Determination of the Most Critical Failure Mode in Limit Analysis of Slopes Using Simulated Annealing Method

Yu-Jie Wang, China Institute of Water Resources and Hydropower Research(IWHR), Beijing, China
Paul Chiasson, Faculté d'ingénierie, Université de Moncton, Moncton, Canada
Jian-Hua Yin, Department of Civil and Structural Engineering, Hongkong Polytechnic University, Hongkong

ABSTRACT
In most slope stability analyses that use for example limit equilibrium methods or the upper bound approach, an important step is to search for the most critical failure mode and corresponding global minima of the factor of safety. Optimisation methods are generally used to accomplish this task. Conventional optimisation methods, however, only accept downhill steps and sometimes get trapped into a local minimum. The simulated annealing method may accept both downhill and uphill steps in searching for the critical failure mode, thus finding the global minima of the factor of safety. Although simulated annealing has been widely used to solve multi-extreme problems in other disciplines, the application of the method to geotechnical engineering problems is rarely reported. This paper presents procedures and applications of the simulated annealing method for slope stability analysis. In particular, the paper discusses how to choose a suitable cooling strategy, random steps and stopping criterion. These are the core issues of the simulated annealing method. Application of the simulated annealing method to two illustrative examples is presented. A comparison is made between the simulated annealing method and the random-search simplex method. The advantages, effectiveness and limitation of the simulated annealing method are presented and discussed.

RÉSUMÉ
Dans la plupart des méthodes de stabilité de pente, comme celles d'équilibre limite ou l'approche par borne supérieure (analyse limite), la recherche de la surface de glissement critique est une étape de grande importance. Car c'est à cette surface critique que correspond le minimum global du facteur de sécurité. Les méthodes d'optimisation sont généralement employées pour accomplir cette tâche. Cependant, les méthodes conventionnelles d'optimisation sont contraintes à chercher seulement suivant des pas descendants. Ainsi, parfois, elles se prennent au piège d'un minimum local. La méthode du recuit simulée, sous certaines réserves, accepte des pas ascendants en plus des pas descendants. Elle permet ainsi la découverte du minimum global du facteur de sécurité. Bien que le recuit simulé ait été largement employé pour résoudre des problèmes à extremums multiples dans d'autres disciplines, l'application de la méthode à des problèmes géotechniques est rarement rapportée. Cet article présente une application de la méthode du recuit simulé au problème de l'analyse de stabilité des pentes. L'article discute du choix d'une stratégie appropriée pour le refroidissement, les pas aléatoires et le critère d'arrêt. Ces questions sont fondamentales à la méthode du recuit simulé. L'application de la méthode du recuit simulé à deux exemples explicatifs est présentée. Une comparaison est faite entre la méthode du recuit simulé et la méthode du simplex à recherche aléatoire. Les avantages, l'efficacité et les limitations de la méthode du recuit simulé sont identifiés et discutés.

1. INTRODUCTION

A number of methods exist for slope stability analysis, including limit equilibrium methods (Bishop 1955, Janbu 1957, Morgenstern and Price 1965, Spencer 1967) and the upper bound method (Chen 1975, Michalowski 1995, Donald and Chen 1997). Due to assumptions regarding distribution and inclination of interslice forces, limit equilibrium methods do not yield a unique solution to the factor of safety. Recently, application of the upper bound theorem of limit analysis to slope stability analysis was made extensively (Chen 1975, Michalowski 1995, Donald and Chen 1997). In either method, limit equilibrium or upper bound, an important step is to search for the most critical failure mode and the global minima of the factor of safety. This may be performed by using optimisation methods.

In this paper, the upper bound method combined with the simulated annealing algorithm is used to search for the most critical failure mode. The upper bound analysis for slope stability problems generally involves two steps: (1) introduce a kinematically admissible failure mode within which the work-energy-balance equation is established and the factor of safety is the only one unknown in the equation which can be solved by iteration; (2) repeat step 1 to generate a large number of kinematically admissible failure modes among which the minimum of the factor of safety and the corresponding critical failure mode can be obtained. When the work-energy-balance equation is established, the main burden to the upper bound analysis is to search for the most critical failure mode by using an optimisation method.

Generally, optimisation methods for locating the critical slip surface in literature are locally convergent. The random search technique is incorporated into conventional optimisation methods to improve global convergence. However, when an optimisation problem involves many degrees of freedom and the slope profile

contains many different soil layers, conventional algorithms may suffer from a premature termination at which the solution is not actually the global minimum. A technique called "simulated annealing method" has attracted significant attention (Kirkpatrick et al. 1983, Bohachevsky et al. 1986). The simulated annealing method may step out of a local minimum with a technique that allows deteriorating (or uphill) steps through an accepting criterion. Thus, the simulated annealing method is likelier to approach the global solution in comparison to conventional methods.

The objectives of this paper are: (a) incorporating the globally convergent simulated annealing method into the upper bound analysis to find the most critical failure mode and the associated global minimum factor of safety, and (b) comparing the simulated annealing method with the random-search simplex method (Chen 1992).

2. DEFINITION OF THE OPTIMISATION PROBLEM IN THE UPPER BOUND APPROACH

The upper bound theorem of limit analysis is a useful tool to solve geotechnical problems. For slope stability analysis, the potential sliding mass is generally induced to the plastic limit state by mobilizing the strength parameters reduced by factor of safety *F*. Thus, the factor of safety, rather than surface loads, is the upper bound to the true value of solution.

An example of a schematic failure mode is shown in Fig.1. The slip surface is divided by a number of nodal points, which are connected by straight lines or smooth curves. Coordinates of *i*-th nodal point are denoted by (x_i, y_i). The inclination of wedge side associated to the *i*-th nodal point is designated as δ_i, which is positive clockwise from positive *y* axis. According to different movement characteristics of nodal points on a slip surface (such as A-B-C-D-E-F in Fig.1), the nodal points are classified into two categories:

(a) Unconstrained nodal points. This type of nodal point can move in any direction, the abscissa x_i and ordinate y_i are indeterminate. Points *B* and *E* as shown in Fig. 1 are examples of unconstrained nodal points.

(b) Constrained nodal points. Very often, development of a failure mechanism is controlled by the presence of a weak seam or structural joint. In this case, such as the points *C* and *D* in Fig. 1, nodal points are constrained to move along a specified direction (i.e. the weak seam). Coordinates of these nodal points are dependent. In addition, the head and end nodal points of a slip surface, such as Points *A* and *F* in Fig.1, are constrained to move along the slope surface. In both cases above, if the abscissa x_i is given, the ordinate y_i is constrained to follow the prescribed direction and can be expressed as function of x_i and the unit vector. Therefore, constrained nodal points involve only one degree of freedom, which is chosen for simplicity as the abscissa x_i.

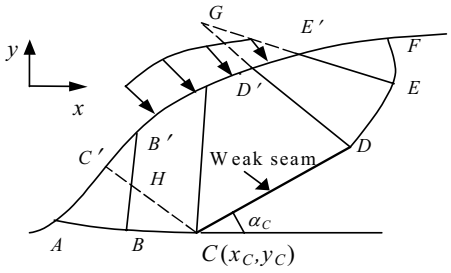

Figure 1. A kinematically admissible failure mode

In addition, the inclinations δ of interslice sides are varied during optimisation. This adds a degree of freedom per interslice side. Suppose that there are *k* number of unconstrained nodal points and *l* number of constrained nodal points for a slip surface (such as *A-B-C-D-E-F* in Fig.1). The *k* unconstrained nodal points are represented with the coordinates x_i and y_i (i=1, 2, …, k). The *l* constrained nodal points are denoted by abscissas X_i, i=1, 2, …, *l*. The inclinations associating with the nodal points except for the head and end nodal points for the *k*+*l* nodal points are denoted by δ_i, i=1, 3, …, *k*+*l*-2. The factor of safety *F* for slope stability analysis or the loading factor η for bearing capacity analysis can then be expressed as a function with respect to x_1, y_1, x_2, y_2, …, x_k, y_k for *k* unconstrained nodal points, X_1, X_2, …, X_l for *l* constrained nodal points and δ_1, δ_2, …, δ_{k+l-1} for *k*+*l*-2 nodal points. In this case, therefore, the number of degree of freedom for this failure mechanism totals to 3*k*+2*l*-2. Generally, variables x_1, y_1, x_2, y_2, …, x_k, y_k; X_1, X_2, …, X_l; δ_1, δ_3, …, δ_{k+l-2} define a space of n=(3*k*+2*l*-2) dimensional variables. For convenience, a vector $\boldsymbol{Z}$ is introduced, whose components are given by

$$\boldsymbol{Z} = (z_1 z_2, \cdots, z_{3k+2l-2})^T = (x_1, y_1, x_2, y_2, \cdots x_k, y_k; X_1, X_2, \cdots, X_l; \delta_1, \delta_2, \cdots, \delta_{k+l-2})^T \quad [1]$$

Thus, the factor of safety *F* is a function

$$\begin{aligned} F &= F(\boldsymbol{Z}) \\ &= F(x_1, y_1, x_2, y_2, \cdots x_k, y_k; X_1, X_2, \cdots, X_l; \delta_1, \delta_2, \cdots, \delta_{k+l-2}) \end{aligned} \quad [2]$$

To guarantee generation of a kinematically admissible failure mode, some constraint conditions must be introduced explicitly and implicitly. Referring to Fig. 1, the five nodal points connecting the slip surface are arranged in the sequence of *A-B-C-D-E-F*. Therefore, the following inequality should be satisfied.

$$X_A < x_B < X_C < X_D < x_E < X_F \quad [3]$$

The intersection point of inclined interslice sides should locate outside the sliding mass or at least at the slope surface rather than inside it. For example, point *G,* at the intersection of lines *DD′* and *EE′*, is outside the sliding mass and is thus validated and accepted. However, point *H*, at the intersection of interfaces BB′ (dash line in Fig. 1) and *CC′*, is inside the sliding mass, and is thus non valid and rejected. For such a case, the generated failure mode is not kinematically admissible and is thus discarded from

the analysis.

The slip surface, which is generated by connecting nodal points with straight lines or splines, is required to be non-concave. Thus, a point over the interval bounded by any two nodal points along the slip surface must fall below the straight line passing through the two nodal points. The inclination of the base of the *i*-th wedge is defined as α_i and is positive anti-clockwise from the positive *x* axis (as shown in Fig. 1). Subscript *i* in α_i is increased from 1 to *k*+*l*-2.

Thus, when the coordinates of nodal points and their associated interslice side inclinations are varied while satisfying constraints implicitly or explicitly, a large number of kinematically admissible failure modes may be generated. For each kinematically admissible failure mode, the factor of safety *F* can iteratively be determined by Eq.2. There are a large number of different kinematically admissible failure modes and thus a large number of corresponding factor of safety values. An optimisation algorithm must therefore be incorporated into the upper bound computer code to search the critical failure mode and its associated factor of safety F.

3. SIMULATED ANNEALING METHOD

The simulated annealing method is a technique for combinatorial optimisation problems with multiple extremes, such as the well-known *N*-city traveling salesman and computer circuit design problems (Kirkpatrick et al. 1983, Bohachevsky et al. 1986). This method can step out local extremes with a specified accepting probability allowing the uphill steps in the optimisation process compared to those optimisation algorithms accepting downhill steps only.

Traditional optimisation methods are similar to the quickly cooled process for a liquid metal, greedily searching for the minima in the downhill direction from the initial starting points. The searching scheme, sometimes, leads to a local rather than the global minimum (Press et al. 1986). The simulated annealing algorithm, however, adopts a technique that allows uphill steps to be accepted by a specified accepting probability to avoid the occurrence of local traps.

The simulated-annealing method for optimisation problems begins with selecting a starting point, $\boldsymbol{Z_0}$, and making a random step $\Delta\boldsymbol{Z}$, which is generated by a particular technique that will be discussed later. At each step, change ΔF in objective function *F* is evaluated.

$$\Delta F = F(\boldsymbol{Z_0} + \Delta\boldsymbol{Z}) - F(\boldsymbol{Z_0}) \qquad [4]$$

If the value of the objective function *F* is reduced, that is, $\Delta F < 0$, the random step is accepted and $\boldsymbol{Z_0}+\Delta\boldsymbol{Z}$ is taken as a new starting point for the next random step. However, if the value of the objective function is increased, $\Delta F > 0$, the random step may still be accepted based on an auxiliary judgment. A random number ρ is generated from the uniform distribution over (0,1) and compared to the value of $exp[-\Delta F / T]$. The random step in the case of $\Delta F > 0$ is accepted if $\rho < exp[-\Delta F / T]$, otherwise the random step is rejected and a new random step is attempted. The parameter *T* plays the role of temperature similar to that in physical annealing. At a given value of temperature *T*, a number of accepted random steps are required. The number of steps required depends on the magnitude of temperature *T*. For the optimisation problems with multi-minimums, as *T* is decreased slowly enough rather than quickly, the objective function can avoid to be trapped into local minima and finally reach the global minimum. The reason is that the condition $\rho < exp[-\Delta F / T]$ allows detrimental steps that lead to increase in the value of objective function.

Thus, procedures of a simulated annealing algorithm for an optimisation problem are (a) defining the objective function; (b) constructing a practical random-step scheme; (c) choosing a cooling strategy (especially, selecting a starting temperature and rules to determine when the current temperature should be lowered and how much); (d) determining a stopping criterion under what conditions the optimisation process is terminated.

3.1 Cooling Strategy

The cooling strategy, which is designed to control the optimisation process (annealing process), consists of: (a) the initial starting temperature, T_0, (b) how to reduce the temperature to satisfy the slow annealing characteristics of optimisation, (c) how many random steps are taken at each given value of temperature. In addition, the cooling strategy is dependent upon optimisation problems and must be determined by a certain number of tests for a particular type of problems.

The starting temperature, T_0, should be sufficiently high to allow uphill moves away from local minima (Kirkpatrick et al., 1983). But too high temperature will accept more deteriorating random steps that will decrease efficiency and waste computing time. A recipe for determining a reasonable starting temperature, therefore, is needed. Before starting simulated annealing, a number of random walks in accepted variable space (generally 100 or higher) are generated to determine the maximum $F_{\max}$ and the minimum $F_{\min}$ of the objective function. Thus, the starting temperature T_0 is defined as

$$T_0 = F_{max} - F_{min} \qquad [5]$$

The temperature during optimisation is reduced by a damping function, that is,

$$T_k = T_0 \cdot (\alpha_T)^k \qquad [6]$$

where T_k is temperature *T* decreased *k* times from the starting temperature, T_0. At each temperature T_k, *m* random steps are simply carried out. Constant α_T is referred to as a temperature damping factor. It is in the range of $0 < \alpha_T < 1$ and should be chosen by trial and

error (Vanderbilt and Louie 1984). Thus, the cooling strategy presented here is: (a) begin with a starting temperature, T_0, determined by Eq.5, (b) decrease the temperature by Eq.6, and (c) carry out a fixed number of m random steps at each temperature.

3.2 Random Step ΔZ

Since simulated annealing is a powerful stochastic technique for optimisation problems, the optimal magnitude and direction of random steps are not known in advance during optimisation. The steps, which are too small, will most likely be accepted and be inefficient in exploring the variable space. The steps, which are too large, will most likely be rejected and time consuming. In either step, there is little information available for determining a reasonable random step for the simulated annealing algorithm. Hence, to find a better minimum value of the objective function and save computing time, it is crucial to determine appropriate random steps for continuous optimisation. Generally, determination of random step $\Delta\boldsymbol{Z}$ involves choosing a randomly generated direction and a reasonable magnitude for the step.

As proposed by Bohachevsky et al. (1986), the procedure for choosing the direction of a random step is to generate n random numbers (corresponding to the n dimensional variables of Eq. 2), $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$, from the uniform distribution on [–1,1] and to convert them into direction cosines $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$, where

$$u_i = v_i \Big(\sum_{i=1}^{n} v_i^2 \Big)^{-1/2} \qquad [7]$$

Assume that the area of interest, Ω, of the optimisation problem, can be represented by a finite set, which is taken to be a hypercube defined by upper and lower bounds on each variable

$$\Omega = \{ \boldsymbol{Z} \,|\, z_i^l < z_i < z_i^u,\ i = 1, \ldots, n \} \qquad [8]$$

To take into account unevenness in size of the domain of each optimisation variable, positive constants, κ_i (i=1, …, n), are proposed

$$\kappa_i = (z_i^u - z_i^l) / max\{ (z_1^u - z_1^l), \cdots, (z_i^u - z_i^l), \cdots, (z_n^u - z_n^l) \} \quad (i = 1, 2, \ldots, n) \qquad [9]$$

The magnitude of random step, Δr, is decreased with temperature by a damping function, but fixed at temperature T_k. The magnitude of random step, Δr_k, at temperature T_k, is defined as

$$\Delta r_k = \Delta r_0 \cdot (\alpha_r)^{k-1} \qquad [10]$$

where α_r is the random damping factor, $0 < \alpha_T < \alpha_r < 1$. Δr_0 is the initial value of a random step, which is dependent on the properties of the objective function, the desired accuracy and resolution. Accordingly, the component of a random step, Δr_k, in the direction of variable z_i is defined as

$$\Delta z_i = \kappa_i \cdot \Delta r_k \cdot u_i \qquad i = 1, 2, \ldots, n \qquad [11]$$

3.3 Stopping Criterion

Specification of a satisfactory termination criterion is the difficult part of the simulated annealing method. When the optimal value of an objective function is known in advance, the stopping criterion is defined in terms of the relative deviation in values of the objective function

$$\frac{< F > - F_{min}}{F_{min}} < \varepsilon_f \qquad [12]$$

where <F> is the average of values of the objective function evaluated by the successful random steps satisfying $\Delta F < 0$ or $\rho < exp[-\Delta F / T]$ in the case of $\Delta F > 0$ at temperature T_k, F_{min} is the minimum of the objective function before temperature T_{k+1}, and ε_f is the allowable error for the objective function (generally, $\varepsilon_f = 1.0 \times 10^{-5}$).

To prevent that Eq.12 is not satisfied during the optimisation and an infinite loop occurs for a given problem, further stepping search is chosen to stop when temperature T_f is less than an allowable error ε_T (generally, $\varepsilon_T = 1.0 \times 10^{-6}$). Thus, the auxiliary stopping criterion is

$$T_f \le \varepsilon_T \qquad [13]$$

4. SIMULATED ANNEALING ALGORITHM

After setting the strategy for random walks and cooling for a specific optimisation problem, steps of the simulated annealing algorithm used for optimisation problems may be described as follows:

(a) Choose a starting point $\boldsymbol{Z_0}$ randomly or judiciously on the basis of some knowledge of the problem and determine the corresponding value of objective function $F(\boldsymbol{Z_0})$ and initial temperature T_0 according to Eq. 5, set $F_{min} = F(\boldsymbol{Z_0})$;

(b) Generate a small random walk $\Delta\boldsymbol{Z}$ based on Eq.11;

(c) Calculate the change in values of the objective function at any position in space of continuous variables $\Delta F = F(\boldsymbol{Z_k}) - F(\boldsymbol{Z_{k-1}})$, and if $F(\boldsymbol{Z_k}) < F_{min}$, then $F_{min} = F(\boldsymbol{Z_k})$;

(d) Check the number of random walks m_k at temperature T_k, if $m_k \le m$ continue, otherwise, check the stopping criteria in Eqs.12 and 13;
If satisfied, stop the search;
Otherwise, $k = k+1$ and calculate temperature and step damping factor α_T, α_r, next value of temperature T_{k+1}, the magnitude of random step, Δr_k ;

(e) For $\Delta F \le 0$ set $\boldsymbol{Z}_{k+1}=\boldsymbol{Z}_k+\Delta\boldsymbol{Z}$, increment k by 1 and go to step (2);

(f) For $\Delta F>0$ calculate the value of $exp[-\Delta F / T_k]$. Generate a random number ρ in the range of $0<\rho<1$, form specified distribution on [0,1] (uniform distribution in most cases), and compare:

If $exp[-\Delta F / T_k] < \rho$, the random step is rejected and go to (b);

If $exp[-\Delta F / T_k] \ge \rho$, the random step is accepted, set $\boldsymbol{Z}_{k+1}=\boldsymbol{Z}_k+\Delta\boldsymbol{Z}$, increment k by 1 and go to (b).

5. CALCULATION OF THE FACTOR OF SAFETY

As mentioned before, the factor of safety for slope stability problems is defined as a function of coordinates of nodal points on a slip surface and associated inclinations of interfaces between wedges. Although a set of variables, which must satisfy the kinematically admissible condition, corresponds to one value of the factor of safety, there is still no analytical relationship between the minimum factor of safety and the variables. Based on the upper bound theorem, the work-energy-balance established on the kinematically admissible failure mode serves as the bridge between the factor of safety (an objective function) and those variables. Fig. 2 shows an *n*-wedge failure mode for the slope stability analysis using the upper bound theorem of plasticity. The work-energy-balance equation in this general failure mode is written as (Donald and Chen 1997, Wang 2001)

$$\begin{aligned}&\sum_{i=1}^{n}(c_e^i+u^i\cdot\tan\phi_e^i)\cdot V_i\cdot\cos\phi_e^i+\\&\sum_{i=1}^{n-1}(c_{ej}^i+u_j^i\cdot\tan\phi_{ej}^i)\cdot V_i^j\cdot\cos\phi_{ej}^i\\&=\sum_{i=1}^{n}W_i\cdot\sin(\alpha_i-\phi_e^i)\cdot V_i\\&+\sum_{i=1}^{n}[T_x^i\cdot\cos(\alpha_i-\phi_e^i)+T_y^i\cdot\sin(\alpha_i-\phi_e^i)]\cdot V_i\end{aligned} \qquad [14]$$

where c_e and ϕ_e are the mobilized cohesion and friction angle respectively. The subscript "*e*" denotes the strength parameters, which have been mobilized using the factor of safety *F*, that is, $c_e=\frac{c}{F}$ and $\tan\phi_e=\frac{\tan\phi}{F}$. In Eq.14, W_i, T_x^i and T_y^i are weight of *i*-th slice, the component of the *i*-th surface load acting on the slope surface in *x* and *y* directions, respectively. It is noted that in Eq.14 "*i*" stands for the physical quantities associated with *i*-th wedge block and "*j*" stands for parameters and forces at the *j*-th interface and others without "*j*" means the parameters and forces at the base of the slice (i.e. on the slip surface).

It should be noted that all geometrical parameters in Eq.14 are related to the vector $\boldsymbol{Z}$ defined in Eq.1. Thus for a kinematically admissible failure mode, the factor of safety, is the only unknown in Eq.14 and can be obtained iteratively.

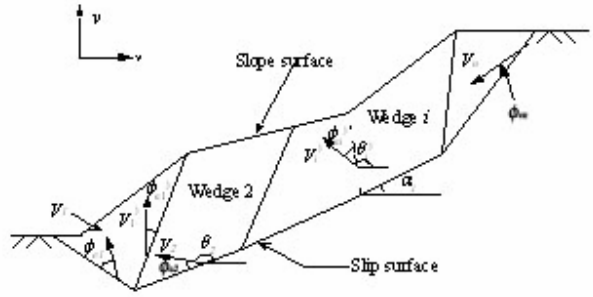

Figure 2. An illustration of *n*-slice failure mode

6. APPLICATION AND COMPARISON

Example1 A two-layer soil slope with a two-wedge failure mode

Figure 3 shows a horizontal 1 to vertical 1 slope with two soil layers. The horizontal interface between the two soil layers is 5 m deep from the slope top. The strength parameters and unit weights for the two types of soils and other geometric properties are presented in Fig. 3. The failure mode for the two-layer soil slope consists of two wedges, that is, triangles *ABE* and *BEF*. In the process of generation of kinematically admissible failure modes for this example, Point *A* is fixed at (0,0), Point *E*, which is denoted by coordinates (x_E, y_E) can move freely, Point *F* denoted by (x_F, 10) is only allowed to move along the top surface. The failure mode, therefore, includes 3 degree of freedom, that is, x_E, y_E, x_F. To guarantee the generated failure mode is geometrically admissible and the slip surface is non-concave, the following relationship must be satisfied

$$\frac{y_E}{x_E}\le\frac{10-y_E}{x_F-x_E} \qquad [15]$$

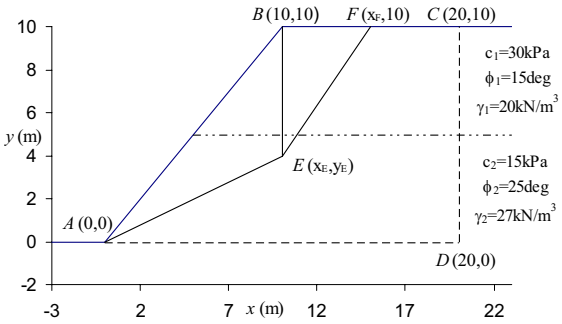

Figure 3. Two-wedge failure mode for two-layer soil profile

In addition, point *F* should not overlap points *B* and *E*. Application of the simulated annealing method involves the following steps:

(a) Select a searching area that approximately covers the critical failure mode by experience and determine the

lower and upper bounds on each degree of freedom. The searching area for Example 1 is the area surrounded by straight lines *AB*, *BC*, *CD* and *AD*. According to the searching area, the lower and upper bound of the variables can be determined and presented in columns 1 and 2 in Table 1. The failure mode generated should be located between the two bounds and satisfy the kinematically admissible conditions of the geometry (such as Eq.3 and so on).

(b) Choose suitable values for temperature damping factor, α_T, number of random steps, m, at each temperature T_k, and random damping factor α_r. The principal criterion for the selection of these parameters is to guarantee the global extreme to be approached in much less time and be achieved generally by trial and error for a specific example. For simplicity, in Example 1, α_T=0.9, α_r=0.98, m=150.

(c) Determine the initial magnitude of a random step.
If $\nabla\chi$ is designated to denote the maximum difference between the upper and lower bounds on the variables. Δr_0 is generally taken as $\frac{\nabla\chi}{20} < \Delta r_0 < \frac{\nabla\chi}{50}$ (Wang 2001). Figure 4 shows variations of the factor of safety with the optimisation process for different value of Δr_0 = 0.1, 0.2, 0.3 and 0.4. Fig. 4(a) shows that the simulated annealing method is likely to give a local minimum when Δr_0 is taken as a small value. Figs. (b), (c) and (d) show that it is likely for simulated annealing to locate the global minimum when using a relatively large value of the initial random step as recommended previously. In Example 1, Δr_0 is taken as 0.3 in the following analysis. It can be also seen from Fig.4 that the simulated annealing method accepts not only the downhill random steps (decreasing F) but also the uphill random steps (increasing F). However, the disadvantage is that approaching the global minimum takes much more time than for other conventionally used optimisation methods.

(d) Choose an initial or seed failure mode.
Since simulated annealing is designed for multi-extreme problems, choice of the initial or seed failure mode should not affect the final solution. Starting from the seed failure mode (Case 1 in Table 1 and Fig.4c), 100 successfully generated random steps give 0.60 for initial temperature T_0. Then using the simulated annealing algorithm proposed in the preceding section, the final global solution to the factor of safety is obtained as 1.616.

Figure 5 presents changes in the factor of safety during the optimisation process (successful iteration number). It shows that the factor of safety decreased steadily during optimisation to finally reach 1.616, which is considered to be the global minimum of Example 1.

Table 1. A summary of initial estimates on failure mechanisms

	χ^-	χ^+	Case 1	Case 2	Case 3	Case 4	Local	Global
x_E	0	20	5.0	10.0	15.0	17.0	11.40	11.99
Y_E	0	10	1.0	1.0	2.0	4.0	4.68	5.48
x_F	10	20	12.0	13.0	17.0	19.0	16.55	16.08
F	-	-	3.190	2.493	2.142	2.060	1.620	1.616

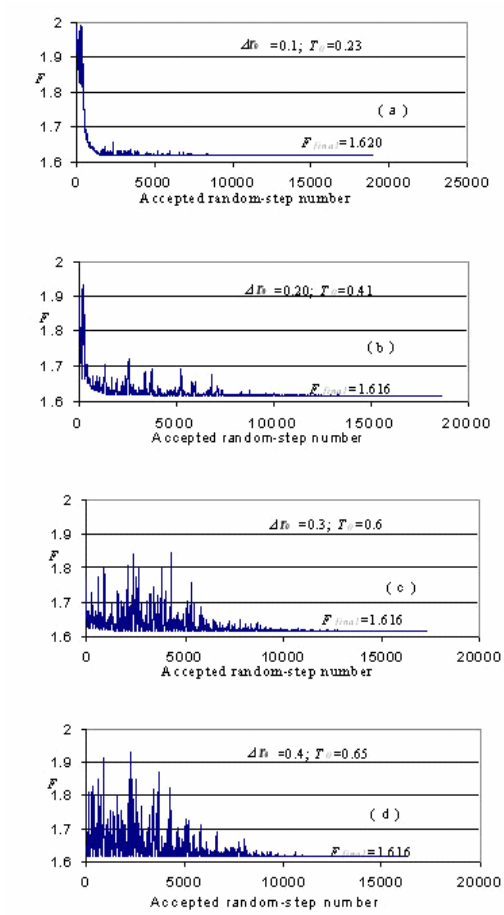

Figure 4. Changes in the factor of safety during optimisation

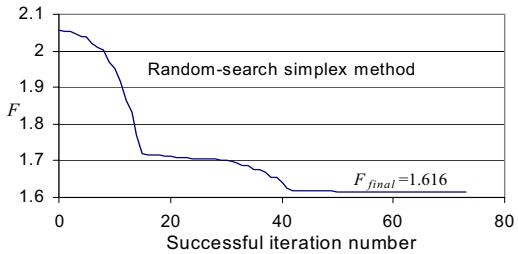

Figure 5. Changes in the factor of safety during optimisation with random-search simplex method

One of the important characteristics for so-called globally convergent optimisation methods is that the final solution to the global minimum is independent of the starting point in the variable space. To compare convergence to the global minimum for Example 1 between simulated annealing (globally convergent method) and the simplex method (locally convergent method), four different starting points (Cases 1 to 4) in the variable space (as shown in Fig. 6 and listed in Table 1) are considered. Simulated annealing gives F_{final}=1.616 whatever the four studied seed failure modes. Simplex method gives F_{local}=1.620 for Cases 1 and 2, and F_{final}=1.616 for Cases 3 and 4. The two failure modes for F_{local}=1.620 and F_{final}=1.616 are shown in Fig. 6 and summarized in Table 1.

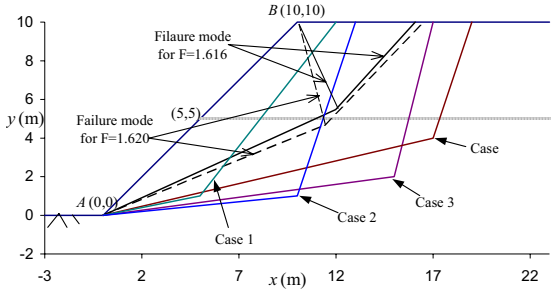

Figure 6. Different initial or seed failure modes (Cases 1, 2, 3, 4), local critical failure mode (F=1.620), and global critical failure mode (F=1.616).

Example 2: A slope with three soil layers

To further illustrate the feasibility of the simulated annealing method in searching for the most critical failure mode for a slope with a complicated soil profile and to compare with the random-search simplex method, a three-layered soil slope with a phreatic line is presented in Figure 7. The effective soil strength parameters for different soil layers are listed in Table 2. The potential slip surface in this example is approximated by four nodal points A, B, C and D connected by smooth curves. Point A is only allowed to move along the slope surface; Points B and C move independently; Point D is also specified to move along the top horizontal slope surface. In order to increase calculation accuracy, the slip surface is further divided into 16 points by interpolation. Since the newly added nodal points by interpolation do not provide additional degrees of freedom, the failure mode for Example 2 has 8 degrees of freedom, that is, X_A, x_B, y_B, δ_B, x_C, y_C, δ_C, X_D. The initial failure mode represented by the 8 variables is presented in Fig. 7 and the values and associated factor of safety are listed in Table 3.

Table 2 The soil strength parameters for Example 2

	c (kN/m^2)	ϕ (°)	γ (kN/m^3)
Soil layer 1	0.0	32.0	19.5
Soil layer 2	5.3	23.0	19.5
Soil layer 3	7.2	20.0	19.5

When α_T=0.9, α_r=0.98, m_k=300, Δr_0 =0.23 and the initial failure mode in Table 3 are adopted, the simulated annealing method yields that F=1.154, the random-search simplex method gives F=1.170. The critical failure modes obtained using these two methods are presented in Fig. 8.

The corresponding critical failure modes are listed in Table 3.

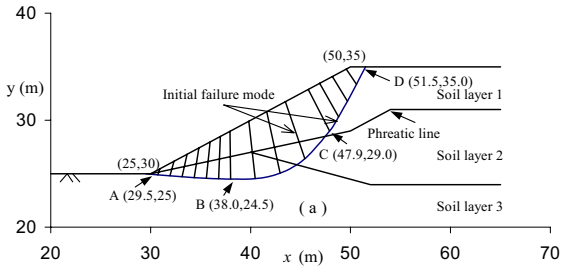

Figure 7. Initial or seed failure mode for three-layered soil slope (example 2).

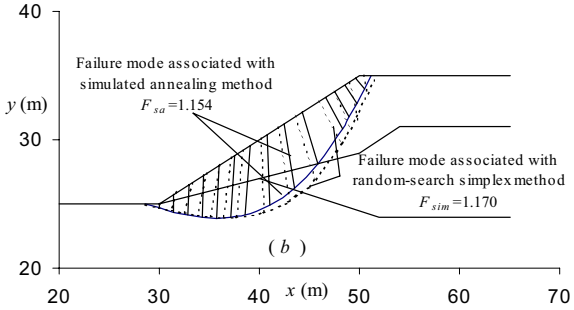

Figure 8. Failure modes associated with simulated annealing and random-search search methods for Example 2.

Table 3. Comparison of the simulated annealing method with the random-search simplex method.

	Initial failure mode	Final failure mode for the simulated annealing method	Final failure mode for the random-search simplex method
X_A (m)	29.50	28.55	28.59
x_B (m)	37.94	38.62	37.72
y_B (m)	28.97	24.13	24.00
δ_B (°)	38.00	4.50	6.68
x_C (m)	45.76	48.27	48.15
y_C (m)	32.88	30.85	29.68
δ_C (°)	47.90	-33.59	-36.16
X_D (m)	51.50	51.19	51.68
F	1.246	1.154	1.170

7. CONCLUSIONS

The upper bound method as a simple and practical tool is more and more commonly appreciated and widely used in geotechnical engineering. Since the upper bound theorem only guarantees the solution is an upper bound to the "true solution", powerful optimisation techniques must be integrated to search for the minimum value of the "upper bound" solutions. This is especially true when problems concerned involve complicated geometries and different soil strength parameters. Simulated annealing is considered to be a so-called globally convergent optimisation procedure (Kirkpatrick et al. 1983, Bohachevsky et al. 1986). In this paper, simulated annealing is applied to solve the optimisation problem involved in upper bound analysis. Two applications of

simulated annealing applied to upper bound analysis are presented and compared to the random-search simplex method proposed by Chen (1992).

Based on this study, analysis and comparison, the following conclusions may be drawn:

(a) The simulated annealing method allows deteriorating or uphill steps to be accepted with a special accepting criterion. This is different from the conventional optimisation methods, which allow only downhill steps. Thus, the simulated annealing method may step out local minima traps and is able to approach the global minimum value of an objective function and locate the corresponding critical failure mode.

(b) Compared with the random-search simplex method, simulated annealing is more likely to obtain a solution closer to the true global minimum.

(c) Since simulated annealing is an analogy to metal annealing in nature, the method would take much more time to approach the global minimum than if using conventional methods, although these may give a local minimum.

ACKNOWLEDGEMENTS

The research and preparation of this paper have received financial supports from a RGC grant (PolyU 5064/00E) of the University Grants Committee of the Hong Kong SAR Government of China, the Hong Kong Polytechnic University, and China Institute of Water Resources and Hydropower Research (IWHR). These financial supports are gratefully acknowledged.

REFERENCES

Bishop, A.W. 1955. The use of the slip circle in the stability analysis of slopes. Geotechnique, **5(**1),7-17.

Bohachevsky, I. O., Johnson, M.E., and Stein, M.L. 1986. Generalized simulated annealing for function optimisation. Technometrics, **28**(3) : 209-217.

Chen, W. F. 1975. Limit analysis and soil plasticity. Amsterdam: Elsevier.

Chen, Z.Y. 1992. Random trials used in determining global minimum factors of safety of slopes. Canadian Geotechnical Journal, **29**(2): 225-233.

Donald, I.B. and Chen, Z.Y. 1997. Slope stability analysis by the upper bound approach: fundamentals and methods. Canadian Geotechnical Journal, **34**(6): 853-862.

Janbu N. 1957. Earth pressures and bearing capacity calculations by generalized procedures by slices. Pro. *4th Int. Conf. Soil Mech*., London 1957; **2**:207-212.

Kirkpatrick, S., Gelatt, C.D. Jr. and Vecchi, M. P. 1983. Optimisation by simulated annealing. Science, Vol 220(4598), pp.671-680.

Michalowski, R.L. 1995. Slope stability analysis: a kinematical approach. Geotechnique, **45**(2): 283-293.

Morgenstern, N.R. and Price, V.E. 1965. The analysis of the stability of general slip surfaces. Geotechnique, **15**:79-93.

Press, W. H., Flannery, B. P., Teukolsky, S. A., Vetterling, W. T. 1988. Numerical recipes: the art of scientific computing, Cambridge University Press.

Spencer, E. 1967. A method of analysis of the stability of embankments assuming parallel inter-slice forces. Geotechnique, **17**(1):11-26.

Vanderbilt, D. and Louie, S.G. 1984. A monte carlo simulated annealing approach to optimisation over continuous variables. The Journal of Computational Physics, **56**: 259-271.

Wang, Y.-J. 2001. Stability Analysis of Slopes and Footings Considering Different Dilation Angles of Geomaterial, Ph.D. Thesis, The Hong Kong Polytechnic University.