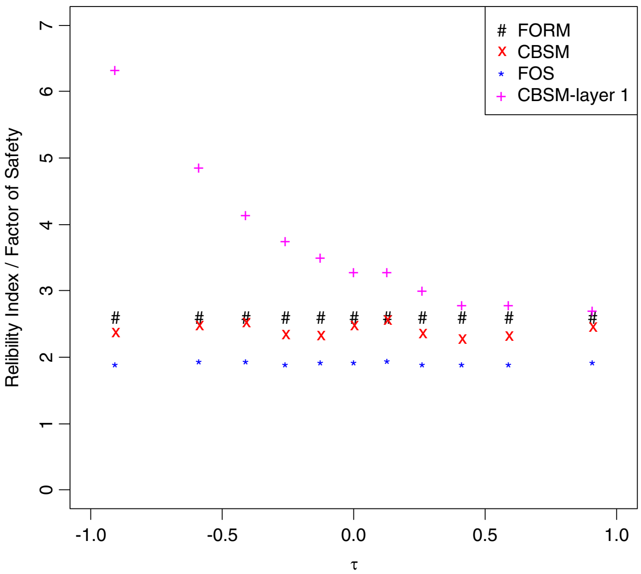

Fig. 11 Reliability index versus ranked correlations of soil properties for the stratified slope

(using the normal copula as an example, as described below). The reliability indices computed for these correlation coefficients are shown in Fig. 9. The reliability indices are expected to decrease as the correlation coefficients decrease. This observation arises from the fact that the variance of shear strength is reduced if there is a strong negative correlation between cohesion and friction angle. The results from the CBSM and the FORM show good agreement, which proves to work well in determining the reliability index using the computation technique presented.

In Fig. 9, the means of the factor of safety are also given by the deterministic limit equilibrium method through those sampled strength pairs (i.e., 10,000 simulations). The correlation coefficients have little influence on the means, which are always inferred from their aggregate behaviour in terms of the mean soil strengths. [38] determined the value of the reliability index using Hasofer and Lind's approximate method. According to their results, the minimum critical factor of safety in conventional design is estimated to be 1.5, and the corresponding reliability index is 2.63. The results obtained in this study show some agreement with their results, although the model input may be slightly different.

The FORM is a powerful tool in probabilistic geotechnical analysis, especially in standard normal spaces. However, partial differential terms of the performance function have to be derived if a gradient-based optimisation method is employed when searching for the shortest distance to the failure state. The CBSM better facilitates allowing the non-normal distribution and nonlinear failure state function. Other types of distributions for soil strength properties, such as the triangular, beta and generalised gamma distributions (suggested by [5, 39, 62]), can also be implemented in this approach. The CBSM is flexible in the sense that different distributions can be used to describe each marginal distribution while still being able to incorporate dependence, i.e., it allows the joint distribution function type to be different from the marginal cumulative distribution function types [49].

Example 2: application to a stratified slope

The cross section of a two-layer slope [30] is shown in Fig. 10. The slope in clay is bounded by a hard layer below and is parallel to the ground surface. The statistics of the soil strength parameters are summarised in Table 6. No water table or external water is considered. The corresponding critical deterministic slip surface, based on the mean values of the soil properties, is also presented in Fig. 9. A similar (circular) surface was reported by [30].

A parametric study was performed by specifying various correlation coefficients between the cohesion and the friction angle. The calculated reliability indices obtained from the FORM and CBSM for various correlation coefficient values are given in Fig. 11. Varying the cross correlation τ from -0.91 to 0.91 was found to have only a minor influence on the stochastic behaviour of the slope stability. This difference was not expected owing to the sliding circle being mostly through layer one (its cohesion is assumed to be zero). Notably, when a

Fig. 12 Cross section of the Cannon Dam

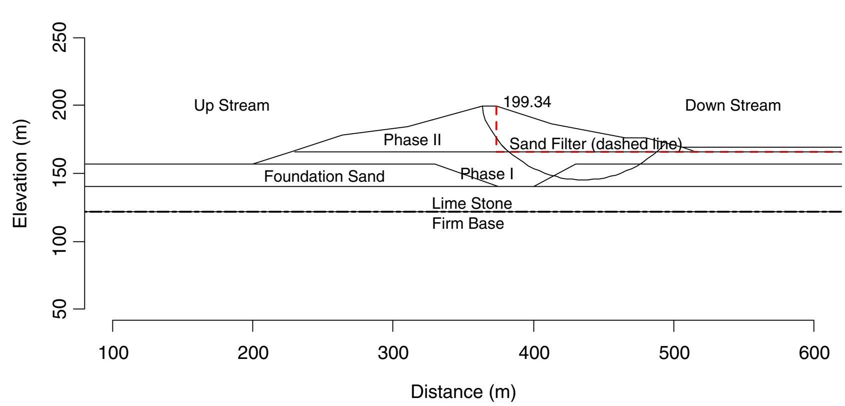

Table 7 Statistical properties of soil parameters for the Clarence Cannon Dam

Soil	Random variable	Unit	Mean	Standard deviation	Distribution
Phase I	c_1	kPa	117.79	58.89	Normal
	$\tan\phi_1$		0.15	0.15	Normal
Phase II	c_2	kPa	143.64	79	Normal
	$\tan\phi_2$		0.268	0.158	Normal

single soil strength parameter is used, consideration of the uncertainties will fall into a class of Monte Carlo sampling method. For instance, a granular material has little or no cohesion, and a clayey material has a very small or even zero friction angle. There is no difference between the CBSM and the conventional Monte Carlo sampling method because no explicit dependence should be represented.

When the cohesion and friction strength parameters of the first layer are set with the same means but with larger standard deviations, as listed in Table 6, c_1^* and $\tan\varphi_1^*$, the computed reliability indices by the CBSM are shown with symbols '+' in Fig. 11. The reliability indices decrease as the correlation increases.

Example 3: Application to the Clarence Cannon Dam

The third typical cross section of the Clarence Cannon Dam previously described by [63] is presented in Fig. 12. The structure consists of two zones of compacted clay, including phase I fill and phase II fill, over layers of sand and limestone. The strength parameters of

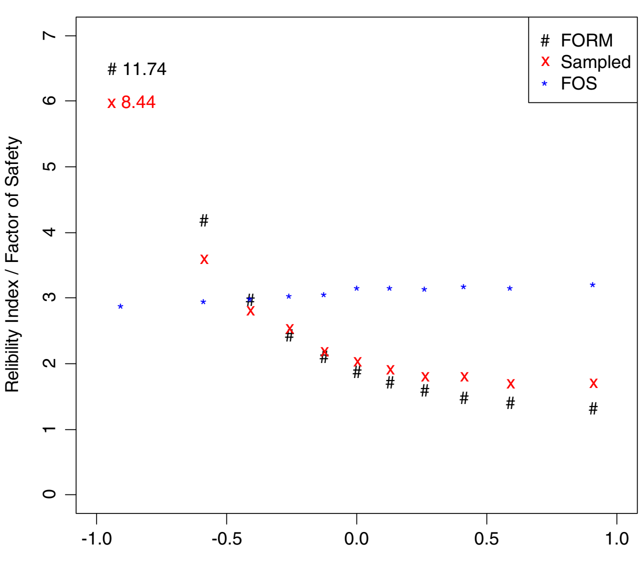

Fig. 13 Reliability index versus ranked correlations of soil properties for the Clarence Cannon Dam

the two clay layers are considered random variables. The statistics for these parameters, based on unconsolidated and undrained shear tests of samples from the embankment [62], are shown in Table 7. The critical deterministic circle is shown in Fig. 12.

A distribution with a high standard deviation, as used here for phase I and phase II clays, implies negative values associated with the low-probability tail of the distribution, which is not admissible for strength parameters. A similar truncated technique [19] is imposed to provide reasonable values.

Figure 13 shows the relations of the factor of safety and the reliability index to the correlation coefficients. The value of the factor of safety increases slightly as the correlation coefficients increase. The reliability index values based on both algorithms increase when the correlation coefficient between c and φ decreases from positive to negative. This is especially evident for the lowest values of the coefficient of variation for the cohesion and friction angle.

5 Discussion

In the probabilistic stability analysis described above, the location of the critical surface is part of the evaluation of the performance function and depends on the values of the strength parameters, which are uncertain. As noted by [30], the difference between the reliability index defined for the critical deterministic surface and the minimum reliability index may be substantial in some cases. Locating this critical probabilistic surface may require additional computational effort, and not doing so may lead to inaccurate measures of reliability. The technique suggested by Hassan and Wolff for locating the surface of the minimum reliability index is used in this study, which examines offset values of each of the random variables while keeping the remaining parameters at their mean values.

Clearly, variations in the correlation parameters of the strengths can substantially affect the reliability index, especially when the correlation approaches negative 1. As determined by [13], the reliability of a slope increases as the correlation between the cohesion and friction angle decreases. Thus, when the cohesion and friction angle are negatively correlated, the reliability index can be much higher than when the shear strength parameters are positively correlated. Therefore, neglecting any negative correlation underestimates the reliability index, while neglecting any positive correlation overestimates the reliability index.

These efforts in the bivariate statistical analysis of soil strength parameters are encouraging but insufficient to obtain an accurate description of the soil uncertainty state, which sometimes dominates multivariate problems. Some other parameters, such as the pore water pressure, unit weight, consolidation coefficient and seepage coefficient should also be considered.

Care should be taken, however, to ensure that the minimum and maximum values of the selected distribution are consistent with the physical limits of the parameter being modelled. For example, shear strength parameters should not imply negative values. If the selected distribution implies negative values in the third case, then the distribution is truncated at a practical minimum threshold. Alternatively, a best-fitting distribution, such as the lognormal, generalised gamma or Weibull distribution, can fit the observed data and avoid negative samples.

The identified copulas can be wrong if a very small number of samples are used. Although more fundamental experiments of shear tests of soils to provide enough data sets should be encouraged greatly, the sample size with around 50 can be acceptable [67]. The size of the data set has been mentioned by some researchers [4, 26] as affecting the confidence regions or dependence structures.

6 Conclusions

An approach to probabilistic slope stability analysis that accounts for the statistical correlation of the input soil strength parameters is presented. A set of reliability indices for varying correlation coefficients yield an objective description of the overall evaluation of slope stability and a better description of the degree of uncertainty. The applicability of the proposed methodology (including the CBSM and the FORM) described herein is examined for a variety of slope stability problems from the literature, such as a homogeneous slope, a stratified slope and the embankment of the Cannon hydroelectric project. The method is proven to be a practical and efficient method for facilitating a probabilistic slope stability analysis of cohesive frictional soils through copula-based samplings. The method does not rely on any assumptions concerning the geometry of the failure surface and can be applied to any complex clay slope geometry, layering and pore pressure conditions.

Invoking the CBSM to take into account the interdependence of soil strength properties, the new method has an advantage in implementation for inputting a combination of soil strength parameters. The approach is simple and can be applied in practice with little effort beyond that needed in a conventional analysis. The method permits practising engineers to locate the surface of the reliability index using existing deterministic slope stability computer programs, without special software, by making a moderate number of multiple runs.

The analysis of the results and the examination of the resulting plots illustrate the importance with respect to the reliability index of the correlation coefficient between soil strength properties, i.e., the reliability decreasing as the correlation increases. Comparing the computed results and the evaluated ones obtained using the FORM method, the CBSM tends to open the way for various marginal distribution types and dependence structure.

Appendix 1

Notation

Symbol	Description
A	indicator variable
b	width of slice
C	copula distribution function
c	cohesion
c'	effective cohesion
c_i	outward normal vector to a hyperplane from the geometry of surfaces, $= \lambda \frac{\partial g}{\partial z_i}$, where λ is arbitrary constant
d_{area}	relative percentage area difference
Cov	covariance of two random variables
F	marginal distribution
F^{-1}	quasi-inverse of F
F_s	factor of safety (defined with respect to shear strength)
g	performance function
H	two-dimensional distribution function
h_a	average height of slice
L_β	distance in standard deviation units
L_β^{corr}	L_β for correlated variables, $= \sqrt{\sum_{i=1}^{2} c_i^T R c_i}$
m	total number of failed cases
m_α	term used in the simplified Bishop method, $= \cos\alpha + \frac{\tan\varphi' \sin\alpha}{F_s}$
n_{sim}	number of simulations
p_c	Cramér–von Mises test statistic
p_m	Anderson–Darling test statistic
Pr	probability of failure
S^n	Cramér–von Mises function
u	pore water pressure
u_i	i^{th} uniform random variable, $=F(Z_i)$
u_j	j^{th} uniform random variable, $=F(Z_j)$
W	weight of slice, $=\gamma h_\alpha b$
Z_i	i^{th} random variable
Z_i^*	dimensionless variable, reduced variable, or standard random variable
z_i	i^{th} realisation of Z_i

Z_j	j^{th} random variable
z_j	j^{th} realisation of Z_j
α	inclination of the slope or the slope of failure surface
α_i	direction cosine, $= \frac{c_i}{L_\beta}$
α_i^{corr}	α_i for correlated variables, $= \frac{Rc_i}{L_\beta^{\text{corr}}}$
β_{HL}	the Hasofer–Lind reliability index
β_{cb}	reliability index by the CBSM
γ	bulk unit weight of soil
λ	degrees of freedom of the Student copula
μ	mean
ρ_{p}	Pearson's correlation coefficient
σ	standard deviation
Σ_2	variance–covariance matrix
τ	Kendall's correlation coefficient
φ	inner friction angle
φ'	effective inner friction angle
φ_θ	generator function
Φ	standard normal distribution
θ	copula parameter

Appendix 2

Calculation of the reliability index by the FORM with correlated variables

For the limit equilibrium analysis of slope stability, two shear strength variables Z_i, Z_j are considered, and a performance function can be written as $g(Z_i, Z_j) = F_s$–1. The FORM, commonly called the Hasofer–Lind method [29], transforms basic input variables from the physical space Z to the standard normal space Z^*, i.e., it uses dimensionless variables

$$Z_i^* = \{Z_i - \mu(Z_i)\}/\sigma[Z_i] \tag{13}$$

to explore the numerical approximation of the performance function. The reliability index β_{HL} defined by this method is measured by the distance L_β from the origin to the failure surface

$$g\left(Z_i^*, Z_j^*\right) = 0 \tag{14}$$

in the space of the dimensionless variables. The point on the failure surface (or curve) is called the 'design point'. This method was originally developed for normal-type or Gaussian-type variables. To extend its application to non-normal variables, the Rackwitz–Fiessler algorithm [53], is a straightforward local approximation of the marginal cumulative distribution function of a non-normal variable by a normal cumulative distribution function that has the same ordinate and slope at the design point.

Given that the limit state function is zero, the reliability index β_{HL} can be found from

$$\beta_{HL} = \min_{g\left(Z_i^*, Z_j^*\right)=0} \sqrt{\left\{Z_i^*\right\}^T \left\{Z_j^*\right\}} \tag{15}$$

Calculating this value is an iterative optimisation process in which the minimum value of a matrix calculation is found, subject to the constraint that the values result in a system failure. Creating an iterative scheme requires an expression for successive approximations Z_i^* and Z_j^*. This can be achieved from a first-order Taylor expansion of $g\left(Z_i^{*+1}, Z_j^{*+1}\right)$ about Z_i^* and Z_j^*

$$g\left(Z_i^{*+1}, Z_j^{*+1}\right) \approx g\left(Z_i^*, Z_j^*\right) + \left(Z_i^{*+1} - Z_i^*\right) \times \frac{\partial g\left(Z_i^*, Z_j^*\right)}{\partial Z_i^*} + \left(Z_j^{*+1} - Z_j^*\right) \times \frac{\partial g\left(Z_i^*, Z_j^*\right)}{\partial Z_j^*} \tag{16}$$

where Z_i^{*+1} represents the value of Z_i^* in the next iterative step. $\frac{\partial g\left(Z_i^*, Z_j^*\right)}{\partial Z_j^*}$ is the outward normal vector to a hyperplane (or curve) from the geometry of surfaces, denoted by c_i. Thus, the total length of the outward normal, L_β, is defined as $L_\beta = \sqrt{\sum_{i=1}^{2} c_i^2}$, which describes the distance between the most probable set of values and the most probable set of values that causes a failure. The direction cosines (also called sensitivity factors by [32]), α_i, of the unit outward normal are defined as $\alpha_i = \frac{c_i}{L_\beta}$, $i = 1{,}2$. There is an α_i value for each random variable considered in the reliability analysis, and the α values range from -1 to $+1$. With a known α_i, the coordinate of the trial point Z_i^* in the initial step can be estimated by

$$Z_i^* = -\alpha_i \beta_{HL} \tag{17}$$

Then, substituting Eq. 17 into Eq. 16 yields

$$Z_i^{*+1} = -\alpha_i \left[\beta_{HL} + \frac{g\left(Z_i^*, Z_j^*\right)}{L_\beta}\right] \tag{18}$$

If the basic variables are correlated, the literature ([59], [5] pp. 393–398; [4] pp. 353–359) recommends transforming the problem into a space of new variables that are uncorrelated and thereafter minimising β_{HL} in that space. Interestingly, Chowdhury and Xu [13] presented a transformation to address the correlated random variables in terms of original basic random variables, based on a linear algebra theory. Herein, a transformation algorithm derived by [66] that does not require the user to leave the original space of correlated variables is used. The technique is functional for the combined case of both a nonlinear failure surface and correlated variables in that it imposes the correlation matrix

R in existing formulas for independent variables. The correlation matrix is a function of the correlation coefficient ρ_{p}^{ij} for the pair Z_i, Z_j. For example, the length of the outward normal is $L_\beta^{\mathrm{corr}} = \sqrt{\sum_{i=1}^{2} c_i^T R c_i}$ and the direction cosines are $\alpha_i^{\mathrm{corr}} = \frac{Rc_i}{L_\beta^{\mathrm{corr}}}$. Readers can refer to [66] for more details on this numerical algorithm.

The scheme can now be summarised as follows:

1. Standardise the basic random variables Z to the standardised normal variables Z^*.
2. Compute the derivative $c_i = \frac{\partial g}{\partial Z_i^*}$ and the direction cosines $\alpha_i = \frac{c_i}{L_\beta}$ (for the independent case) or $\alpha_i^{\mathrm{corr}} = \frac{Rc_i}{L_\beta^{\mathrm{corr}}}$ (for the dependent case).
3. Evaluate $g\left(Z_i^*, Z_j^*\right)$.
4. Compute Z_i^{*+1} using Eq. 18 and β_{HL}^{+1} using Eq. 15.
5. Check whether β_{HL}^{+1} and Z_i^{*+1} have converged; if not go to step [2].

References

1. Alonso, E.E.: Risk analysis of slopes and its application to slopes in Canadian sensitive clays. Geotechnique **26**, 453–472 (1976)
2. Akaike, H.: A new look at the statistical model identification. IEEE Trans. Autom. Control. **AC-19**(6), 16–722 (1974)
3. Anderson, T.W., Darling, D.A.: A test of goodness-of-fit. J. Am. Stat. Assoc. **49**, 765–769 (1954)
4. Ang, A.H.-S., Tang, W.H.: Probability Concepts in Engineering Planning and Design, p. 562. Wiley, New York (1984)
5. Baecher, G.B., Christian, J.T.: Reliability and statistics in geotechnical engineering. Wiley, New York (2003)
6. Bergado, D.T., Anderson, L.R.: Stochastic analysis of pore pressure uncertainty for the probabilistic assessment of the safety of earth slopes. Soils and Found. **25**(2), 87–105 (1985)
7. Bishop, A.W.: The use of the slip circle stability in analysis of slopes. Geotechnique **1**, 7–17 (1955)
8. Boyer, B., Gibson, M., Loretan, M.: Pitfalls in tests for changes in correlation. International Finance Discussion Paper 597, Board of Governors of the Federal Reserve System (1999)
9. Brejda, J., Moorman, J., Smith, T.B., Karlen, J.L., Allan, D.L., Dao, T.H.: Distribution and variability of surface soil properties at a regional scale. Soil Sci. Soc. Am. J. **64**, 974–982 (2000)
10. Cherubini, C.: Data and considerations on the variability of geotechnical properties of soils. In: Proceedings of the International Conference on safety and reliability (ESREL 97) Lisbon, vol. 2, pp. 1583–1591 (1997)
11. Cho, S.E., Park, H.C.: Effect of spatial variability of cross-correlated soil properties on bearing capacity of strip footing. Int. J. Numer. Anal. Meth. Geomech. **34**, 1–26 (2010)
12. Chowdhury, R.N., Tang, W.H., Sidi, I.: Reliability model of progressive slope failure. Geotechnique **37**(4), 467–481 (1987)
13. Chowdhury, R.N., Xu, D.W.: Reliability index for slope stability assessment-two methods compared. Reliab. Eng. Syst. Safe. **37**(2), 99–108 (1992)
14. Chowdhury, R.N., Xu, D.W.: Rational polynomial technique in slope stability analysis. J. Geotech. Eng. Div. ASCE **119**(12), 1910–28 (1993)
15. Christian, J.T., Ladd, C.C., Baecher, G.B.: Reliability applied to slope stability analysis. J. Geotech. Eng. ASCE **120**(12), 2180–2207 (1994)
16. Clemen Robert, T., Reilly, T.: Correlations and copulas for decision and risk analysis. Manag. Sci. **45**(2), 208–224 (1999)
17. Duncan, J.M.: Factors of safety and reliability in geotechnical engineering. J. Geotech. Geoenviron. ASCE **126**, 307–316 (2000)
18. Dupuis, D.J.: Using Copulas in hydrology: benefits, cautions and issues. J. Hydrol. Eng. ASCE **12**(4), 381–393 (2007)
19. El-Ramly, H., Morgenstern, N.R., Cruden, D.M.: Probabilistic slope stability analysis for practice. Can. Geotech. J. **39**, 665–683 (2002)
20. Embrechts, P., McNeil, A.J., Straumann, D.: Correlation and dependence in risk management: properties and pitfalls. In: Dempster, M. (ed.) Risk Management: Value at Risk and Beyond, pp. 176–223. Cambridge University Press, Cambridge (2002)
21. Fenton, G.A., Griffiths, D.V.: Bearing capacity prediction of spatially random $c-\phi$ soils. Can. Geotech. J. **40**(1), 54–65 (2003)
22. Ferson, S., Hajagos Janos, G.: Varying correlation coefficients can underestimate uncertainty in probabilistic models. Reliab. Eng. Syst. Saf. **91**, 1461–1467 (2006)
23. Forrest William, S., Orr Trevor, L.L.: Reliability of shallow foundations designed to Eurocode 7, Vol. 4, pp. 186-207 (2010)
24. Frees, E.W., Valdez, E.A.: Understanding relationships using copulas. North Amer. Actua. J. **2**(1), 1–25 (1998)
25. Genest, C., MacKay, J.: The joy of copulas: bivariate distributions with uniform marginals. Am. Stat. **40**, 280–283 (1986)
26. Genest, C., Favre, A.C.: Everything you always wanted to know about copula modelling but were afraid to ask. J. Hydrol. Eng. ASCE **12**(4), 347–368 (2007)
27. Genest, C., Remillard, B., Beaudoin, D.: Goodness-of-fit tests for copulas: a review and a power study. Insur. Math. Econ. **44**, 199–214 (2009)
28. Harr, M.E.: Reliability based design in Civil Engineering. McGraw-Hill, New York (1987)
29. Hasofer, A.A., Lind, A.M.: Exact and invariant second moment code format. J. Engrg. Mech. Div. ASCE **100**(1), 111–121 (1974)
30. Hassan, A., Wolff, T.: Search algorithm for minimum reliability index of earth slopes. J. Geotech. Geoenviron. Eng. ASCE **125**, 301–308 (1999)
31. Hata, Y., Ichii, K., Tokida, K.-i.: A probabilistic evaluation of the size of earthquake induced slope failure for an embankment. In: Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards (2011). doi:10.1080/17499518.2011.604583
32. Honjo, Y., Suzuki, M., Matsuo, M.: Reliability analysis of shallow foundations in reference to design codes development. Comput. Geotech. **26**, 331–346 (2000)
33. Husein Malkawi, A.I., Hassan, W.F., Abdulla, F.: Uncertainty and reliability analysis applied to slope stability. Struct. Saf. J. **22**, 161–187 (2000)
34. Joe, H.: Multivariate models and dependence concept. Chapman and Hall, New York (1997)
35. Kojadinovic, I., Yan, J.: Modeling multivariate distributions with continuous margins using the copula R Package. J. Stat. Soft. **34**(9), 1–20 (2010)
36. Kotz, S., Balakrishnan, N., Johnson, N.: Continuous multivariate distributions. Wiley, New York (2000)
37. Lambert, P., Vandenhende, F.: A copula-based model for multivariate nonnormal longitudinal data: analysis of a dose titration

safety study on a new antidepressant. Stat. Med. **21**, 3197–3217 (2002)
38. Li, K.S., Lumb, P.: Probabilistic design of slopes. Can. Geotech. J. **24**, 520–535 (1987)
39. Lumb, P.: Safety factors and the probability distribution of soil strength. Can. Geotech. J. **7**, 225–242 (1970)
40. Lü, Q., Low, B.K.: Probabilistic analysis of underground rock excavations using response surface method and SORM. Comput. Geotech. **38**, 1008–1021 (2011)
41. Marshall, A.W., Olkin, I.: Families of multivariate distributions. J. Am. Stat. Assoc. **83**, 834–841 (1988)
42. McNeil, A.J., Frey, R., Embrechts, P.: Quantitative risk management: concepts, techniques and tools. Princeton University Press, Princeton (2005)
43. Montgomery, D.C., Runger, G.C.: Applied Statistics and Probability for Engineers. Wiley, New York (1999)
44. Nadim, F.: Tools and strategies for dealing with uncertainty in geotechnics. In: Griffiths, D.V., Fenton, G.A. (eds.) Probabilistic methods in geotechnical engineering, pp. 71–96. Springer, Wien (2007)
45. Nelsen, R.B.: An introduction to copulas, 2nd edition. Springer, New York (2006)
46. Nguyen, V.U., Chowdhury, R.N.: Probabilistic study of spoil pile stability in strip coal mines-two techniques compared. Rock. Mech. Min. Sci. Geomech. Abstr., 303–212 (1984)
47. Phoon, K.K., Nadim, F.: Modelling non-Gaussian random vectors for FORM: state-of-the-art review. International workshop on risk assessment in site characterization and geotechnical design. India Institute of Science, Bangalore, India, pp. 26–27 (2004)
48. Phoon, K.K., Kulhanny, F.H.: Evaluation of geotechnical property variability. Can. Geotech. J. **36**, 625–639 (1999)
49. Poulin, A., Huard, D., Favre, A., Pugin, S.: Importance of tail dependence in bivariate frequency analysis. J. Hydrol. Eng. ASCE **12**(4), 394–403 (2007)
50. Pouillot, R., Delignette-Muller, M.L.: Evaluating variability and uncertainty separately in microbial quantitative risk assessment using two R packages. Int. J. Food Microbiol. **142**, 330–340 (2010)
51. R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN:3-900051-07-0. http://www.R-project.org (2008)
52. Rackwitz, R.: Reviewing probabilistic soils modeling. Comput. Geotech. **26**(3), 199–223 (2000)
53. Rackwitz, R., Fiessler, B.: An algorithm for calculation of structural reliability under combined loading. P Berichte zur Sicherheitstheorie der Bauwerke, Lab. f. Konstr Ingb. Munchen, Germany (1977)
54. Salvadori, G., De Michele, C., Kottegoda, N.T., Rosso, R.: Extremes in Nature. Springer, Dordecht (2007)
55. Schweizer, B.: Thirty years of copulas. In: Dall'Aglio, G., Kotz, S., Salinetti, G. (eds.) Advances in probability distributions with given marginals. Kluwer, Dordrecht (1991)
56. Sklar, A.: Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique Université de Paris **8**, 229–231 (1959)
57. Tamimi, S., Amadei, B., Frangopol, D.M.: Monte Carlo simulation of rock slope stability. Comput. Struct. **33**(6), 1495–1505 (1989)
58. Tang, W.H., Yucemen, M.S., Ang, A.H.-S.: Probability-based short term design of slopes. Can. Geotech. J. **13**(3), 201–215 (1976)
59. Thoft-Christensen, P., Baker, M.J.: Structural reliability theory and its applications. Springer, New York (1982)
60. Tobutt, D.C.: Monte Carlo simulation methods for slope stability. Comput. Geosci. **8**(2), 199–208 (1982)
61. Venables, W.N., Ripley, B.D.: Modern applied statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0 (2002)
62. Wolff, T.F.: Analysis and design of embankment dam slopes: a probabilistic approach, Ph.D. Thesis. Purdue University, Lafayette, Indiana (1985)
63. Wolff, T.F., Hassan, A., Khan, R., Ur-Rasul, I., Miller, M.: Geotechnical reliability of dam and levee embankments. Technical report prepared for U.S. Army Engineer Waterways Experiment Station. Geotechnical Laboratory, Vicksburg, MS (1995)
64. Yan, J.: Enjoy the joy of copulas: with a package copula. J. Stat. Softw. **21**(4), 1–21 (2007)
65. Yan, J., Kojadinovic, I.: Copula: multivariate dependence with copulas. R package version 0.9-5. http://CRAN.R-project.org/package=copula (2010)
66. Zahn, J.J.: Empirical failure criteria with correlated resistance variables. J. Struct. Eng. **116**(11), 3122–3137 (1989)
67. Zhang, L., Singh, V.P.: Trivariate flood frequency analysis using the Gumbel-Hougaard copula. J. Hydrol. Eng. **12**, 431–439 (2007)