DESIGN OF OPTIMAL FLOOD PROTECTION EMBANKMENT: COMBINING DIFFERENT FORMS OF ARTIFICIAL INTELLIGENCE

R. van der Meij
Deltares, Delft, The Netherlands
Raymond.vanderMeij@Deltares.nl

Delft University of Technology, Department of Civil Engineering, Section of GeoEngineering, Delft, The Netherlands

ABSTRACT: *This paper describes a way the most essential design criteria of the VTV have been transferred into artificial neural networks. These design criteria are stability according to Bishop and Van, internal erosion by Sellmeijer and horizontal deformations with FEM. The artificial neural networks are used within a genetic algorithm to automatically generate multiple good flood protection embankment designs in a few seconds time. The designs adhere to the criteria, maximize safety and minimize cost at the same time.*

1 INTRODUCTION

Designing flood protection embankments is very difficult. The Dutch "Voorschrift Toetsen op Veiligheid" (VTV) is a document that describes the rules a flood protection embankment has to adhere to. This document is almost five hundred pages and contains much technical information. On top of this, knowledge of calculation techniques is required to implement the rules given in the VTV.

It is difficult to make changes to, or a construction nearby an embankment because the design of such an embankment is very difficult. Planners often use rules of thumb, but they do not know how large the impact of the VTV will be. Having a simple rule regarding the width of an embankment often causes large problems when the specialist finds out the rule of thumb is too optimistic. On the other hand, it is also possible that good alternatives are discarded in the beginning because they do not seem possible, even though they are.

This paper describes a way the most essential design criteria of the VTV (stability according to Bishop (1955), stability with uplift conditions according to Van (2001), internal erosion by Sellmeijer (1988), horizontal deformations with FEM (Van der Meij 2008)) have been transferred into different artificial neural networks. These different artificial neural networks are used within a genetic algorithm to automatically generate multiple flood protection embankment designs in a few seconds time. The designs adhere to the mentioned failure mechanisms, maximize safety and minimize cost at the same time.

This tool can easily be used by planners in an early stage of a construction project because it is fast and easy to use. Because conventional computational techniques are too time consuming, it is only possible to have such a design tool using artificial intelligence techniques. In the near future, extensions to this design method will be made, for example to design emergency measures for flood protection embankments.

2 CONCEPTUAL DESIGN OF THE ALGORITHM

In a safety evaluation, input data is put into a conceptual model in order to make a prediction. Making a design is the inverse of such an evaluation. Given a desired output, the optimal input data needs to be found. Two difficulties arise: the inverse of the conceptual model is not known, and the search space is extremely large due to the curse of dimension (Bellman 1957). This makes finding the optimal design practically impossible.

The considered design - a flood protection embankment - is so complicated that it is not realistic to find a relationship to go directly from the design criteria to the optimal design. This means an iterative procedure needs to be used in order to make the inverse analysis.

The designer normally performs this iterative procedure. Because it is so time consuming to use the conceptual models, only two or three iteration steps are made. The conceptual models need to be simplified in order to have a proper iterative design procedure. In this paper, the simplification of the conceptual model happens through the training of an Artificial Neural Network (ANN) (Bishop, 1995). This is possible, because large amounts of input and output data can be prepared by sequentially running all the possible cases through the available conceptual model. In this case, the ANN is only used as a regression tool.

What's left is a good search procedure to find the optimum design. A Genetic Algorithm (GA) (Barricelli, Nils Aall 1957) is a very good tool to search the optimum in a very complex multi-dimensional search area with many local maxima and minima. The combination of such a Genetic Algorithm to find the optimum with the Neural Networks representing the conceptual models is a good combination to let the computer calculate the optimal design of a flood protection embankment.

3 COST FUNCTIONS USING NEURAL NETWORKS

As explained in the previous section, the conventional conceptual models take too much time to evaluate in an optimization procedure. Several ANNs are available to replace the conceptual models and are discussed in this section together with some analytical formulas. These networks and functions form the basis of the cost function in the optimization procedure.

Stability Bishop and Van

Two types of neural networks are available to perform the stability analysis. One is based on a very large batch of calculations and the other is based on a large database with real cases from the area called the "Alblasserwaard", east of Rotterdam.

The database based on the real cases determines the safety factor based on two failure mechanisms: Bishop's method and Van's method. Both networks have five input parameters: the average slope, the height of the embankment, the water height difference, the soil thickness and the factor of resistance against uplift of the hinterland. This database based on the real cases has two disadvantages. First, it makes the genetic algorithm only applicable in a similar area. Second, it limits the boundaries where the fitness can be determined to the bandwidth available in the Alblasserwaard. The artificial neural network based on the large batch of calculations has a much broader bandwidth in which it is valid.

The database based the batch of calculations does not have such strict maximum and minimum height requirements and the uplift safety is not necessary to use the network. A disadvantage of the network based on the calculations is that only Bishop's method is available.

For now, the database with the real cases for Van's method is used for the algorithm. For the case "Bishop", the ANN based on the large batch of calculations is used.

Internal erosion by Sellmeijer

For erosion underneath the dike, so called “piping”, only one artificial neural network is available. It is based on Sellmeijer’s method (Sellmeijer 1988) and contains four input parameters. Where “L” is the length of the base of the embankment, the following parameters are input for the network: HG/L, HS/L, log(KR/KG), log(KS/KG). Figure 2 shows the meaning of the variables. The output of the network is the critical slope of the water table under the embankment. If the true slope is steeper than the critical slope, the safety is not sufficient.

Horizontal deformations with FEM

An artificial neural network is available to predict horizontal deformations under a newly built embankment. The building and verification of the network is described in Van der Meij (2008).

Several neural networks are available. Some of them take into account the building history of the embankment. This information is not available in the genetic algorithm, therefore it is chosen to use a relatively simple network with only three input parameters: the height of the embankment, the stiffness of the subsoil and the thickness of the subsoil.

Other cost functions

The volume of the embankment is minimized by giving a cost for every cubic metre per metre embankment. As a future extension, it is easy to add a cost per square metre for buying ground outside the current profile. It is also possible to charge a cubic metre outside the current profile with a higher cost than a cubic metre within the current profile. This is a suggestion for the future development of the algorithm.

Two other possible future extensions are possible to design in a less conventional way. It is possible to add two variables to the embankment properties: the cohesion and the internal friction angle of the embankment. By giving a cost to a higher strength parameter, ground improvement can be taken into account. Alternatively, it is possible to allow for steep or vertical slopes with their respective cost.

If such extensions are made, one has to make sure the artificial neural networks that calculate the safety can deal with such designs. Consequently, such slopes are not permitted.

4 THE GENETIC ALGORITHM

This algorithm will have as input an array of variables representing the boundary conditions. Given these boundary conditions, the design of the levee needs to be optimized. The array of numbers that need to be optimized is called the “genome”. This section shows how the fitness of the embankment is calculated and how the fitness of the population will increase over time. The description of the embankment contains variables to be optimized. These variables are shown in this section together with the bandwidth by which they may vary. This bandwidth is usually imposed by the neural networks because they only function with a certain range.

The current variables in the description of the embankment are the height of the embankment (HK), the width of the crest (LK), the inner slope after the crest (M1), the height of the berm (HB), the length of the berm (LB) and the slope between the berm and the surface level (M2). They are drawn in figure 1.

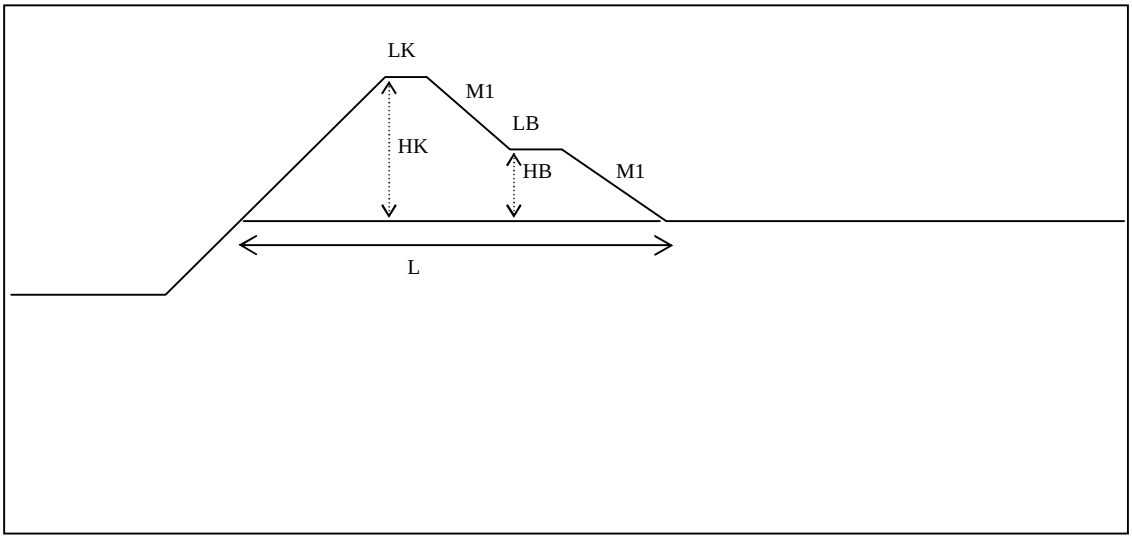

Fig. 1. Parameters to be optimized

The range in which the variables may vary depends on the range for which the calculation of the fitness function is valid. On top of that, a number of logical borders are set, like ‘the crest needs to be higher than the water table’ and ‘a berm may not be lower than the surface level’. There is also a maximum value for the width of the crest and the width of a berm to reduce the solution space and thereby reduce the computational costs. An overview of the variables with the maximum and minimum values is given in Table 1.

Table 1. Bandwidth of the parameters

Variable	*Unit*	*Minimum value*	*Maximum value*
HK	*metre*	*High water level*	*High water level +6 metre*
LK	*metre*	*3 metree*	*15 metre*
M1	*coefficient*	*1:10*	*1:2*
HB	*metre*	*Surface level*	*High water level +6 metre*
LB	*metre*	*0 metre*	*20 metre*
M2	*coefficient*	*1:10*	*1:2*

Boundary conditions

Boundary conditions are entered for the problem. The boundary values have been added to the embankment variables (the genome) in Figure 2. The variable names are clarified and the boundaries are given in Table 2. It has been assumed that behind the flood protection embankment, the surface consists of a soft soil layer with underneath different granular materials.

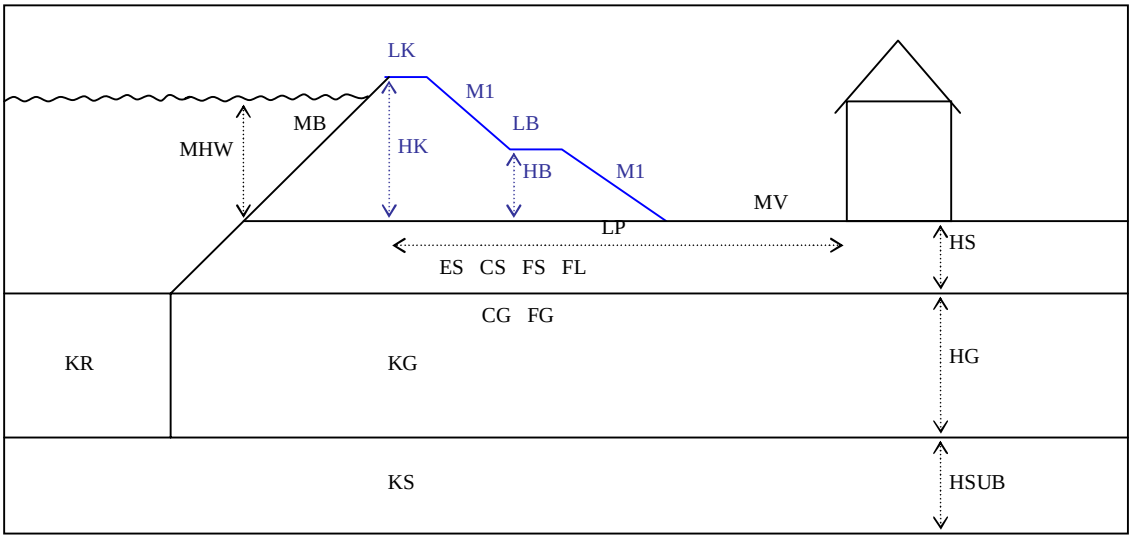

Fig. 2. Parameters to be optimized in combination with boundary values

The surface level, MV, is a reference level and is therefore not an input value. For clarity, it is given in the figure above and the table below.

Table 2. Bandwith of the boundary values

Variable	*Abbr.*	*Sample value*	*unit*	*max*	*Min*
Surface level	*MV*	*0.0*	*Metres above reference level*	*0*	*0*
Rep high water	*MHW*	*6.0*	*Metres above surface level*	*1*	*12*
Stiffness soft layer	*ES*	*1000*	*kPa*	*700*	*7000*
Cohesion soft layer	*CS*	*10.0*	*kPa*	*2*	*20*
Int. frict. angle soft layer	*FS*	*25.0*	*Degrees*	*15*	*25*
Height soft layer	*HS*	*8.0*	*Metre*	*3,6*	*11*
Uplift safety soft layer	*FL*	*1.2*	*coefficient*	*0.5*	*99*
Outer slope	*MB*	*2.5*	*coefficient*	*0*	*10*
Cohesion granular layer	*CG*	*3.0*	*kPa*	*0*	*10*
Int. Frict. granular layer	*FG*	*33.0*	*Degrees*	*20*	*40*
Thickness granular layer	*HG*	*10.0*	*Metre*	*1*	*50*
Perm. granular layer	*KG*	*3.0*	*Metre*	*0.001*	*1000*
Permeability river bed	*KR*	*3.0*	*Metre/day*	*0.001*	*1000*
Permeability sub layer	*KS*	*1.0*	*Metre/day*	*0.001*	*1000*
Thickness sub layer	*HS*	*20.0*	*Metre*	*1*	*50*
Distance to buildings	*LP*	*40.0*	*metre*	*5*	*999*

In addition, the β in Sellmeijer's formula is required and the minimal safety factor for the failure mechanisms according to Bishop and Van. For every failure mechanism, the minimum requirement needs to be entered.

Besides these geo-mechanical parameters, a set of parameters needs to be defined to optimize the search algorithm. For this algorithm, a mutation rate, the size of the elite and the population size need to be entered. The cross-over happens by using the first half of the first genome and the second half of the other. A mutation happens in a certain percentage of the cases when another random value for a property is chosen. Fighting happens through tournament selection. Elitism is used.

5 RESULTS

Many calculations under different boundary conditions have been run. The results consequently seem to give the optimal design in the solution space. Figure 3 shows both the boundary conditions and the results of such a run. Boundary conditions that are not displayed in the figure are the β in Sellmeijer's formula (0.15), the required safety factor of the embankment (1.13) and the minimal crest width (3 metre).

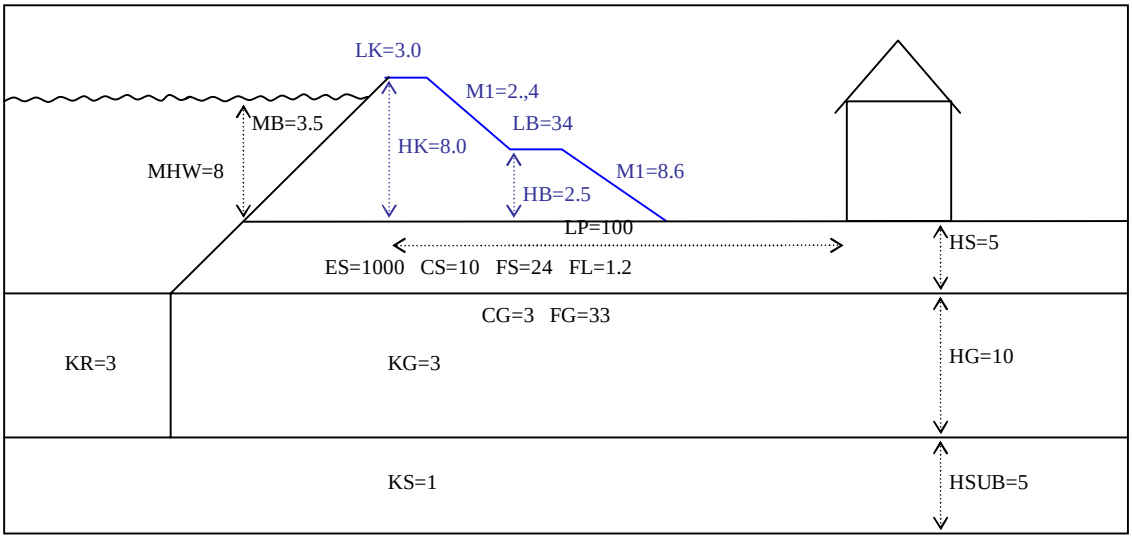

Fig. 3. Boundary conditions and results of example

The program results in a genome with properties that satisfy all the safety and environmental requirements and minimizes the volume of the embankment at the same time. The safety factors of the embankment are exactly 1.13 and the critical piping height is higher than the height difference of the water levels.

To test both the speed and the quality of the design an alternative search mechanisms has been programmed. The results of the genetic algorithm are compared to a hill climbing (HC) algorithm. The hill climbing algorithm makes steps of 1 metre. The begin point of the search is in the middle of the search space. The algorithm stops if no improvement in fitness is reached in an iteration.

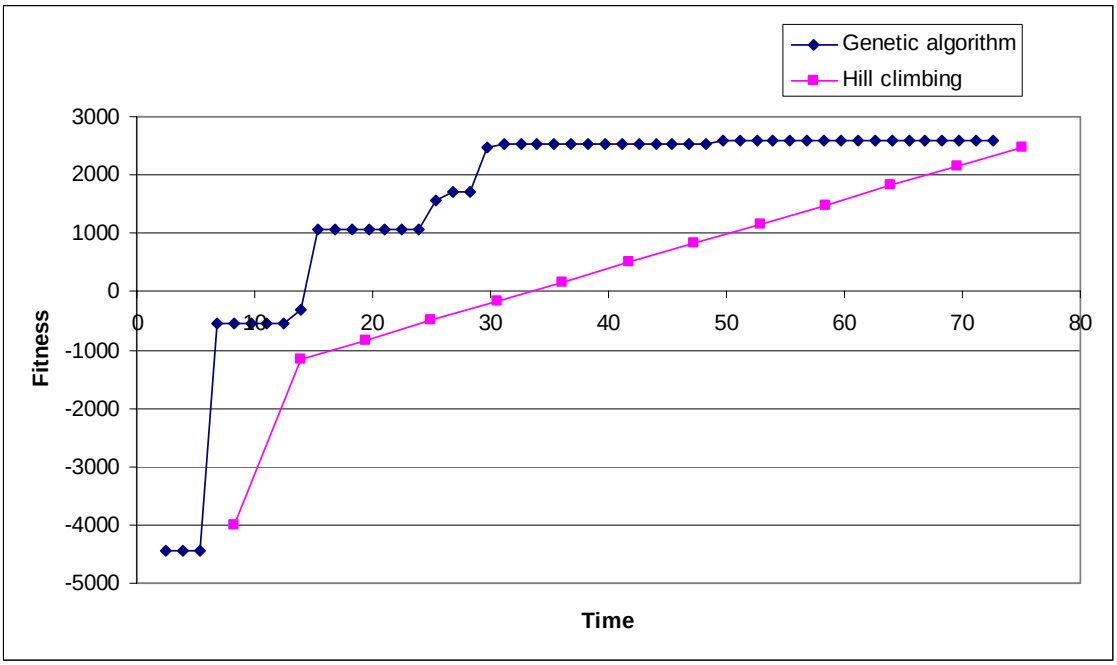

Fig. 4. Speed of the Genetic Algorithm verses a Hill Climbing algorithm

Search method	***Height crest***	***Width crest***	***Upper slope***	***Height of the berm***	***Width of the berm***	***Lower slope***
Genetic	*8.04*	*3.04*	*2.44*	*2.50*	*33.56*	*8.57*
Hillclimbing	*8.00*	*3.50*	*4.00*	*2.50*	*34.00*	*5.00*

Table 4.11 comparison of the designs with a GA and HC

The calculation time is much in the advantage of the GA. Within the GA, it is not known, though, when the optimal solution is found. Figure 4.7 shows the fitness of both algorithms is about the same in the end. Both algorithms have found the same minimum.

6 CONCLUSIONS AND RECOMMENDATIONS

The most important conclusion is that it is possible to combine different types of AI to make a quick design of a flood protection embankment. This case can be seen as a proof of concept to find an optimal design in a very limited amount of time. The design is relatively fast, and adheres to the boundary conditions supplied. Even though the case presented is relatively course (only a few mechanisms are checked and they have a relatively small bandwidth in which they function) it gives enough trust to built a specific tool that uses these techniques.

Best parameters for the Genetic Algorithm

Relevant parameters for the Genetic Algorithm are the number of generations, the population size, the mutation rate and the number of elites. Experience has learned that a relatively small population size can give enough diversity. A population of 50 individuals can contain sufficient information to have a population that has enough diversity to find good solutions quickly. With such a small group, it is better to have a relatively high mutation rate in order to have sufficient mutations to perfect the good solutions. A rate of 5 percent can be

sufficient but an even higher rate facilitates the convergence to the minimum. Even a rate of 15 percent does not cause too much disturbance in the convergence.

Experience shows that the different individuals in the elite usually represent the same geometry, so having a multiple elite does not lead to a number of different output designs. Having a relatively large elite does have the advantage that more often good pieces of DNA are combined with different strings. By trying different configurations, it can be concluded that for this case it is a good practice to use an elite of around ten percent of the total population.

With these parameters it takes about 20 generations to find a number of good solutions. In another 20 generations, these solutions are improved every now and then. In one minute of calculation time, a number of good design can be presented.

GA versus Hill Climbing

For final judgment regarding the hill climbing algorithm verses the genetic algorithm, the hill climbing algorithm needs to be optimized. It is unfair to judge the hill climbing algorithm the way it is, because the performance depends much on the starting point and the search algorithm can be optimized. At the moment, it seems logical to start the search with a genetic algorithm and after a few generations continue with the hill climbing algorithm.

It seems like the search area is too complex for hill climbing. The calculation time is very long because there are too many neighboring nodes to check. Optimization of this method is possible but it will become more complex. The biggest problem with the hill climing algorithm is that it can converge to a local maximum. Taking a smaller step size does not overcome this problem.

An advantage of the hill climbing algorithm is that it is known when a local optimum is found. When using the genetic algorithm, it is not known if an optimum is found. This is yet another reason to combine the advantages of the hill climbing and the genetic algorithm.

In the case presented in this document, the solution space is too smooth to come up with multiple designs. The entire elite in the GA always converges to the same design. With a more complex search space, there is a chance more good designs will appear.

REFERENCES

Barricelli, Nils Aall (1957). "Symbiogenetic evolution processes realized by artificial methods". Methodos: 143–182

Bellman, R.E. (1957). Dynamic Programming. Princeton University Press, Princeton, NJ.

Bishop, C. M. (1995). Neural Networks for Pattern Recognition. Oxford University Press, ISBN 0-19-853864-2

Bishop, W. (1955). “The use of the slip circle in the stability analysis of slopes”. Geotechnique, Vol 5, 7-17.

Sellmeijer, J. B. (1988). On the mechanism of piping under impervious structures. Technical University, Delft, The Netherlands.

Van, M. A. (2001). “New approach for uplift induced slope failure”. XVth International Conference on Soil Mechanics and Geotechnical Engineering, Istanbul. 2285-2288

Van der Meij, R. (2008). “Predicting horizontal deformations under an embankment using an Artificial Neural Network”. The 12th International Conference of International Association for Computer Methods and Advances in Geomechanics (IACMAG), Goa, India 2008.