Improved knowledge-based clustered partitioning approach and its application to slope reliability analysis

Xiao-Song Tang\textsuperscript{a}, Dian-Qing Li\textsuperscript{a,∗}, Yi-Feng Chen\textsuperscript{a}, Chuang-Bing Zhou\textsuperscript{a}, Li-Min Zhang\textsuperscript{b}

\textsuperscript{a}State Key Laboratory of Water Resources and Hydropower Engineering Science, Key Laboratory of Rock Mechanics in Hydraulic Structural Engineering (Ministry of Education), Wuhan University, 8 Donghu South Road, Wuhan 430072, PR China
\textsuperscript{b}Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

**A R T I C L E   I N F O**

Article history:
Received 31 January 2012
Received in revised form 6 May 2012
Accepted 8 May 2012
Available online 31 May 2012

Keywords:
Slope stability
Rock slopes
Reliability analysis
Knowledge-based clustered partitioning
Monte Carlo simulation
Optimization

**A B S T R A C T**

A knowledge-based clustered partitioning (KCP) approach is improved to determine the reliability index and probability of failure of a rock slope. The Nataf transformation is adopted to transform the correlated non-normal random variables involved in the KCP approach into independent standard normal variables. An improved KCP technique is proposed to search the design point and calculate the reliability index. Two illustrative examples are presented to demonstrate the capability and validity of the proposed approach. The results indicate that the improved KCP-based reliability method can be applied to evaluate the reliability of rock slopes involving multiple correlated non-normal variables accurately and efficiently. Its accuracy is shown to be higher than that of the traditional KCP using the bisection method, and it is much more efficient than Monte Carlo simulation. The improved KCP-based reliability method is especially suitable for dealing with an implicit performance function with a large number of random variables, which is often involved in slope reliability analysis.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Slope stability analysis often involves many uncertainties. Reliability-based methods should be used for slope stability analysis because the probability of failure or reliability index provides a more consistent measure of safety of slopes (e.g., \cite{[2,33,17,19]}, several methods are available to determine the slope reliability index. These methods include response surface method (RSM) (e.g. \cite{[38]}), stochastic response surface method (e.g. \cite{[17]}), polynomial technique \cite{[8]}, First-order reliability method (FORM) implemented with an Excel Solver \cite{[23,25]}, Second-order reliability method (SORM) \cite{[28]}, point estimation method \cite{[7]}, Monte Carlo simulation (MCS) (e.g. \cite{[37]}), and random finite element method (RFEM) \cite{[14]}. As pointed out by Lee et al. \cite{[15]}, however, these methods may not be well suited for reliability problems that involve a complex performance function and a large number of random variables. In this study, a knowledge-based clustered partitioning (KCP) technique (e.g., \cite{[30,34]}) is modified and used to determine the reliability index.

The KCP technique was initially used for grouping books in a library \cite{[30]}. Thereafter, Shi et al. \cite{[34]} used the KCP technique to solve a traveling salesman problem that involves finding the shortest route among a number of cities. Recently, the KCP technique was applied to geotechnical problems. For example, Lee et al. \cite{[15]} employed the KCP technique to estimate the annual liquefaction probability. It was further used to determine the reliability index and failure probability of rock wedge stability \cite{[16]} and the reliability results using the KCP technique were compared with those from an Excel Solver-based method \cite{[23]}.

The studies by Lee et al. \cite{[15,16]} show that the KCP approach is an effective algorithm for searching the design point and reliability index. However, the KCP approach used by Lee et al. \cite{[15,16]} cannot deal with problems involving correlated non-normal random variables. It is widely accepted that the geotechnical engineering literature is replete with correlations between two engineering parameters. For instance, negative correlation between cohesion and friction angle \cite{[2]}, curve-fitting parameters of soil–water characteristic curves \cite{[33]}, and two hyperbolic parameters underlying a pile load–displacement curve \cite{[19]}. Furthermore, these parameters usually follow non-normal distributions \cite{[2,33,17,19]}. Therefore, it is necessary to extend the KCP approach to handle correlated non-normal random variables involved in reliability analysis. In addition, a bisection method underlying the KCP approach was adopted to search the design point and reliability index \cite{[15,16]}. Its accuracy may not be sufficient for some special slope reliability problems.

The objective of this study is to propose an improved knowledge-based clustered partitioning (KCP) approach to search the design point and compute the reliability index. The Nataf transformation...
[31] is adopted for the transformation of correlated non-normal random variables. A changing-increment method is proposed to replace the bisection method underlying the KCP approach. This article is organized as follows. In Section 2, the improved KCP approach is introduced in detail. Two practical reliability analyses of slope stability are demonstrated in Section 3. The first example is about the reliability of a rock slope that involves a closed-form performance function. This example is used to validate the KCP approach by comparison with results obtained from FORM and direct MCS. The second example is analyzed to illustrate the capability of the proposed KCP approach in dealing with an implicit performance function, which is compared with the FORM with response surface method and direct MCS.

2. Improved knowledge-based clustered partitioning (KCP) approach

2.1. Hasofer and Lind’s reliability index

The matrix formulation of the Hasofer–Lind second moment reliability index, \( \beta_{HL} \), is expressed as [24]

\[
\beta_{HL} = \min_{x \in \Omega} \left[ \frac{x_i - \mu_i}{\sigma_i} \right]^T [R]^{-1} \left[ \frac{x_i - \mu_i}{\sigma_i} \right]
\]  

where \( X \) is a vector representing the input random variables \( x_i; \mu_i \) and \( \sigma_i \) are the mean and standard deviation of \( x_i \), respectively; \( R \) is the correlation matrix of \( X \); \( \Theta \) is the failure region.

The reliability index is regarded as the minimum distance from the mean of the variables to the boundary of the failure region in the vector direction of directional standard deviation [12]. Essentially, the solution of \( \beta_{HL} \) is an optimization problem subjected to constraints. In this paper, an improved KCP approach is proposed to solve the optimization problem for determining the reliability index.

2.2. Knowledge-based clustered partitioning (KCP) approach

As discussed by Lee et al. [15,16], the KCP approach is a global optimization technique, which is well suitable for solving a large number of cases at once. In this aspect, it is very different from the case-by-case solution using the Excel Solver as proposed by Low and Tang [23]. The KCP approach is very efficient although it is conceptually more complex than other optimization approaches. It adopts a systematic algorithm for locating the promising regions and shortens the iteration time in the optimization process. Let \( K_i \) and \( X_i \) be a standard normal random variable and a non-normal random variable, respectively. Consider the following isoprobabilistic transformation (e.g., [1]),

\[
\begin{align*}
\Phi(k_i) &= F_{X_i}(x_i) \\
\xi_i &= F^{-1}_X(\Phi(k_i))
\end{align*}
\]

in which \( F_{X_i}(\cdot) \) is the cumulative distribution function (CDF) of \( X_i; F^{-1}_X(\cdot) \) is the inverse CDF of \( X_i \); \( \Phi(\cdot) \) is the CDF of a standard normal random variable. Applying Eq. (2), each random variable \( X_i \) in Eq. (1) can be transformed to a standard normal random variable \( K_i \). Taking the lognormal random variable as an example, the corresponding transformation is derived as

\[
X_i = \exp \left( \ln \mu_i - \frac{1}{2} \ln \left( 1 + \sigma_i^2/\mu_i^2 \right) \right) + K_i \sqrt{\ln \left( 1 + \sigma_i^2/\mu_i^2 \right)},
\]

where \( n \) is the number of random variables; then, a non-normal data space \( X \) can be converted to a standard normal data space \( K \). In the transformed space, the design point can be expressed in terms of the random variables \( K_i \) by Eq. (3). If the random variables are correlated non-normal random variables, the Nataf transformation [31] can be employed to transform the correlated non-normal random variables into independent standard normal random variables, as shown in Appendix A.

Consider a reliability problem involving 10 random variables. Note that for the reliability analysis of the Jinping rock slope to be presented later, the implicit performance function involves 10 random variables. The design point in the transformed space can be represented by 10 numbers, \( K_i (i = 1, 2, \ldots, 10) \). Determining the reliability index and the design point requires solving the optimization problem involving 10 random variables. Such an optimization process is time-consuming with traditional optimization algorithms such as the Excel Solver. To improve the efficiency, the 10-dimensional space is transformed into a polar coordinate system with one length \( L \) and four angles \( \theta_1, \theta_2, \theta_3 \) and \( \theta_4 \) using the KCP technique. The corresponding expressions are derived as follows:

\[
\begin{align*}
\theta_1 &= L \times \sin(\theta_1) \times \sin(\theta_2) \times \sin(\theta_3) \times \sin(\theta_4), \\
\theta_2 &= L \times \sin(\theta_1) \times \sin(\theta_2) \times \sin(\theta_3) \times \cos(\theta_4), \\
\theta_3 &= L \times \sin(\theta_1) \times \sin(\theta_2) \times \cos(\theta_3) \times \sin(\theta_4), \\
\theta_4 &= L \times \sin(\theta_1) \times \sin(\theta_2) \times \cos(\theta_3) \times \cos(\theta_4), \\
\end{align*}
\]

where \( \theta_i (i = 1, 2, \ldots, 10) \) are possible values of \( K_i \) in the independent standard normal space; the angles (\( \theta_1, \theta_2, \theta_3 \) and \( \theta_4 \)) and the length \( L \) are unknown parameters to be determined.

Similarly, expressions for any number of random variables can be derived. For convenience, the expressions of \( \theta_i \) for 2–8 random variables are presented in Appendix B. It should be pointed out that the number of angles is related to the number of input random variables. If \( d \) angles are used, the maximum number of random variables involved in a reliability problem can be up to \( 2^d \). For instance, three angles will suffice to deal with a reliability problem involving eight random variables (\( 2^3 = 8 \)). If nine variables are involved, then four angles should be used. One can simply increase the number of “angle” variables so that the “dimension” of \( u \), or the number of “addressable” variables, can be increased.

It can be proved that the following relationship holds [15,16]:

\[
\sqrt{\sum_{i=1}^{n} (u_i)^2} = \sqrt{\sum_{i=1}^{n} (K_i)^2} = L
\]  

Thus, \( L \) is the distance between any given point on the limit state surface and the origin. If all the “points” on the limit state surface can be found and analyzed, the reliability index, namely the smallest \( L \), can be determined. Applying the polar coordinate system, the optimization problem to find the design point and reliability index involving \( n \) random variables in the original space is reduced to \((1+d)\) random variables. Regarding the problem shown in Eq. (4), the optimization problem involving 10 random variables now becomes one that involves five random variables. In this way, the computational efficiency is improved greatly.

In order to convert into the polar coordinate system, the terms \( K_i \) in Eq. (3) should be replaced by any \( u_i \) but each \( u_i \) can only be selected once for a given \( K_i \). Accordingly, the first variable \( K_1 \) can take its value from ten possibilities, \( u_1, u_2, \ldots, u_{10} \). For the second through 10th variables (\( K_2 \) through \( K_{10} \)), the number of possibilities reduces to nine, eight, seven, six, five, four, three, two and one, respectively. Therefore, there is a total of 3,628,800 (i.e., \( 10! \)) combinations. Accordingly, the analysis needs to be repeated 3,628,800 times. It
is a very time-consuming process. For this reason, the KCP technique [34,15] is employed to reduce the total number of combinations. In general, it includes the following five steps: (1) partitioning, (2) random sampling, (3) calculation of length L, (4) backtracking and (5) determination of the reliability index and design point. In the following, these five steps except the third step are briefly introduced, and a detailed description of the third step is given because a new algorithm is proposed to search the reliability index. It should be noted that the aforementioned five steps were initially proposed by Lee et al. [15,16], and are used in this study only for clearly explaining the improved KCP approach.

**Step 1 – Partitioning**

It can be observed from Eq. (4) that terms $u_1$ and $u_2$ are conjugated because they are the term $L \times \sin(\theta_1) \times \sin(\theta_2)$ multiplied by $\sin(\theta_4)$ and $\cos(\theta_4)$, respectively, and the square root of the sum of $u_1^2$ and $u_2^2$ is equal to $L \times \sin(\theta_1) \times \sin(\theta_2)$, Similarly, terms $u_3$ and $u_4$ are conjugated. Additionally, the square root of the sum of $u_1^2$ and $u_2^2$ and that of $u_3^2$ and $u_4^2$ are also conjugated. For the sake of computational efficiency, $u_1, u_2, u_3$ and $u_4$ may be clustered in the same partition because of the similarity of the function. Applying the similar approach, $u_5, u_6, u_7$ and $u_8$ may be clustered in the same partition, and $u_9$ and $u_{10}$ may be clustered in another partition. In this way, $u_i$ $(i = 1, 2, \ldots, 10)$ may be divided into three partitions. According to Lee et al. [15], within a partition, the sequence of $u_i$ is not important. For instance, $[u_1, u_2]$ and $[u_3, u_4]$ are considered the same.

**Step 2 – Random sampling**

Having obtained the partitions, the number of possibilities in assigning $u_i$ to a given $K_i$ can be reduced substantially. For example, if $u_i$ has been assigned to $K_i$, then $u_i$ will not be assigned to $K_2, K_3$ and $K_4$. By substituting $K_i$ in Eq. (3) with $u_i$, the 10 variables $K_i$ $(i = 1, 2, \ldots, 10)$ may also be separated into three partitions. As shown in Fig. 1, the entire sampling process is carried out in three stages. In the first stage, the sampling starts with $K_1, K_2, K_3$ and $K_4$, the corresponding $u_i$ are selected according to the partitioning principle discussed in Step 1. There are 210 possible combinations ($C_{10}^2 = 210$) to assign $K_1, K_2, K_3$ and $K_4$. In the second stage, sampling for $K_5, K_6, K_7$ and $K_8$ is conducted and the corresponding $u_i$ can be determined. There are 15 possible combinations ($C_{15}^2 = 15$). In the third stage, sampling for $K_9$ and $K_{10}$ is carried out and there is only one alternative. The number of all possible combinations is $210 \times 15 \times 1 = 3150$. This number is about 0.1% of 3,680,800 corresponding to the total number of combinations without partitioning. Such result indicates that the KCP approach can significantly save the computational time as the number of combinations to search is greatly reduced.

**Step 3 – Calculation of length L**

An improved searching algorithm is shown in Fig. 2. This algorithm involves 3 levels of search for length L and 2 levels of search for the angles $\theta_1, \theta_2, \theta_3$ and $\theta_4$. Note that the searching range of L is set as $[-10,10]$ because the probability of failure will approach one or zero if the reliability index is smaller than $-10$ or greater than $10$. Thus, the following two cases are taken into consideration:

1. When the central factor of safety (CFS) calculated by Lee et al. [15,16] is less than 0, then the reliability index is positive, the searching range of L is set as $[0,10]$. The searching ranges of angles $\theta_1, \theta_2, \theta_3$ and $\theta_4$ are $[0^\circ, 360^\circ]$ and the starting values of angles $\theta_1, \theta_2, \theta_3$ and $\theta_4$ are all set to 0. For each set of angles ($\theta_1, \theta_2, \theta_3$ and $\theta_4$), $L_{min}$ is determined by the bisection method based on the calculated FS, as shown in Fig. 2. This process is repeated until $|a-b| < \varepsilon = 0.0001$. In the end, $L_{min}$ is determined as $(a+b)/2$ for a given set of angles.

2. When the CFS is below 1 and the resulting reliability index is negative, the search range of L is set as $[-10,0]$ and the starting value of L is set to $-\sqrt{n}$. The searching ranges of angles $\theta_1, \theta_2, \theta_3$ and $\theta_4$ are $[0^\circ, 360^\circ]$ and the starting values of angles $\theta_1, \theta_2, \theta_3$ and $\theta_4$ are all set to 0. For each set of angles ($\theta_1, \theta_2, \theta_3$ and $\theta_4$), $L_{min}$ is determined by the bisection method on the calculated FS, as shown in Fig. 2. This process is repeated until $|a-b| < \varepsilon = 0.0001$. In the end, $L_{min}$ is determined as $(a+b)/2$ for a given set of angles.

In order to search the reliability index and the optimal angles, a bisection method was used by Lee et al. [15,16]. In their study, the initial increment for all three angles is set to 45°. After each alteration of angle, the increment is set to be one half of the previous angle increment. This process is continued until the angle increment is less than 1° in all three angles. Then, the initial increment in all three angles is set to 45° again, and the process described above is repeated until all the three angles reach their upper bounds. However, the bisection method may not search the angles evenly and may result in insufficient accuracy. To overcome this disadvantage, a changing-increment method is proposed in this study. This method includes the following three steps:

1. A large increment of angle is set to be $\Delta \theta_1$, then the tentatively optimal $L_{min}$ and $\theta_{opt1}$ are determined by the search algorithm. The searching ranges of all four angles are $[0^\circ, 360^\circ]$ in this step.

2. An increment $\Delta \theta_2$ smaller than $\Delta \theta_1$ and searching ranges of $[\theta_{opt1} - \Delta \theta_2, \theta_{opt1} + \Delta \theta_2]$ for the four angles are used to determine $L_{min2}$ and $\theta_{opt2}$, which are better than $L_{min1}$ and $\theta_{opt1}$, respectively.

3. An increment $\Delta \theta_3$ smaller than $\Delta \theta_2$ and searching ranges of $[\theta_{opt2} - \Delta \theta_3, \theta_{opt2} + \Delta \theta_3]$ are used. Such a process is repeated until the angle increment is less than 1° for all four angles. Finally, the minimum value of all the $L_{min}$ values obtained from all possible sets of angles is the optimal solution of the minimum length for a given $K_i$ combination. And the entire process is repeated for all possible $K_i$ combinations. It is evident that the angle increment decreases gradually from the first step to the third step. However, the angle increment remains the same within each step.

To improve the computational efficiency, the angle increment $\Delta \theta$ is changed two times and the minimum increment is set to 1° herein. That is, $\Delta \theta_1 = 45^\circ, \Delta \theta_2 = 5^\circ$ and $\Delta \theta_3 = 1^\circ$ are adopted for the subsequent analyses. Note that it takes 8 increments for angles $\theta_i (i = 1, 2, 3, 4)$ to increase from 0° to 360° in the first step; 18 increments from $\theta_{opt1} - 45^\circ$ to $\theta_{opt1} + 45^\circ$ in the second step; and 10 increments from $\theta_{opt2} - 5^\circ$ to $\theta_{opt2} + 5^\circ$ in the third step. Thus the total number of increments is 36, which is slightly larger than 28 for the bisection method used by Lee et al. [15,16]. However, the design point is determined more accurately than that using the
The bisection method because a small angle increment of $1^\circ$ is used for searching the design point. The result will still converge to a unique value when different starting values of $\Delta \theta_i$ are used.

**Step 4 – Backtracking**

During the searching process for $L_{\text{min}}$, if $x_i$ does not fall in the preset range in a certain sampling, angles $\theta_1$, $\theta_2$, $\theta_3$ and $\theta_4$ are changed in turn until a set of $\theta_1$, $\theta_2$, $\theta_3$ and $\theta_4$ is produced that yields a satisfactory $x_i$. Another backtracking point involves a case when the stopping criteria ($L > 10 \text{ or } L < -10$, $\theta_i > 360^\circ$, or $\theta_i > \theta_{\text{opt}} + \Delta \theta_1$, $\theta_i < \theta_{\text{opt}} - \Delta \theta_2$) are reached. In that case, a new combination of $u_i$ has to be created based on a new partition, and all the previous steps need to be repeated.

**Step 5 – Determination of reliability index and design point**

The smallest value of the minimum $L$ values obtained for all possible combinations of $u_i$ is taken as reliability index $b_{\text{HL}}$. Then substituting the corresponding $K_i$ into Eq. (3) yields the design point of all variables $X_i$. The probability of failure, $p_f$, is calculated by $p_f = 1 - \phi(b)$.

### 3. Illustrative examples

#### 3.1. Example 1 – Reliability analysis of Sau Mau Ping slope in Hong Kong with a closed-form performance function

For the purpose of illustration, the Sau Mau Ping slope in Hong Kong [25] is used. The geometry of the slope is illustrated in Fig. 3. Following Low [25], deterministic parameters are adopted for the analyses as below. The overall slope height is $H = 60$ m. The angles of the slope face and the potential failure surface are $\psi_p = 35^\circ$, respectively. The specific weights of rock and water are $\gamma = 26$ kN/m$^3$ and $\gamma_w = 10$ kN/m$^3$, respectively. The reinforcing force is $T = 0$ and the inclination of the reinforcing force is $\theta = 0^\circ$. The statistical parameters for the input random variables in the rock slope stability model are listed in Table 1. Detailed information of the statistical parameters can be referred to Low [25] and Li et al. [17].

The deterministic formulation for the factor of safety is (e.g., [25])

\[
FS = \frac{cA + N\tan \phi}{W(\sin \psi_p + \pi \cos \psi_p) + V \cos \psi_p - T \sin \theta}
\]

where
A = (H – z)/sin ψp
z = H(1 – \sqrt{\cos ψp \tan ψp})
N' = W(cos ψp – α sin ψp) – U sin ψp + T cos θ
W = 0.5γH[(1 – (z/H)^2) cot ψp – cot ψf]
U = 0.5(γb + rzA)
V = 0.5γb + rz\bar{z}
r = z_0\omega
z_0 = \sqrt{A_0/C_1}

In Eq. (6), c and ψ are the cohesion and internal friction angle of the sliding surface, respectively; z is the depth of tension crack; \z_0 is the depth of water in tension crack; α is the gravitational acceleration coefficient defined by the ratio of horizontal earthquake to gravitational acceleration.

The performance function for the rock slope stability analysis can be expressed as

\[ g(X) = F_S(c, \phi, z, r, x) - 1 \]  (7)

in which X is the random vector including the variables underlying the slope stability analysis. All random variables are assumed to be uncorrelated first. Since the performance function shown in Eq. (7) includes five random variables as shown in Table 1, the expressions of u_i (i = 1, 2, ..., 5) in the transformed space can be determined by Eq. (B.4). One length includes five random variables as shown in Table 1, the expressions of \( u_i \) (i = 1, 2, ..., 5). Applying the proposed KCP with changing increments, the reliability index, probability of failure, and design point can be obtained. To validate the proposed KCP-based reliability method, the reliability analysis results of the rock slope model shown in Fig. 3 using the KCP with changing increments, the KCP with the bisection method, and the FORM [25] are provided in Table 2. Taking the reliability index using the FORM [25] as an exact solution, the relative error in reliability index for the KCP with changing increments is only 0.6%, which is smaller than 3.6% for the KCP with the bisection method. The design point obtained from the KCP with changing increments is also more accurate than that obtained from the KCP with the bisection method. Additionally, the probability of failure is 6.5% and the corresponding reliability index is 1.514 obtained from MCS [25]. Compared with the MCS, the relative error in reliability index using the KCP with changing increments is only 3.4%, so the improved KCP approach can produce sufficiently accurate results.

Table 3 compares the global optimal angles, \( \theta_{opt} \), obtained from the KCP with changing increments with those obtained from the KCP with the bisection method [15,26]. The difference in \( \theta_{opt} \) between the two KCP approaches varies from 6.5° to 11.0°. According to the principle of the KCP with changing increments, the maximum error in angles for the improved KCP is 1°. This is because the KCP with changing increments searches the reliability index and design point with an angle increment 1° for all three angles in the neighborhood of the design point. Therefore, the more the difference in \( \theta_{opt} \) between the two KCP approaches is, the larger error the KCP with the bisection method may have.

It is widely accepted that the cohesion and the internal friction angle are negatively correlated (e.g. [13]). That is, the cohesion tends to decrease as the internal friction angle increases and vice versa. Table 4 summarizes the correlation coefficients between cohesion and internal friction angle reported in the literature. Furthermore, the tension crack depth, z, and the extent to which it is filled with water, characterized by the ratio of \( z_0 \) to z, are also negatively correlated [25]. A parametric analysis is performed here to investigate the effect of \( \rho_{c,b} \) and \( \rho_{r,z} \) on the probability of failure of the rock slope. The correlation coefficients are assumed to be negative ranging from 0 to –0.98. Two scenarios are analyzed as below: (1) only the correlation between c and ψ is considered, and (2) the correlations both between c and ψ and between z and r are considered. Fig. 4 shows the probabilities of failure of the rock slope for various correlation coefficients (i.e. \( \rho_{c,b} \) only, and both \( \rho_{c,b} \) and \( \rho_{r,z} \)) using the two KCP approaches and the FORM [25]. In Fig. 4a, both c and ψ are assumed to be normally distributed. Note that the probabilities of failure obtained from the improved KCP approach match those obtained from the FORM quite well, while the probabilities of failure using the KCP with the bisection method are slightly different from those using the FORM.

In Fig. 4b, both correlation coefficients \( \rho_{c,b} \) and \( \rho_{r,z} \) are incorporated into the reliability analysis of the rock slope. r is assumed to follow a truncated exponential distribution, which is non-normal. In this case, the probability of failure using the KCP with the bisection method is not available because the KCP approach used by Lee et al. [15,16] cannot deal with correlated non-normal random variables. It can be observed that the probabilities of failure using the improved KCP approach are consistent with those using the FORM [25]. The slight difference is due to the fact that the Nataf transformation used in the improved KCP approach accounts for the changes in correlation coefficient in the transformed space, while the FORM used by Low [25] does not account for such changes. This result implies that the improved KCP approach can evaluate the reliability of rock slope stability involving correlated non-normal variables accurately and efficiently. As expected, the probabilities of failure decrease significantly as the negative correlations between shear strength parameters become stronger. Neglecting such negative correlations between shear strength parameters will greatly underestimate the reliability of the rock slope.
Fig. 7 shows the forces acting on a typical slice $i$. In Fig. 7, $l_i$ is the length of the base of slice $i$; $h_{i,1}$ and $h_i$ are the heights of the two sides of slice $i$; $h_{w,1} = R_a h_{i,1}$ and $h_{w,i} = R_a h_i$ are the water levels of slice $i$ in which $R_a$ denotes the ratio of the water level to the height of slice side boundary; $\alpha_{i,1}$ and $\alpha_i$ are the base inclinations of slices $i-1$, $i$ and $i+1$; $\delta_i$ is the angle of inclination of external load of slice $i$; $Q_i$ is the external surcharge on slice $i$; $W_i = 0.5 \gamma_i h_i$ is the weight of slice $i$ in which $\gamma$ denotes the unit weight of rock mass; $K_h$ is the horizontal seismic coefficient; $c_i$ and $\phi_i$ are the cohesion and friction angle of soils and rocks.

Table 4
Reported correlation coefficients between cohesion and friction angle of soils and rocks.

<table>
<thead>
<tr>
<th>Correlation coefficient</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-0.70, -0.37]</td>
<td>Lumb [27]</td>
</tr>
<tr>
<td>[-0.49, -0.24]</td>
<td>Yucemen et al. [43]</td>
</tr>
<tr>
<td>-0.47</td>
<td>Wolff [39]</td>
</tr>
<tr>
<td>-0.31, -0.34</td>
<td>Young [42]</td>
</tr>
<tr>
<td>-0.55</td>
<td>Wolff [40]</td>
</tr>
<tr>
<td>-0.61</td>
<td>Cherubini [5]</td>
</tr>
<tr>
<td>-0.3</td>
<td>Hoek [13]</td>
</tr>
<tr>
<td>-0.5</td>
<td>Low [25]</td>
</tr>
<tr>
<td>-0.5</td>
<td>Molon et al. [29]</td>
</tr>
<tr>
<td>-0.5</td>
<td>Li and Low [21]</td>
</tr>
</tbody>
</table>

A typical section, Section I1–I1, of the left abutment slope is selected for reliability analysis (Fig. 6a). The potential sliding surface is highlighted with a thick line in Fig. 6a. Note that the slope being considered here is the natural slope before excavation. For the slope stability model considered, the cohesions and friction coefficients of the materials are treated as random variables. The statistical parameters of 10 random variables of the slope materials are summarized in Table 5. $c_i$ and $\phi_i$ denote the cohesions and friction coefficients of five types of materials. The means of the shear strength parameters are determined based on field test, laboratory test supplemented with engineering judgment, such as direct shear test and triaxial test [48]. The COVs of the shear strength parameters are adopted from the literature [32,5,2,25,29]. For convenience, all the random variables are assumed to be uncorrelated. Additionally, the specific weight of rock is treated as a deterministic quantity, $\gamma = 27 \text{kN/m}^3$. For the sake of illustration, the factor of safety is calculated by the residual thrust method (e.g., [4,36]). The residual thrust method assumes that the interslice force direction is parallel to the slip surface of the upper slice. The recursive expression for the interslice forces can be referred to Chen [4] and Wang et al. [36]. The considered slip surface and slices of Section I1–I1 of the left abutment slope is shown in Fig. 6b. Taking point No. 18 as the origin, Table 6 lists the coordinates of slices and associated shear strength parameters along the sliding surface.

Fig. 4. A deeply cut V-shaped valley at the dam site of Jinping I Hydropower Station.
is the lateral water pressure on the right side of slice $i$, \( P_{Wi} \) is the lateral water pressure on the right side of slice $i$, \( U_i = 0.5\gamma_w (h_{wi} + h_{mi}) \) is the water pressure at the base of slice $i$; \( N_i \) is the normal force at the base of slice $i$; \( T_i \) is the shear strength at the base of slice $i$; \( f_{r1} \) and \( f_{r2} \) are the residual thrust forces on the two sides of slice $i$; \( c_i' \) is the cohesion at the base of slice $i$; \( f_i' \) is the friction coefficient at the base of slice $i$.

The slope stability is evaluated under three working conditions, namely natural condition, rainfall condition, and seismic condition. For the natural condition, the groundwater table is below the slip surface of the slope. If the groundwater table is above the slip surface of the slope, the resulting load should be taken into consideration, which corresponds to the rainfall condition. If the seismic load is considered, it corresponds to the seismic condition. For the rainfall condition, a ratio of the water level to the height of slice \( h_w \) is set to be 0.1 based on Zhou et al. [48] to consider an earthquake of seven seismic intensity, and the resulting horizontal seismic acceleration is set to be 0.1g in which $g$ is the gravitational acceleration.

The general performance function shown in Eq. (7) for slope reliability analysis is used again. There is no explicit function for the factor of safety in the residual thrust method. The resulting performance function is an implicit function, which is not easily solvable by the traditional FORM. It should be noted that the FORM still can deal with the reliability problems involving implicit performance functions via bridging techniques such as the response surface method (RSM) and the artificial neural network (ANN) method, as demonstrated by Chan and Low [3]. The improved KCP approach

Table 5

<table>
<thead>
<tr>
<th>Materials</th>
<th>Random variables</th>
<th>Distribution</th>
<th>Mean</th>
<th>COV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lamprophyre dike X</td>
<td>( f_i )</td>
<td>Lognormal</td>
<td>0.3</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>( c_i' )</td>
<td>Lognormal</td>
<td>20</td>
<td>0.25</td>
</tr>
<tr>
<td>Fault ( f_{r2} )</td>
<td>( f_i' )</td>
<td>Lognormal</td>
<td>0.3</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>( c_i' )</td>
<td>Lognormal</td>
<td>20</td>
<td>0.30</td>
</tr>
<tr>
<td>Class II2 rock mass</td>
<td>( f_i )</td>
<td>Lognormal</td>
<td>0.9</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>( c_i' )</td>
<td>Lognormal</td>
<td>1020</td>
<td>0.15</td>
</tr>
<tr>
<td>Class IV1 rock mass</td>
<td>( f_i )</td>
<td>Lognormal</td>
<td>0.6</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>( c_i' )</td>
<td>Lognormal</td>
<td>700</td>
<td>0.18</td>
</tr>
<tr>
<td>Class IV2 rock mass</td>
<td>( f_i )</td>
<td>Lognormal</td>
<td>0.4</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>( c_i' )</td>
<td>Lognormal</td>
<td>600</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Table 6

<table>
<thead>
<tr>
<th>Points Coordinates (x,y)</th>
<th>Slices Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>2 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>3 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>4 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>5 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>6 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>7 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>8 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>9 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>10 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>11 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>12 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>13 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>14 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>15 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>16 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>17 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
<tr>
<td>18 (0,0)</td>
<td>1 - 2 - 3 - 4 - 5</td>
</tr>
</tbody>
</table>
Comparison of reliability analysis results from the improved KCP approach and the FORM + RSM for the left abutment slope.

Table 7 measures for the left abutment slope of Jinping I Hydropower Station have been suggested and implemented. Detailed information can be referred to Zhou et al. [48].

To account for the effect of the negative correlation between cohesion and friction coefficient, the probabilities of failure using the improved KCP approach, the FORM + RSM, and MCS for the natural condition are plotted against the correlation coefficient in Fig. 8. The probabilities of failure using the improved KCP approach agree well with those using the direct MCS with \(10^6\) samples in which the probability of failure is the average of six MCSs. The slight difference between them could be attributed to the nonlinearity of the performance function and the approximation of \(p_f = 1 - \Phi(\beta)\). In comparison with the results shown in Fig. 4, the results using the improved KCP approach are slightly closer to the results using MCS than those using the FORM + RSM. This is because the RSM uses an approximated performance function rather than the actual performance function that is a high dimensional nonlinear function in the present study. Such results further indicate that the improved KCP approach can be applied to evaluate the reliability of the slope involving correlated non-normal random variables accurately.

4. Conclusions

This paper has proposed an improved knowledge-based clustered partitioning (KCP) approach for reliability analysis involving correlated non-normal variables. A KCP technique with changing increments is proposed to search the design point and calculate the reliability index. Two examples of rock slope stability are investigated to illustrate the proposed approach. Several conclusions can be drawn from this study:

1. The improved KCP-based reliability method can deal with slope reliability problems involving correlated non-normal random variables, which substantially extends the application of the KCP-based reliability method.
(2) The KCP technique with changing increments is more accurate than the KCP technique with the bisection method, while the efficiencies of the two methods are comparable. Generally the KCP-based reliability method can produce reliability results with a sufficient accuracy, and its efficiency is significantly higher than the traditional Monte Carlo simulation.

(3) The improved KCP approach can be applied to effectively determine the design point and reliability index. Within the KCP approach, the evaluation of factor of safety and reliability analysis are decoupled. There is no need to evaluate the derivatives underlying the FORM; only the values of performance function need to be calculated. The improved KCP approach is especially suitable for dealing with implicit performance functions and a large number of random variables involved in slope reliability analysis.

Acknowledgments

This work was supported by the National Basic Research Program of China (973 Program) (Project No. 2011CB013506), the National Natural Science Foundation of China (Project Nos. 50839004 and 51028901), and the Program for New Century Excellent Talents in University, Ministry of Education of China (Project No. NCET-08-0415).

Appendix A. Basic properties of the Nataf transformation

The Nataf transformation [31] is adopted to transform correlated non-normal random variables into independent standard normal variables, which is briefly introduced here.

Let the correlated random input vector $\mathbf{X}$ be

$$\mathbf{X} = (X_1, X_2, \ldots, X_n)$$

Assuming the marginal CDF $F_i(\cdot)$ of each random variable $X_i$ and the correlation matrix $\mathbf{R} = (\rho_{ij})_{n \times n}$ are known. The random vector $\mathbf{X}$ can be transformed to the standardized normal random vector $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ as follows:

$$\begin{align*}
\begin{cases}
\Phi(Y_i) = F_i(X_i), \\
Y_i = \Phi^{-1}(F_i(X_i))
\end{cases}
\quad i = 1, 2, \ldots, n
\end{align*}$$

in which $\Phi^{-1}(\cdot)$ is the inverse standard normal CDF. Let $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ be $n$-dimensional standard normal vector with joint probability density functions $\phi_\mathbf{y}(\mathbf{y}, \mathbf{p}_0)$ having zero means, unit standard deviations and correlation matrix $\mathbf{p}_0 = (\rho_{0ij})_{n \times n}$, which are given by

$$\phi_\mathbf{y}(\mathbf{y}, \mathbf{p}_0) = \frac{1}{\sqrt{(2\pi)^n \det(\mathbf{p}_0)}} \exp \left( -\frac{1}{2} \mathbf{y}^T \mathbf{p}_0^{-1} \mathbf{y} \right)$$

Then, the approximate joint probability density function $f_\mathbf{x}(\cdot)$ is derived as

$$f_\mathbf{x}(\mathbf{x}) = \frac{f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)}{\phi_{X_1}(\phi_{X_2}(\phi_{X_3}(\cdots \phi_{X_n}(\cdot))))} \phi_\mathbf{y}(\mathbf{y}, \mathbf{p}_0)$$

Generally, this distribution model is referred to as the Nataf distribution.

Applying the Cholesky decomposition, $\mathbf{p}_0$ can be decomposed as

$$\mathbf{p}_0 = \Gamma_0^T \mathbf{I}_n$$

where $\Gamma_0^T$ is the upper triangular matrix obtained from the Cholesky decomposition. Then, the independent standard normal random vector $\mathbf{K}$ can be obtained as

$$K = \mathbf{I}_n^T \mathbf{Y}$$

So far, the dependent non-normal random vector $\mathbf{X}$ has been transformed to an independent standard normal random vector $\mathbf{K}$.

It should be noted that, as mentioned in Section 2, the initial sampling is carried out from the independent standard normal random variables instead of the original correlated non-normal random variables. They are later mapped into the dependent non-normal random variables. Such a transformation corresponds to the inverse Nataf transformation as follows:

$$\mathbf{Y} = \mathbf{I}_n \mathbf{K}$$

The original non-normal random variables can be given by

$$x_i = F_i^{-1}(\Phi(Y_i)), i = 1, 2, \ldots, n$$

Then, the sampling points for the independent standard normal random variables are mapped into those for the original dependent non-normal random variables.

Appendix B. Derivations of expressions of $u_i$ for 2–8 random variables.

$$u_1 = L \times \sin(\theta_1) \quad (n = 2)$$

$$u_2 = L \times \cos(\theta_1)$$

$$u_1 = L \times \sin(\theta_1) \times \sin(\theta_2)$$

$$u_2 = L \times \sin(\theta_1) \times \cos(\theta_2) \quad (n = 3)$$

$$u_3 = L \times \cos(\theta_1)$$

$$u_1 = L \times \sin(\theta_1) \times \sin(\theta_2)$$

$$u_2 = L \times \sin(\theta_1) \times \cos(\theta_2) \quad (n = 4)$$

$$u_3 = L \times \cos(\theta_1) \times \sin(\theta_2)$$

$$u_4 = L \times \cos(\theta_1) \times \cos(\theta_2)$$

$$u_1 = L \times \sin(\theta_1) \times \sin(\theta_2) \times \sin(\theta_3)$$

$$u_2 = L \times \sin(\theta_1) \times \sin(\theta_2) \times \cos(\theta_3)$$

$$u_3 = L \times \sin(\theta_1) \times \cos(\theta_2) \times \sin(\theta_3)$$

$$u_4 = L \times \sin(\theta_1) \times \cos(\theta_2) \times \cos(\theta_3)$$

$$u_5 = L \times \cos(\theta_1) \times \sin(\theta_2) \times \sin(\theta_3)$$

$$u_6 = L \times \cos(\theta_1) \times \cos(\theta_2) \times \sin(\theta_3)$$

$$u_7 = L \times \cos(\theta_1) \times \cos(\theta_2) \times \cos(\theta_3)$$

$$u_1 = L \times \sin(\theta_1) \times \sin(\theta_2) \times \sin(\theta_3) \times \sin(\theta_4)$$

$$u_2 = L \times \sin(\theta_1) \times \sin(\theta_2) \times \cos(\theta_3) \times \cos(\theta_4)$$

$$u_3 = L \times \sin(\theta_1) \times \cos(\theta_2) \times \sin(\theta_3) \times \cos(\theta_4)$$

$$u_4 = L \times \sin(\theta_1) \times \cos(\theta_2) \times \cos(\theta_3) \times \cos(\theta_4)$$

$$u_5 = L \times \cos(\theta_1) \times \sin(\theta_2) \times \sin(\theta_3) \times \cos(\theta_4)$$

$$u_6 = L \times \cos(\theta_1) \times \cos(\theta_2) \times \sin(\theta_3) \times \cos(\theta_4)$$

$$u_7 = L \times \cos(\theta_1) \times \cos(\theta_2) \times \cos(\theta_3) \times \cos(\theta_4)$$
where $n$ is the number of random variables; $u_i (i=1,2,3,\ldots, n)$ are possible values of $K_i$ in the transformed space; and the angles ($\theta_1$, $\theta_2$, and $\theta_3$) and the length ($L$) are the unknowns to be determined.

References