CRITICAL FAILURE SURFACE LOCATION USING SIMPLE GENETIC ALGORITHM AND MULTIPLE WEDGE SLOPE STABILITY

Paul McCombie
Department of Architecture and Civil Engineering, University of Bath, United Kingdom

ABSTRACT: *Established slope stability analysis methods cope well with moderately non-circular shear surfaces, and the Simple Genetic Algorithm (SGA) has been used successfully to find the critical slip surface. In complex situations, for example with concentrated soil reinforcement, strong or weak layers, or concentrated applied loading, the assumptions made to allow 'rigorous' analysis are likely to be simplistic. The results of individual analyses could be in error, and the location of the critical slip surface is thus likely to be wrong. The factor of safety could be overestimated because of the impossibility of full shear strength being mobilised all along a shear surface at once, or underestimated because the need to distort the sliding mass of soil as movement takes place is not properly accounted for. In such cases, an engineering judgment has been necessary to identify the potential problem (if it has been recognised at all), and hence an automated optimisation routine cannot work. The multiple wedge method as used here avoids these problems, and so can be used successfully with SGA to identify the critical surface. This is demonstrated here.*

1 INTRODUCTION

The Simple Genetic Algorithm (SGA) has been used by a number of workers for the task of locating the critical slip surface in slope stability analysis (Goh 1999; Goh 2000; McCombie and Wilkinson 2002). In the case of circular failure surfaces the optimisation problem that this represents is relatively straightforward, the search space being only three dimensional (circle centre co-ordinates and radius), whereas for non-circular surfaces the dimensions of the search space are equal to the number of lines defining the surface, if the surface begins and ends at the ground surface. This represents a much more serious optimisation problem. The use of SGA in tackling this problem is well described by Goh (1999), in which the method of analysis used is as described by Donald and Giam (1989). SGA has also been used by Zolfaghari et al (2005), using the method of Morgenstern and Price (1965).

Methods of stability analysis for non-circular failure surfaces tend to be either methods of slices (e.g. Janbu, 1957, Morgenstern and Price, 1965, Spencer, 1967) or methods of wedges (e.g. Sarma, 1979, Donald and Giam, 1989), though composite analyses using finite element analysis have been developed (e.g. Krahn, 2003). Rather than simply determining a factor of safety against failure, which is what is needed in practice, most methods attempt a so-called 'rigorous' analysis which demonstrates equilibrium of both forces and moments, requiring substantial assumptions to be made before obtaining the solution, which may be critically dependent upon these assumptions. These assumptions usually relate to the degree of shear force mobilisation on slice or wedge boundaries, and normal stress distribution. All methods require an estimate to be made of the shape of the slip surface, but multiple wedge methods also require the orientation of inter-wedge boundaries to be determined.

The method presented by McCombie (2009) avoids both of these problems by not attempting a full 'rigorous' solution, simply using force equilibrium to obtain a factor of safety, and by considering kinematics to determine the orientation of inter-wedge boundaries. This makes the method particularly well suited for optimisation routines. The fact that interwedge boundary angles are generated rather than specified reduces the number of unknowns by approximately one third. Because there is no need to explore different assumptions to allow 'rigorous' analysis, the process is very much more straightforward, especially for significantly non-uniform soil masses or loading, for which the routinely used assumptions are particularly likely to be seriously wrong. In adopting this approach, the method attempts to reflect the unsurpassed elegance of Bishop's simplified method, which is very widely used in practice because its solution is not crucially dependent upon assumptions made on the basis of very little information (if any at all). Bishop's method focuses on obtaining the required factor of safety, via a very simple assumption about the overall effect of the unknowns for circular failure surfaces which has been justified by successful use in practice over very many years. To be precise, the determination of the shear strength on the slip surface is not dependent upon knowledge of the magnitude or distribution of the normal forces on the sides of the slices; the effect of shear forces on individual slices was discounted by Bishop in the simplified method, though the actual consequence of this simplification was to neglect the overall effect of shear forces for all the slices, a much more robust assumption which only becomes seriously erroneous if there is heavy concentrated loading. Shear force and normal force distribution on the sides of the slices are simply accepted as unknown, and not necessary for obtaining the factor of safety.

2 METHOD OF ANALYSIS

A full description of the method of analysis used in this work is presented in McCombie (2009). The method uses kinematics to determine the orientation of inter-wedge boundaries, as illustrated for a four wedge mechanism in Figure 1. Hodographs are used to describe relative motions (Donald and Chen, 1997). The motions of each of wedge relative to the underlying ground are shown as vectors ($V_{1..4}$). The hodograph on the left in dotted lines shows the case for no dilation. If the displacement along the shear surface is uniform, as would be expected in the simplest situations, then the lengths of each of these vectors will be the same. The vectors ($V_{12..34}$) then show the relative displacement between each wedge. The orientation of the wedge boundaries must correspond to the directions of these vectors, which will bisect the angle formed by the bases of the wedges. If dilation occurs as the displacement takes place and the shear strength is mobilised, then the directions of the displacements will not be the same as the orientations of the surfaces, as shown in the hodograph on the right, in which ψ denotes the angle of dilation. It can be seen that the inter-wedge boundaries must be rotated backwards by twice the angle of dilation.

It may be noted that in a classic plasticity approach, the angle of dilation would be set at the angle of friction, and displacements along the shear surface could probably no longer be uniform but would become progressively larger towards the toe, so that peak friction might be developed at the toe well before it is developed at the crest. Such an approach could in any case only apply if the displacements were very small, because once they reached an order similar to the size of the particles the angle of dilation would have dropped to a function of the difference between critical state and peak angles of friction (0.8 is used here, after Bolton, 1986). The need for perfectly plastic behaviour would therefore never arise, as the approach would already have failed to represent the real mechanisms of slope instability being considered.

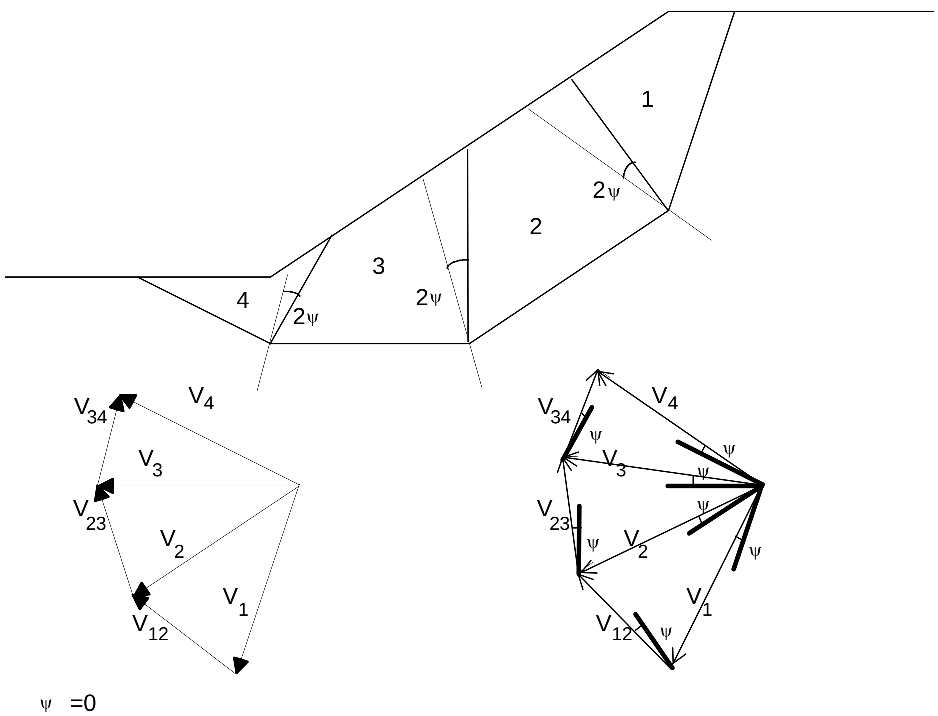

Fig. 1. Hodographs for a four wedge mechanism. The thick lines in the hodograph on the right denote the orientations of the surfaces.

This approach works well until a situation develops in which wedge boundaries overlap (Figure 2). It is not sufficient to simply disallow such configurations, however, as there may be constraints on the position of the slip surface which force this situation to arise. It could be that the wedge boundaries would indeed intersect below the ground surface, perhaps resulting in a block of soil which rotates as the sliding wedges pass beneath it. However, it seems more likely that the displacements along the shear surface will not be uniform, which is what must happen if the boundaries are rotated to remove overlaps. This implies that the full shear strength cannot be mobilised everywhere along the shear surface at the same time unless the soil is perfectly plastic.

The method uses the relative displacements on the slip surface and inter-wedge boundaries to determine the shear force mobilised. The degree of mobilisation of both frictional and cohesive shear strength can be simply related to the displacement on the surface by reference to shear box testing, for example, and modeled in as simple or as complex a way as the engineer requires. This would be expected to include some post-peak behaviour, so that the shear strength reduces as deformations become large. For a simple analysis in which the peak strength is not reached anywhere, a reliable factor of safety will be obtained using a simple linear development of strength with displacement up to peak, and even a simple representation of post-peak behaviour will give an indication of what might happen if the slope is marginally stable or will fail.

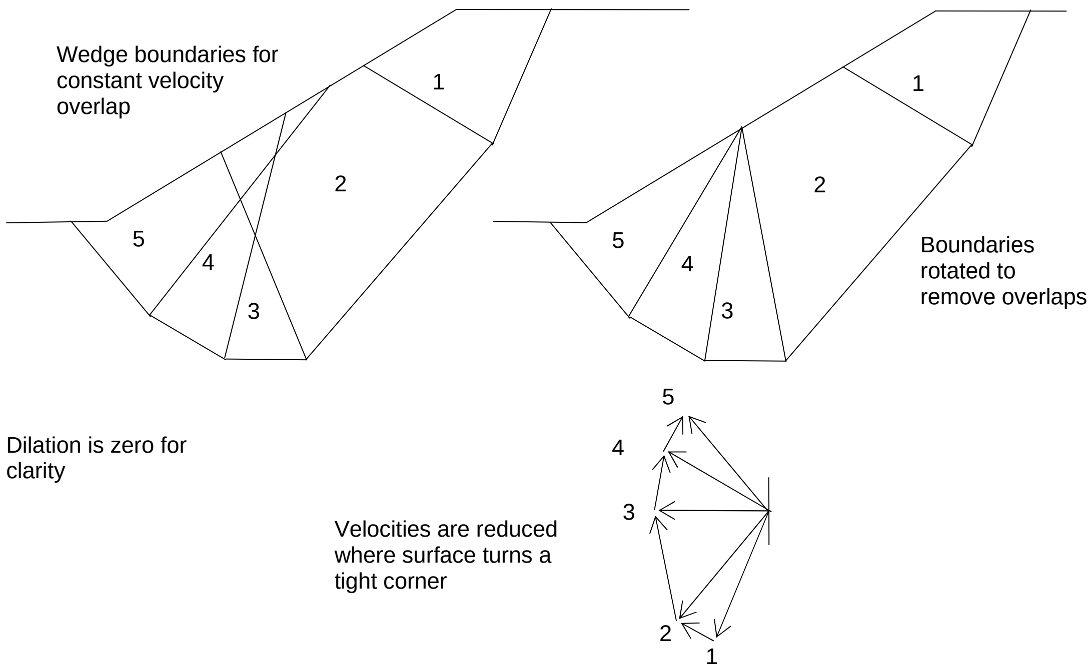

Fig. 2. Overlapping inter-wedge boundaries.

The analysis is carried out by the application of increments of displacement until force equilibrium can be obtained for all the wedges. Each wedge is analysed in turn. The frictional shear force is related to the effective normal force on a boundary by the mobilised angle of friction, as shown in figure 3.

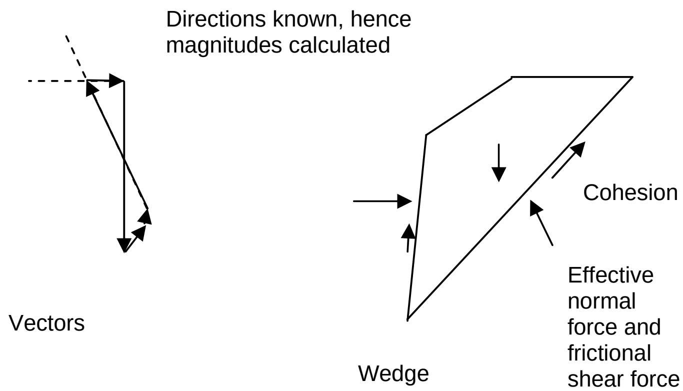

Fig. 3. Determination of forces on the first wedge.

Because the directions of the combined effective normal and shear forces are known, the unknown magnitudes on the two boundaries can be determined through the vector diagram shown, which is equivalent to solving two simultaneous equations. For each subsequent wedge there are similarly two unknown forces, until at the final wedge there needs to be an out-of-balance force applied to maintain equilibrium. The applied displacements, and hence mobilisation of friction, are increased until this out-of-balance force becomes zero. If the displacement along the boundary is not uniform, some parts of the slip surface may have passed peak strength while other parts have not yet reached peak strength. As more and more of the surface passes the displacement for peak strength, it may become impossible to obtain this zero out-of-balance force, and a factor of safety of less than one is obtained by considering the forces at the displacement which attained the minimum out-of-balance force.

It is possible that equilibrium cannot be achieved at all, even though it might have been could the strength be mobilised everywhere at once. This effect is generally not considered in slope stability analysis; the term 'progressive failure' is normally taken to apply to large slopes where compression of the soil results in the non-uniform displacements that lead to uneven mobilisation of shear strength. The possibility that the geometry of the shear surface alone may account for uneven mobilisation of shear strength has probably not been recognised. This effect turns out to be sufficiently important for the failure of Carsington Embankment (see below) that it can completely account for the observed failure, without any need to call on compression of the soil. Given that non-circular failures are most likely to occur in situations which are significantly non-homogenous, the effect of these conditions in forcing the failure surface away from an approximation of a circle is probably very important. The ability to account for these effects must be an important factor in the production of an optimisation routine aimed at finding the critical failure surface.

There is some dependence in this method on the initial conditions in the slope, in that the development of shear resistance is related to the displacement on a surface. For example, soil towards the toe of an excavated slope would be expected to have a greater part of the shear strength already mobilised due to some kind of K_o condition, whereas an overconsolidated clay fill would probably hold substantial pore water suction, and so the degree to which friction is mobilised would start small and slowly increase as pore water pressures increase towards equilibrium values.

3 IMPLEMENTATION IN TENSARSLOPE

The method described above has been implemented in an experimental version of TensarSlope, the reinforced soil slope stability analysis program belonging to Tensar International. Details are given in McCombie (2009). The programming environment used is Codegear (formally Borland) Delphi; the versatility of data types and object orientation were particularly useful in programming this method and the SGA. The program already allowed the definition of complex problems, with soils defined in zones rather than in layers, for example, and scope for applying surcharge loads, horizontal line loads, soil reinforcement, complex pore water pressure regimes, and external standing water. However, the definition of soil strength was altered to allow use of both peak and critical state angles of friction so that the angle of dilation could be deduced. A control was provided to allow the user to view the detailed results for each individual wedge once an analysis was completed, and to export them to a comma separated value file to facilitate checking.

A new form was developed to contain the processes necessary for the genetic algorithm, including entering values for the bounds of the problem, and the parameters for the optimisation. One option is produce each generation of chromosomes by a random process, rather than following the genetic algorithm. This is still far more efficient than using completely random generation of point co-ordinates, because the rational system described below is used to generate a surface from each chromosome, but it allowed the optimisation process itself to be evaluated.

4 SIMPLE GENETIC ALGORITHM

The SGA used here employs binary coding to represent the variables used to generate each slip surface to be analysed. A simple way to generate a slip surface would be to encode the co-ordinates of each wedge junction, but most of the surfaces thus produced would not be feasible. The approach used has therefore been to use the binary chromosome to represent a series of integers, each of which define where a value lies within a range of feasible values. The surface is generated from one end to the other, and the range of feasible values takes into account the geometry of the surface generated thus far. The encoding thus does not define the co-ordinates directly, but rather the sequence of steps by which they must be generated. This does result in the number of wedges used being fixed. However, in the method of analysis adopted, if two wedges have their bases aligned then there will be no relative movement between them and they will behave as a single wedge. It is therefore possible for the search to in effect reduce the number of wedges if this optimises the solution.

The following information is therefore used to generate a surface:

- minimum and maximum x co-ordinates for the toe of the slip surface;
- minimum and maximum x co-ordinates for the crest of the slip surface;
- minimum and maximum angles for the slip surface at the toe;
- minimum and maximum angles for the slip surface at the crest.

These values apply to every surface, and would provide all the information required for a two wedge mechanism. For each individual surface, a single number is used to define where each actual x co-ordinate or angle lies within its range of possible values. This might be thought of as a real number between 0 and 1, but is actually encoded as a binary value between 0 and 255. This reduces the precision of the solution compared with using a floating point value, but given that the range is restricted to realistic possibilities rather than encompassing all possible values, this precision is more than sufficient. In real coded algorithms, crossover would have to operate at data boundaries, and changes in actual numbers can only occur through mutation. The binary coded approach adopted here has the advantage of being able to produce substantial changes through splitting an integer encoding in two at crossover, which far outweighs the potential for more precise definition of numbers through a large number of random mutations in real coded algorithms.

The positions of the intervening wedge junctions are defined by two values at each junction, which are each encoded as binary values in the same way. One value is used to determine the x co-ordinate. For each node, a realistic range of values may be set using whatever rule is considered appropriate, but in this instance a lower bound has been set at one fifth of an equal share of the horizontal distance to the end of the surface, and the upper bound is set at 1.5 times this equal share, up to a maximum of 0.8 times the distance to the end of the surface.

The other value is used to define the angle of the wedge starting from that point. If the lower bound for the angle is the angle of the base of the previous wedge, then a concave surface is guaranteed; this might not be required, in which case a different lower bound may be defined. Given a requirement for a concave failure surface, an upper bound for the angle can be defined in terms of the angle to be achieved by the end of the failure surface; if this were achieved before the last point, then remaining points would have to lie on a straight line, and the range of permissible angles would be reduced to zero. The angle may also be constrained by the need to reach the end of the slope within the required range of x co-ordinates. As the last line on the surface has its position defined already, the penultimate point may be found by the intersection of this line with the previous line. Checks may be made that the surface lies beneath the ground at all points along its length, but provided any inappropriate surface is given a high value of factor of safety, then the optimisation will succeed.

The integer values thus required are combined into a chromosome on which the genetic algorithm operates. The conventional way of encoding for the SGA has all the data represented in a single binary string. This could take the form of examples a) or b) in Fig. 4.

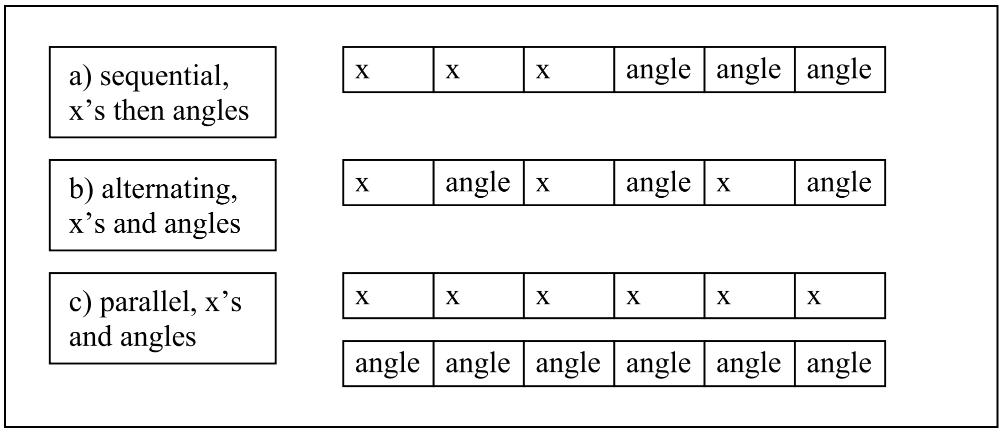

Fig. 4. Encoding the surface.

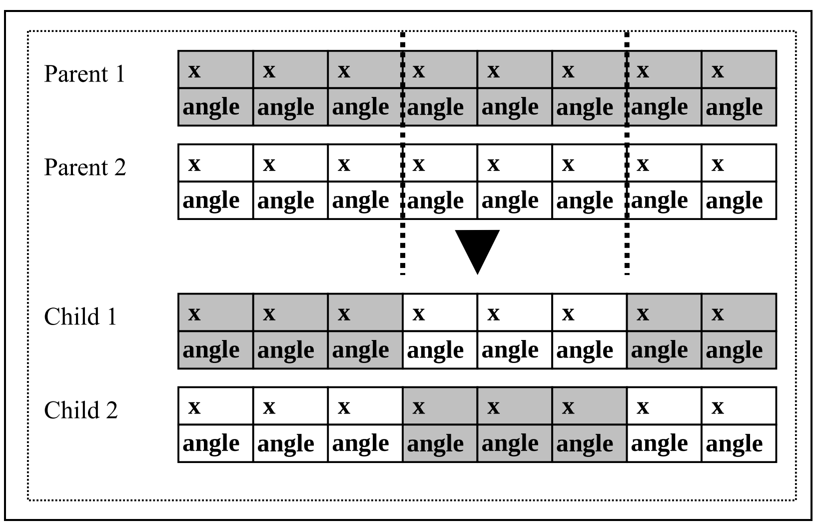

Fig. 5. Two point crossover for parallel chromosomes.

In the conventional approach, all the genetic operations take place on single chromosomes, with each chromosome containing all the information required to define the member of the population, or failure surface. In case a), a genetic process could result in either the angles for a part of a surface being swapped, or the x co-ordinates, but never both angles and x co-ordinates for the same section of slip surface. This constrain the power of the algorithm to generate new solutions or optimise existing solutions. In case b), this problem is removed, but there is still a possibility of separating an x-coordinate from its associated angle, and it seems more appropriate to use two parallel chromosomes, representing each set of data. These chromosomes are then manipulated in parallel, as if they were fixed together so that they are cut and exchanged at the same places (figure 5). This method has been used in the current work, and has the additional advantage of being relatively straightforward to program, as the two chromosomes are read in parallel as the surface is generated.

5 DEMONSTRATION

The analysis method and the optimisation process have been evaluated using the example of the Carsington Embankment failure, using parameters from work by Potts et al. (1990), as described in McCombie (2009). An image from the computer screen is shown in Fig. 6 after 66 generations of a population of 40, which took under two minutes. The entire population of surfaces is shown, indicating the scope for significantly divergent solutions to be generated even after a large number of generations.

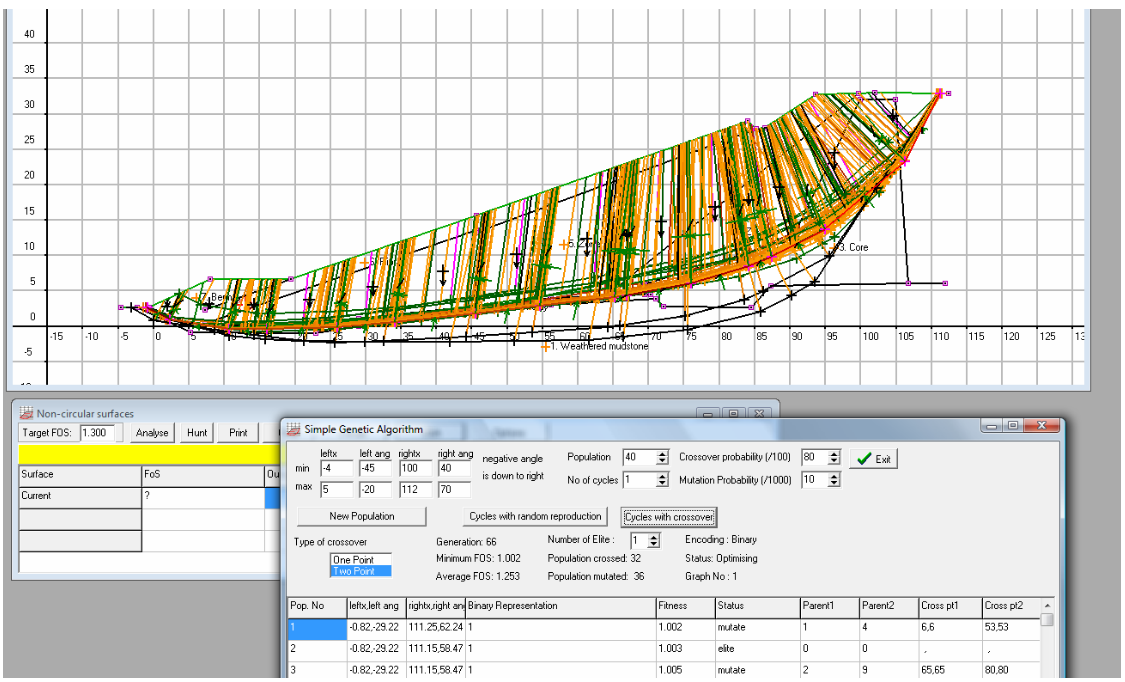

Fig. 6. SGA analysis of Carsington Embankment failure.

A second screen shot is shown in Fig. 7, following a final stepwise adjustment of the slip surface position, which gives a slight improvement in the resolution for this example in which the weak layer is very thin. This figure also allows the final surface to be seen clearly on its own.

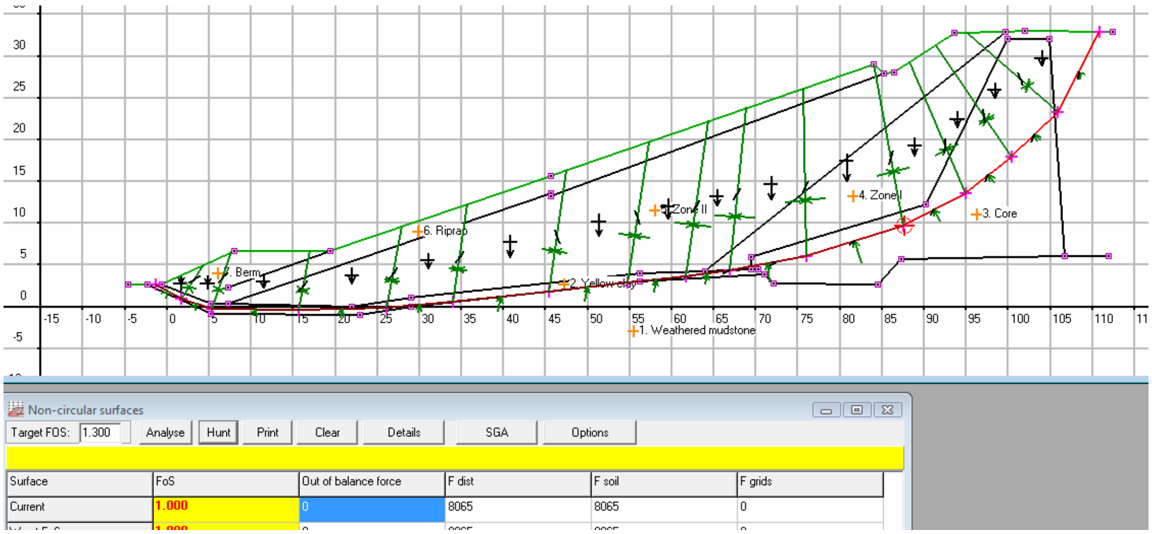

Fig. 7. Final refinement of the analysis.

The critical mechanism as found by the SGA corresponds closely to the pattern of vectors from the finite element analysis, as described by McCombie (2009). The resulting factor of safety of 1.000 found in an analysis carried out for this paper also corresponds with the result reported by Potts et al. (1990), and the fact that the embankment actually failed on a surface very close to that described above. Potts et al. (1990) reported that limit equilibrium methods available at the time could not achieve this result.

6 CONCLUSIONS

The SGA has been shown to work well in finding an accurate optimum solution for a relatively difficult problem, when used with the multiple wedge method as described by McCombie (2009). Importantly, both the analytical procedure and the optimisation routine are easy and quick to use for routine design work.

ACKNOWLEDGEMENT

The author wishes to express his particular thanks to Tensar International Limited for their permission to use the program WinSlope, later TensarSlope, in the work reported here.

REFERENCES

Bolton, M.D. (1986). "The strength and dilatancy of sands". Géotechnique Vol. 36(1), 65-78.

Donald, I.B. and Chen, Z. (1999). "Modern Wedge Methods of Slope Analysis – Their Power and Their Versatility". Proceedings of 2nd Int. Conf. on Landslides, Slope Stability & The Safety of Infra-Structures, Singapore, 13-24.

Donald, I.B. and Giam, P.S.K. (1989), "Improved comprehensive limit equilibrium stability analysis". Department of Civil Engineering Report No. 1/1989, Monash University, Melbourne, Australia.

Goh, A.T.C. (1999). "Genetic algorithm search for critical slip surface in multiple wedge stability analysis". Can. Geotech. J., Vol. 36, 382-391.

Goh, A.T.C. (2000). "Search for critical slip circle using genetic algorithms". Civil Eng, and Env. Syst., Vol. 17, 181-211.

Krahn, J. (2003). "The 2001 R.M. Hardy Lecture: The limits of limit equilibrium analyses". Can. Geotech. J., Vol. 40, 643-660.

Janbu, N. (1957). "Earth pressures and bearing capacity calculations by generalized procedure of slices ". Proceedings 4th Conference of ISSMFE, London, Vol. 2, 207-212.

McCombie, P. (2009). "Displacement based multiple wedge slope stability analysis". Computers and Geotechnics, Vol. 36(1-2). 332-341.

McCombie, P. and Wilkinson, P. (2002). "The use of the simple genetic algorithm in finding the critical factor of safety in slope stability analysis". Computers and Geotechnics, Vol. 29(8), 699-714.

Morgenstern, N.R., and Price, V.E. (1965). "The analysis of the stability of general slip surfaces". Géotechnique, Vol. 15(1), 79-93.

Potts, D.M., Dounias, G.T. & Vaughan, P.R. (1990). "Finite element analysis of progressive failure of Carsington embankment". Géotechnique Vol. 40 (1), 79-101.

Sarma, S.K. (1979). "Stability analysis of embankments and slopes". J. Geotech. Engng Div. Am. Soc. Civ. Engrs., Vol. 105(GT12), 1511-1524.

Spencer, E. (1967). "A method of analysis of embankments assuming parallel interslice forces ". Géotechnique, Vol. 17(1), 11-26.

Zolfaghari AR, Heath AC, McCombie PF. (2005). "Simple genetic algorithm search for critical non-circular failure surface in slope stability analysis". Computers and Geotechnics, Vol. 32(3), 139-152.