Failure analysis of soil slopes with advanced Bayesian networks

Longxue He[1], António Topa Gomes[2], Matteo Broggi[1], Michael Beer [1,3,4*]

[1] Institute for Risk and Reliability, Leibniz Universität Hannover, Germany
longxue@irz.uni-hannover.de; broggi@irz.uni-hannover.de; beer@irz.uni-hannover.de
[2] CONSTRUCT, Faculty of Engineering, University of Porto, Portugal,
atgomes@fe.up.pt
[3] Institute for Risk and Uncertainty, University of Liverpool, UK.
[4] CERSM, Tongji University, China.
[*] Corresponding author, e-mail: beer@irz.uni-hannover.de

Abstract

To prevent catastrophic consequences of slope failure, it can be effective to have in advance a good understanding of the effect of both, internal and external triggering-factors on the slope stability. Herein we present an application of advanced Bayesian networks for solving geotechnical problems. A model of soil slopes is constructed to predict the probability of slope failure and analyze the influence of the induced-factors on the results. The paper explains the theoretical background of enhanced Bayesian networks, able to cope with continuous input parameters, and Credal networks, specially used for incomplete input information. Two geotechnical examples are implemented to demonstrate the feasibility and predictive effectiveness of advanced Bayesian networks. The ability of BNs to deal with the prediction of slope failure is discussed as well. The paper also evaluates the influence of several geotechnical parameters. Besides, it discusses how the different types of BNs contribute for assessing the stability of real slopes, and how new information could be introduced and updated in the analysis.

Keywords

Failure probability, Slope stability, Water table, Drainage, Advanced Bayesian Networks

1 Introduction

Slope failure are a potential catastrophic threat by leading to casualties and economic loss in many areas around the world. Therefore, the slope stability problem, as a classical research topic, has attracted much attention in geotechnical engineering [1]. A slope failure event may be triggered by miscellaneous factors such as geotechnical factors, rainstorm, earthquakes, anthropogenic activity and so on. Water plays a significant role in the process, affecting the slope stability. Kristo et al. [2] demonstrated the increasing rain intensity had a detrimental influence on the slope stability. Also, the level of water table has a negative correlation with the factor of safety. Otherwise, soil properties and the presence or absence of vegetation can also potentially affect the slope stability [3, 4]. Furthermore, it is pivotal for decision-makers to achieve the information which the key failure-inducing factors are more sensitive to destabilizing the slope in order to avoid the highly economical and life loss.

Due to the unavoidable uncertainties existing in vague environmental condition, varying soil properties as well as insufficient information affecting the slope failure, the probabilistic method plays an important role in the estimation of the probability of failure for slopes [5, 6]. Traditional Limit equilibrium methods are normally used to analyze the stability of slopes, and the different shapes of potential failure surface are defined in advance to compute the factors of safety. Considering the most critical slip surface regarding the slope stability, the probability of failure for the slopes can be computed with this response surface.

The common approach is to model probabilistic slope stability as the system reliability problems. Various attempts have been applied in calculating the failure probability. For instance, slope stability problems associated with Structural Reliability Methods (SRMs) have been conducted by means of first-order reliability method (FORM) [7] and simulation approaches, such as Monte Carlo Simulation [8], Importance Sampling [9] and Subset Simulation [10]. These studies demonstrated the feasibility of structural reliability analysis for computing the probability of slope failure in geotechnical engineering. Artificial neural networks also have been adopted to predict the stability of slopes with geometric or geological data, influential factors [11, 12]. However, this approach is not good at quantifying the uncertainty and characterizing the impact of individual risk factors on the slope stability using information updating.

Bayesian Networks (BNs), as the causal probabilistic models, have been developed and successfully applied to natural hazards, safety, and reliability engineering for over two decades since their first introduction by Pearl [13]. Compared to the aforementioned numerical tools, BNs carry advantages over other available methods to calculate the probability of slope failure and identify the important factors regarding a given structure. In particular, they show the following advantages:

- Simple graphical visualization. The failure of a slope can be affected by geo-environmental parameters, weather condition, natural hazards (e.g. earthquakes and storm) as well as human activities. BNs can not only integrate these elements into a rigorous framework but provide a visual cause-effect relation among events in a graphical model. In particular, BNs help decision makers and even non-expert without a strong background in geotechnical engineering to gain a good understanding of the failure mechanisms. For a detailed overview on how to construct a graphical framework for risk assessment of rock-fall hazard with a BN model, see Straub [14].
- Uncertainty quantification. BNs are developed successfully to capture the uncertainties affecting the problem and benefit from the capability of the forward and backward propagation of probabilities according to the axioms of Bayesian probability theory [15].
- Information update from new observation. Updating of the event probabilities in BNs can be efficiently performed in near-real-time by mean of Bayesian updating to respect the information carried by the new observation. Thanks to this, the BN model can provide the decision makers with up-to-date information on the slope failure mechanisms as soon as new evidence is presented.

Traditional BNs (i.e., mainly discrete probability values and binary event are considered) have been already extensively employed to analyse slope stability [16-18]. Nevertheless, the slope stability problem is clearly influenced by both discrete events and continuous variables, thus it is impractical to obtain discrete probabilities of all the factors affecting a slope. Moreover, traditional BNs are precise probabilistic model, which fail to solve geotechnical problems with scarce information. Based upon this context, an extended and robust model: the advanced BNs including enhanced Bayesian Networks (eBNs) and Credal Networks (CNs), is proposed to deal with the geotechnical problems.

The main purpose of this work is to present how to estimate the failure probabilities of slopes, obtaining real-time results. Also, an attempted is made to capture the uncertainty by measuring the effect of variation of the induced-factors on the slope failure. Thus the paper is organised as follows: Section 2 introduces the methods of the advanced BNs, where a detailed review of eBNs and CNs is presented. Two examples are employed in Section 3 and 4 to evaluate the feasibility of models. We present how to build the failure analysis model for the slopes. We investigate two different failure types of slopes in a graphical model and combine the BNs with neural networks. Besides, the structure of CN of a slope is also presented. The final part summarizes the relevant results.

2 Methodology

2.1 Bayesian Networks

BNs, also known as Bayesian belief networks or causal networks, originate from artificial intelligence and statistics. They were developed as a powerful modeling tool for decision support and quantification of uncertainties, especially for low probability events. They have been applied to risk analysis in many studies since 2001 [19].

In a nutshell, a BN (see Fig. 1) is a directed acyclic graph, in which a set of variables are represented by nodes. The relation between each node is represented in terms of parent-child and linked by an arrow, denoting the conditional dependencies between these variables. Conditional Probability Tables (CPTs) are attached to each node and consider all the possible states of a variable. Then, the probabilities of the nodes are determined by marginalization calculation of the joint probability. The joint probability is the function of all the random variables in BNs. For any BN, it can be given mathematically by a product of the CPTs entries,

$$P(X_1, X_2, .., X_i) = \prod_i^n P\big(X_i|pa(X_i)\big).$$

(1**Fehler! Textmarke nicht definiert.**)

where X_i={$X_1, \dots, X_n$} denote the nodes of the BN, $pa(X_i)$ are the set of parents of X_i, and $P\big(X_i|pa(X_i)\big)$ represent the entries of the CPTs. The effective methods for general inference in BNs can be accessed in literature [20] and it is also applicable for probability updating. For instance, in the case where evidence is assigned to an observed node X_k=e, this information will propagate through the prior probabilities to the posterior probabilities as follows,

$$P(X_i|e) = \frac{P(X_i,e)}{P(e)} = \frac{\prod_i^n P(X_i|pa(X_i),e)}{\sum_{X_i \backslash X_k} P(X_i,e)}.$$

(2**Fehler! Textmarke nicht definiert.**)

note that the joint distribution $P(X_i, e)$, obtained by using Eq. (1), associates with the evidence value e, and compute $P(e)$ from $P(X_i, e)$ by marginalizing out all the variables except the node X_k. If a node with no children has no associated evidence, it is called “barren node”, meaning that the conditional probability is useless for the calculation of the marginal probabilities of non-barren nodes.

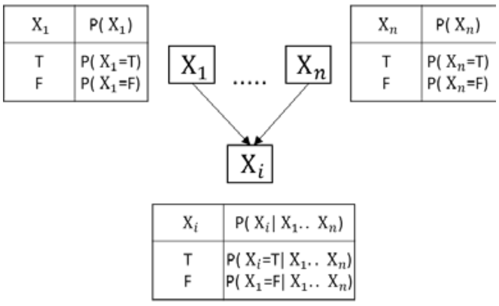


Fig. 1 A simple graph of a general BN (T=True; F=False)

In general, as for the ability of belief propagation in the network, marginal posterior probabilities of the query nodes can be achieved through both top to bottom and inverse reasoning by means of the inference algorithms, including exact algorithms and approximate algorithms. In comparison to approximate algorithms, exact algorithms, which are suitable for computing discrete BNs, are guaranteed to gain correct answers and hence, it is a more robust computational method. In case of continuous variables in a BN, however, given the difficulty of defining the prior probability distributions as the discrete form, unavoidably impeding the application of BNs for practical purposes.

BNs consisting of discrete and continuous variables are referred to as hybrid BNs. With consideration of exact algorithms, there are three special approaches for extending discrete BNs to continuous BNs or hybrid BNs. The first is to restrict continuous nodes to Gaussian random variables while allowing them to link only towards their non-discrete children. The second method is to define the continuous nodes as a mixture of truncated exponential distributions (MTEs), which is a generalization allowing to approximate any distribution function, but still requires further scrutiny [21]. The final methodology is eBNs, implemented by joining BNs with SRMs, and was successfully applied in risk and reliability analysis by Straub and Kiureghian [22]. An introduction to this method is given in detail in the following section.

2.2 Enhanced Bayesian Networks

Enhanced BNs approach [23] is to combine structural reliability methods with BNs, where continuous nodes can be involved in the BNs and removed with SRMs. With this model, exact inference algorithms can be conducted for a BN including both discrete and continuous nodes.

In a structural reliability problem, the outcome domain of an event, determined by a set of continuous random variables with known distributions, can be divided into failure and safe region by the relevant limit state functions. The failure probability of an event is the integral of the probability density function in the failure domain. In light of this, for an eBNs, the continuous nodes must have at least an offspring, which is a discrete node defined as a domain in the outcome space of these continuous nodes. That is, the continuous nodes should meet the requirement of well-established SRMs, and it is the key condition for using eBNs approach. Then, all the continuous nodes can be removed from eBNs according to node elimination algorithm [23]. Thus hybrid BNs are reduced to discrete BNs.

An example of computation of the total probability of an eBN and the process of node elimination is described by Eq. (3) to Eq. (5) for the simple case represented by Fig. 2. From Eq. (1), the joint probability of all the nodes for the eBN can be written as:

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_4|X_2, X_3)f(X_3)f(X_2|X_1)$$

(3**Fehler! Textmarke nicht definiert.**)

in which $P(X_1)$ and $P(X_4|X_2, X_3)$ represent conditional probabilities of discrete nodes X_1 and X_4, while $f(X_3)$ and $f(X_2|X_1)$ are the probability density functions of continuous nodes X_2 and X_3, respectively. The joint probability of the discrete nodes $P(X_1, X_4)$ can be obtained by marginalization calculation.

In the case that the domain of node X_4 can be determined by the outcome space of its parent nodes, then $P(X_1, X_4)$ can be written by:

$$P(X_1, X_4) = P(X_1) \iint_{\Omega_{X_4}(X_2, X_3)} f(X_3)f(X_2|X_1)dX_2dX_3 \quad (4)$$

where $\Omega_{X_4}(X_2, X_3)$ represents variable X_4 as a domain in the outcome space of variables X_2 and X_3. The form of Eq. (4) is in line with the definition of structural reliability problems, and hence can be estimated by means of SRMs.

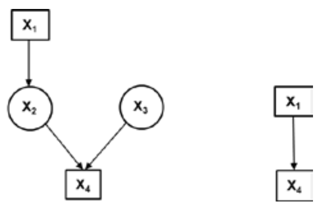


Fig. 2 An example of reduction of an eBN into BN (circle represents continuous node and rectangle represents discrete node)

2.3 New observation on continuous nodes

As already stated, BNs show a powerful capability in updating probabilistic propagation through given observations. As previously discussed, the evidence is inserted to replace certain prior probability on observed nodes, and the probabilities of the other nodes are updated using exact algorithms in discrete BNs. Similarly, in eBNs, it is necessary to discretize continuous nodes with evidence at first, and then the corresponding discrete nodes are kept in place of the continuous nodes in the reduced BNs.

A plethora of discretization methods for continuous nodes in the BNs has been investigated for many years [24-26]. Currently, there are no formalized approaches for the discretization of continuous random variables. Thus, for the problem studied in this paper, a credible discretization approach for eBNs [23] is used.

The previously introduced example is now reintroduced to explain how to discretize continuous nodes in eBNs. As shown in Fig. 3, node X_3 is substituted with two nodes, a discrete variable $X_{3discrete}$ and a continuous variable $X_{3continuous}$.

$X_{3discrete}$ has i states that are defined by the outcome space of X_3 with conditional cumulative distribution function $F_{X3}[x_3]$, and the number of its states is identical to corresponding intervals of the divided domain of X_3. Each sub-domain of X_3 can be represent by $[\underline{x}_{3i}, \overline{x}_{3i}]$, where $\underline{x}_{3i}$ and $\overline{x}_{3i}$ denote the lower and upper bounds of the interval, respectively. Then the probability mass function of $X_{3discrete}$ given the state i can be achieved as,

$$P\left(X_{3discrete}^{i}\right) = F_{X_3}[\overline{x}_{3i}] - F_{X_3}[\underline{x}_{3i}]$$

(5**Fehler! Textmarke nicht definiert.**)

On the other hand, $X_{3continuous}$, as the child of $X_{3discrete}$, inherits all the descendants and outcome space of X_3. The continuous variable $X_{3continuous}$ is eliminated from the model after it becomes a barren node by used of SRMs, and the discretized node $X_{3discrete}$ is retained to facilitate new observation updating the model.

In the same way, for inserting the evidence on X_3, the process of discretization is to split the domain of X_3 given the evidence into the sub-domains, each of which is obtained with a discrete probability value. In this study, the sub-domains on the observed continuous node are defined with the same length [23].

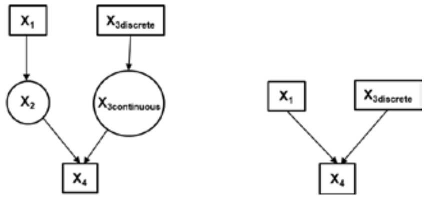


Fig. 3 An example of the discretization procedure

2.4 Credal networks

In the case imprecise probabilities are introduced to BNs, they are referred to as CNs since the node corresponding to an imprecise event is associated with a credal set instead of a CPT or a PDF. Credal sets are defined as closed convex sets associated with a set of probability distribution functions, which are used to represent imprecise probabilities in the graphical models. Fagiuoli and Zaffalon [27] used convex sets to compute posterior probabilities in a discrete BN with exact algorithms and first referred to this kind of model as CNs. A detailed introduction of CNs can be found in [28].

The inference for CNs is more complex than for BNs, still being in its infancy stage of development [29-31]. Thanks to the development of inference algorithms in CNs, some exact and approximate inference algorithms can be used for the reasoning of CNs although imprecise probabilities propagation in CNs is still under study. In this paper, the integration of CNs and SRMs [32] is adopted to analyze the stability of slopes.

2.4.1 Inference computation in CNs

The same as the elimination procedure of eBNs, continuous variables and interval variables in CNs also should be removed in the first step. As Fig. 4 shows, a simple CN consists of three types of nodes: discrete node X_1, continuous node X_2 with a known distribution, and an imprecise node X_3. The deterministic node X_4 is dependent of all the other three nodes.

Considering the simulation methods for the model elimination, direct Monte Carlo approach is a robust and feasible method to compute the probability of failure. It is a classical simulation tool suited for the reduction of eBNs. Nevertheless, it requires a very high number of samples in the case of small failure probabilities. This is especially the case in the analysis of slope failures, where failure probabilities are typically in the order of 10^{-4} or smaller. Therefore, advanced line sampling [33] is considered herein. It is a recently developed advanced Monte Carlo methods, based on line sampling [34] and an adaptive algorithm to adapt the important direction to the shape of limitation state surface. Most importantly, it allows for sets of probability distributions to be included in the estimation of imprecise failure probabilities, which are bounded with upper and lower probabilities. Because of these advantages, advanced line sampling is adopted for node elimination.

Then, after removing the continuous and imprecise nodes, the network only contains two types of conditional probability in discrete nodes: point probabilities and bounded probabilities. Afterwards, exact inference for BNs such as the variable elimination algorithm [15], can be applied here to estimate probability propagation in CNs.

Both of discrete nodes X_1 and X_4 are assumed as binary variables, and then the joint probability for identifying upper and lower bounds of nodes in the CN can be expressed as,

$$\boldsymbol{P}\left(\boldsymbol{X_1},\underline{\overline{\boldsymbol{X}}}_{\boldsymbol{4}}\right)=\boldsymbol{P}(\boldsymbol{X_1})\boldsymbol{P}\left(\underline{\overline{\boldsymbol{X}}}_{\boldsymbol{4}}\middle|\boldsymbol{X_1}\right)$$

(6**Fehler! Textmarke nicht definiert.**)

in which $\underline{\overline{X}}_4$ denotes the upper and lower bounds in node X_4 with two states x_{41} and x_{42}. Then, according to variable elimination, exact bounds of marginal probability with upper bound in the state x_{11} of node X_1 can be obtained as,

$$P(\overline{x}_{11})_{exact}=max\left(\sum_{\underline{\overline{X}}_4}P(X_1)P\left(\underline{\overline{X}}_4\middle|X_1\right)\right)$$

$$=\boldsymbol{max}\begin{bmatrix}\boldsymbol{P}(\boldsymbol{x_{11}})\boldsymbol{P}(\underline{\boldsymbol{x}}_{\boldsymbol{41}}|\boldsymbol{x_{11}})+\boldsymbol{P}(\boldsymbol{x_{11}})\boldsymbol{P}(\overline{\boldsymbol{x}}_{\boldsymbol{42}}|\boldsymbol{x_{11}})\\ \boldsymbol{P}(\boldsymbol{x_{11}})\boldsymbol{P}(\overline{\boldsymbol{x}}_{\boldsymbol{41}}|\boldsymbol{x_{11}})+\boldsymbol{P}(\boldsymbol{x_{11}})\boldsymbol{P}(\underline{\boldsymbol{x}}_{\boldsymbol{42}}|\boldsymbol{x_{11}})\end{bmatrix}$$

(7**Fehler! Textmarke nicht definiert.**)

The lower bound of the marginal probability can be obtained similarly with the minimum operator. Although traditional exact inference algorithms are efficient to compute the exact bounds, the exact inference is highly inefficient and leads to a combinatorial explosion in the case of complex networks, since it requires the evaluation of every possible bound combination for every node.

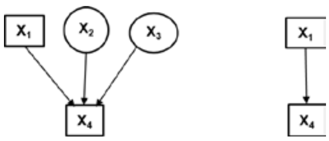


Fig. 4 An example of elimination procedure in a CN (Circle, rectangle, Ellipse denotes continuous, discrete and imprecise node, respectively)

A novel algorithm has been introduced to avoid this combinatorial explosion encountered by exact inference [35]. The outcome from this approach can get the inner bounds, which can be equal to the exact bounds if no nodes with probability interval are observed. For a query node, briefly, instead of computing the true bound identifying all of the combinations of the bounds in input, the key step is to compare the conditional probabilities of the query variable given the related nodes in CNs. Therefore, it is obvious that the result by use of this kind of inner approximation is exact if there is no evidence involved in the bounded nodes.

It has been testified that this approach makes the computation low-cost, and it is effective to obtain real-time results concerning the imprecise nodes in the model [35].

3 Illustrative Example 1: Failure analysis of the soil slope with eBN

3.1 Problem description

Two models are studied herein. One model is constructed with an infinite slope, which has a soil layer 4 m thick at an inclination of 3H to 2V. Another model with the same slope angle including two materials: 4 m thickness of the soil layer and bedrock at the height of 10 m is studied. Furthermore, the types of slopes failure are considered by two methods of stability analysis (see Fig. (5)). Specifically, the infinite slope has an assumed translational slip surface (Failure Model 2), is studied by considering the driving forces and resisting forces, comparing them and calculating the Factors of Safety. Meanwhile a slope without the assumed sliding surface (Failure model 1), is analyzed by finite element method (FEM). The detailed process is presented in the following section.

3.2 The structure of the network

In this section, different shapes of failure plane as well as two different analysis methods of slope stability are combined with eBNs approach. Based on the cause-effect relation, a BN is built in Fig. (5). Two failure models are studied as the consequence events, and connected with the crucial factors affecting the slope failure.

The failure models represent different shapes of failure surface, maybe a circle or a non-circle, and most of time it cannot be achieved the failure mechanism of a slope in advance. So, FEM is used herein to analyze a slope with the uncertain slip surface, where the failure event is denoted by *Failure Model 1* (FM1) in the network. Moreover, the uncertain soil parameters: cohesion, friction angle and the varying position of the groundwater table are considered as input for the response of the factors of safety, which is determined by the shear strength reduction (SSR) method in geotechical software RS2v7.0 [36].

For *Failure Model 2* (FM2), an infinite slope with a known slip surface is studied herein. Limited equilibrium technique is used to analyze the slope stability. Based on this, the cause-effect relationship is built in the BN, where nodes *Cohesion* and *Friction Angle* are the resisting parameters preventing the occurrence of a failure. Meanwhile, the geometrical parameters of the slope are the slope inclination and slope's height, being also two important factors for slope stability. The angle of a slope defines how much driving force is distributed in the parallel direction along the slope surface. Small angles mean small pulling force on the downslope movement while large angle provides the large pulling force. In this model, the total height and angle of the slope are constant, so they are not considered into this BN.

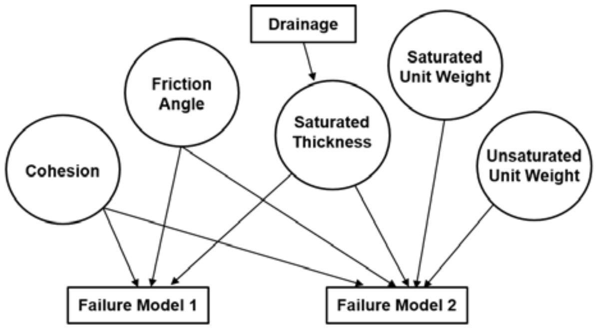


Fig. 5 The hybrid BN model of an infinite slope

Furthermore, the nodes *Unsaturated Unit Weight*, *Saturated Unit Weight* and *Saturated Thickness* are selected in the slope model according to effective stress principle [34], in which pore water pressure is defined by the unit weight of soil and the corresponding soil thickness. In such conditions, it was also considered the influence of the water table in the slope stability.

The position of the water table is an unfavourable variable defining the slope safety. The node *Saturated Thickness* can represent the depth of saturated soil, which is the level of the water table. This random variable is governed by the drainage condition. To be specific, the water table is away when drainage takes place. If not, the depth of saturated soil will assume random values ranging under the soil surface. In general, the event of *Drainage* affects the node *Saturated Thickness*.

3.3 The quantification of a network

3.3.1 Limited equilibrium function

Factors of safety are frequently computed to identify whether a slope is safe, which can be obtained by the ratio of resisting and driving stresses along a potential slip surface. This calculation, however, is not based on a unique equation, since there are a variety of methods [37, 38] that can be selected to obtain the factor of safety according to different conditions. These conditions also depend on the type of failure surface and its extension.

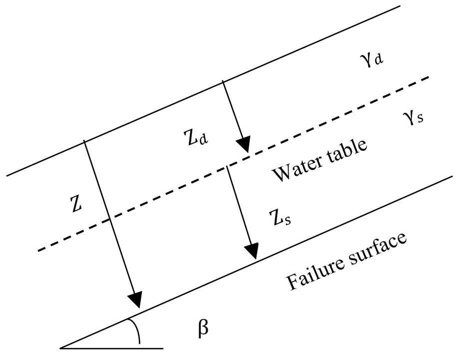


Fig. 6 The slope with translational slip

In the analysis of a given failure surface, as Fig. (6) shows, the equation for the factor of safety in terms of effective stress analysis is given by

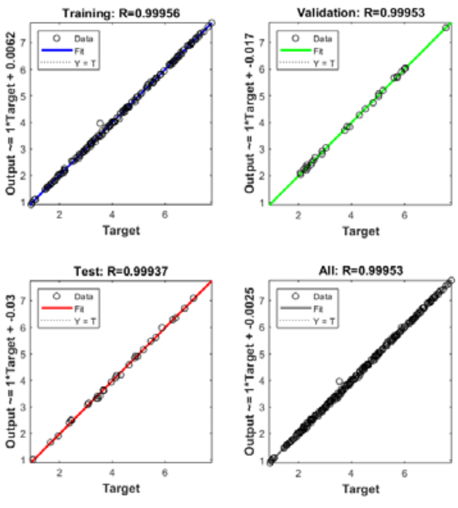
$$\boldsymbol{FOS} = \frac{\boldsymbol{c + (\gamma_d Z_d + \gamma_s Z_s - \gamma_w Z_s) * \cos\beta * \tan\phi}}{\boldsymbol{(\gamma_d Z_d + \gamma_s Z_s) * \sin\beta}}$$

(8**Fehler! Textmarke nicht definiert.**)

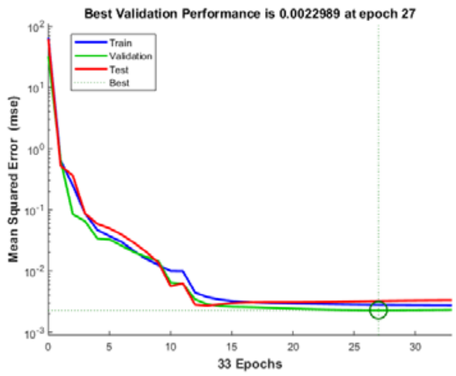
here, the drained parameters of cohesion (c) and friction angle (ϕ) are parameters governing the soil strength. Z_d and Z_s are the thickness of unsaturated and saturated soil layer, respectively, and the sum of them is the total thickness of soil (Z). β is the slope inclination and γ_w is the unit weight of water, 9.81kN/m^3. For the layer above and below the water table, soil unit weight should be split into two parts: dry unit soil weight (γ_d) and saturated unit soil weight (γ_s). This analysis has been completed using the equilibrium of an infinite [39]. Moreover, FOS ≤ 1 means the slope fails, whilst the FOS larger than 1 indicates the slope is safe. All the calculations are performed in effective stresses but, for the sake of simplicity, the effective parameters, cohesion and friction angle, are simply denominated as c and ϕ, as there is no risk to misunderstand effective and total strength resistances.

3.3.2 Finite element analysis

A set of response results is computed by FEM, where 200 experiment data are selected based on full factorial design, wherein the number of levels for c, ϕ and Z_s is 5, 5, 8, respectively and the results of FOS are carried out by the experimental runs on c, ϕ and Z_s. Then the response relationship is built via artificial neural networks (ANN) approach. Matlab R2018a 'nftool' is used to train and test the proposed model, where the ANN includes three layers: input (c, ϕ and Z_s), hidden layer and output layer (FOS). In Fig. (7), the results from training, validation and test data (140, 30, 30 samples, respectively) all show the good linear relationship, and the mean squared error of them is at the level of around 10^{-3}. Afterwards, this black-box of input-output can be saved as 'net' in the workspace, then put it to work in a BN model on new inputs, wherein the node is defined with this 'net'.



(a) Linear regression plots



(b) The performance plots

Fig. 7 Results of ANN: (a) linear regression (b) the performance

3.3.3 Evidence observation

The definitions of variables involved in the BN are shown in Table 1. A coefficient of 0.85 [40] is adopted to describe the correlation between γ_d and γ_s.

The probabilities of two failure models are computed by the limit state function G(X),

$$\mathrm{G(X)} = \mathrm{FOS} - 1 \tag{9}$$

in which the node is a discrete variable with two states: G(X)>0, the node denotes the probability of a stable slope, otherwise, it is the failure probability of the slope.

According to Eq. (9), the probabilities of the slope state can be expressed as:

$$\mathrm{P}(SF) = \begin{cases} P_f, & G(X) < 0 \\ P_s, & G(X) \geq 0 \end{cases} \tag{10}$$

herein P_f denotes the failure probability of the slope while the safe probability is P_s.

Table 1 Input parameters of slopes

Parameters	Variable type	CPD*
c (kPa)	Continuous	logN(22, 10)
ϕ (°)	Continuous	N(35,3)
γ_d (kN/m^3)	Continuous	N(17, 0.4)
γ_s (kN/m^3)	Continuous	N(19, 0.5)
Z_s (m)	Continuous	U(0, 4) or 0
Drainage (D)	Discrete	[0.5, 0.5]
Slope Failure	Discrete	[P_f, P_s]

* N, logN, and U represent normal, lognormal and uniform distribution with mean and standard deviation, respectively.

To characterize the relationship between slope stability and its influence factors, one easy way is to check the sensitivity of slope failure by inserting new evidence on the induced-factors in the BN, respectively. Then in this work, we initially make some observations on continuous nodes by giving specific distribution range of random variables. According to the expert knowledge, initially, the ranges of distribution of c, ϕ, Z_s, γ_d and γ_s are defined with the closed interval: [0, 100], [25, 45], [0, 4], [16, 19], [18, 21], respectively. The further observation is made to identify the key factors by changing the range of distribution of each parameter, in which the interval of distribution is narrowed to about 50% of the initial observed range.

3.4 Results from example 1

The results of the two failure models are obtained simultaneously. In Table 2 (computational time is about 2.99 seconds), the failure probabilities of the two Failure modes: FM1 and FM2 are similar, 7.77% and 7.21%, respectively. Given the condition of drainage, the occurrence of failure of the two slopes are close to 0 and 0.06%, respectively, which are much lower than the state of no drainage, whose results are 8.01% and 7.50%. That means that if drainage takes place, it can stabilize the slope. Therefore, it can be reasonably achieved that drainage is decisive to the soil slope. In light of this, the decision maker knows the disaster can be avoided if he spends money in draining the slope.

Table 2 The effect of Drainage on slope safety

State	P(FM)	P(FM|D=false)	P(FM|D=true)
FM1	7.77e-02	8. 01e-02	1.00e-08
FM2	7.21e-02	7.50e-02	6.00e-04

In geotechnical problems, it is common that the soil characterization is performed in different phases and, therefore, new observations can be obtained in an advanced step of the study. These new results (the elapsed CPU time is lower than 10 seconds) serve to identify the influence of soil parameters on the slope stability. The adoption of the discretized approach allows considering these new results as evidence, updating the probabilities in the model. From the results in Table 3, the failure probability of FM2 varies from 7.25% to 7.38%, which is very close to the original result. Similarly, FM1 also shows a slight variation around the initial result, but the new information indicates a negative tendency on slope stability.

Table 3 BNs updated with evidence

Node	c	ϕ	Z_s	γ_d	γ_s
Evidence	[0, 100]	[25, 45]	[0, 4]	[16, 19]	[18, 21]
P(FM1)	7.93e-02	7.81e-02	7.91e-02	-	-
P(FM2)	7.37e-02	7.26e-02	7.28e-02	7.38e-02	7.25e-02

Such a small variation in the failure probability contributes to the large range given by the first observation. Hence, the outcome will be much more distinct if the observed intervals are narrower, which could be a result of additional geotechnical tests. Through further observation in Table 4, showing more obvious the effect on the results, where P(FM1) is 7.79%, 7.72% and 0.35%, respectively, with the corresponding limited ranges. With the results of FM1, we can infer that Z_s greatly affect the reduction of the slope failure, in comparison with c and ϕ having a smaller effect on the slope failure. Likewise, for FM2, the slope stability is mainly affected by the varying c and Z_s comparing to ϕ, γ_d and γ_s. Generally, the uncertainty of Z_s has more influence on FM1 while the variation of c and Z_s are more sensitive for the slope stability with FM2.

Table 4 BNs updated with further observation

Node	c	ϕ	Z_s	γ_d	γ_s
Evidence	[25, 75]	[30, 40]	[1, 3]	[17, 18]	[19, 20]
P(FM1)	7.79e-02	7.72e-02	3.50e-03	-	-
P(FM2)	1.00e-04	7.02e-02	9.37e-02	7.36e-02	7.23e-02

In spite of the coarse results, the decision maker will immediately obtain real-time information about the possibility of slope failure. This real-time information support can be useful for the requirement of real-time analysis of the risk of potential failure.

4 Illustrative Example 2: Failure analysis of the soil slope with CNs

4.1 Problem description

Igneous rock like granite or gneiss is present in some regions, where the weathering of the rock produces so-called "residual soils". These materials are very common in mountainous countries as the case of Portugal, Spain, Brazil, China, Hong Kong, Singapore, and Africa.

An extensive geotechnical characterization of residual granite soils has been carried out in the northern part of Portugal [41-43]. The common strength parameters are found in the residual soil from Granite in the Porto region. The mean values for strength parameters of this type of soil from Porto, such as cohesion and friction angle, are represented by interval-valued quantities to cope with the lack of information, and are represented by means of p-boxes. Unsaturated and saturated unit soil weight are both defined based on expert knowledge. Additionally, for a typical design, a slope in residual soils is typically designed with a fixed inclination of 3H to 2V, and the total soil thickness of this slope is assumed as 4 m in this study. The failure surface is considered parallel to the surface of the slope, as shown in Fig. (8).

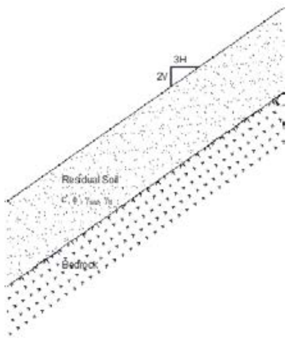


Fig. 8 A residual soil slope

Table 5 Input parameters of the residual soil

Parameters	Scenario 1	Scenario 2		Scenario 3*
		Min	Max	
c	logN(20, 4)	0	70	logN(μ_c,4), $\mu_c \in [16, 22]$
ϕ	N(37, 1.85)	25	47	N(μ_f,1.85), $\mu_f \in [36, 38.5]$
γ_d	N(18.5,0.5)	17	20	[17, 20]
γ_s	N(20, 0.6)	18	22	[18, 22]

* μ indicates the mean of the distributions. The notes of Table 1 also apply here.

Three different situations of information available in c, ϕ, γ_d and γ_s, are studies herein (see Table 5). A BN of the slope is used here as a reference, and for the other two scenarios, interval analysis is adopted to cope with the limited information. If further information about the variables can be achieved, such as input distribution with a bound on its mean, then the parametric p-boxes is introduced in the imprecise nodes *Cohesion* and *Friction Angle*. Thus it is possible to observe the change of the results in comparison with only interval nodes in the model.

4.2 The structure of the network

The CN based on the previous BN model is built to estimate the probability of slope failure with limited information subject to drainage influence.

This model presents nine nodes, including discrete variables, continuous variables, interval variables and parametric p-boxes. These corresponding nodes are represented by rectangular, circle, ellipse and trapezoid, respectively (see Fig. (9)). If there is scarce information provided, for example, the parameters c, ϕ, γ_d and γ_s change with geological/geotechnical conditions. Then without any geotechnical test, it cannot be known in advance the exact properties of them. In this case, they are associated with imprecise information. Such as scenario 2, these imprecise nodes can be defined by interval-values from expert judgement.

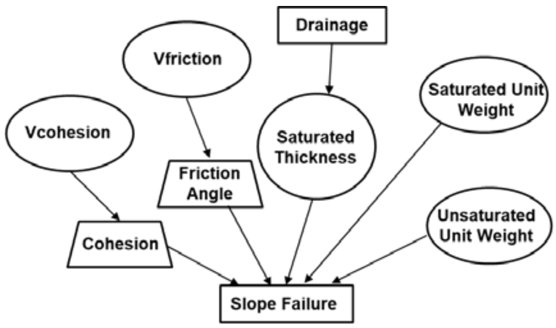


Fig. 9 The CN model of an infinite slope

However, if further information is available, such as the distribution types of nodes *Cohesion* and *Friction Angle* are known and the distribution parameters are uncertain, then the two soil parameters can be described by the parametric p-boxes. In this CN, the imprecise information is presented by a combination of the nodes *Vcohesion* and *Cohesion*, *Vfriction* and *Friction Angle*. Comparing to the previous BN, the nodes *Cohesion* and *Friction Angle* in the CN model are substituted by the respective parametric p-boxes.

Slope Failure (SF) is the node of interest in the CN, whose failure state of the node can predict the occurrence of a shallow landslide. The probability of slope failure is inferred by marginal probability calculation in the reduced CN. Furthermore, an analysis can be conducted to demonstrate the effect of the node *Drainage* on the slope stability. The analysis is conducted in the software OpenCossan [44-45]. The computation tool provides eBN methodology and the above-mentioned inference for CN. Traditional and advanced Monte Carlo methods also are included in this tool. For this model, adaptive line sampling [33] is used to estimate the lower and upper bounds of the failure probability. Additionally, the computation takes a few seconds in the software.

4.3 Results from example 2

From the results (Table 6), it can be seen that an exact probability of slope failure can be obtained with the precise input for the conservative model. The 2.74% failure probability indicates a reasonable degree of stability for this existing slope with precisely specific parameters. However, in the case of poor information, the input uncertainty affects the precision of output so that the results are denoted with the probability bounds. When the input nodes *Cohesion*, *Friction Angle*, *Unsaturated Unit Weight*, and *Saturated Unit Weight* only can be defined as interval variables with the limited information, the probability bound of slope failure is between 0 and 1. The result is too wide to provide useful information regarding the slope stability. In other words, each combination of the different values of the factors can produce any possibility of the slope states, failure or safe. Hence, the feasible way is reducing the uncertainty input to increase the precision of the output, what can be done by producing additional geotechnical information or by approaching the reliability problem with different methods. For example, a practical common geotechnical solution would result from performing additional boreholes in the slope and laboratory test what would allow to more precise geotechnical parameters.

Table 6 Slope failure probability

Different information	Scenario 1	Scenario 2	Scenario 3
P(SF)	0.0274	[0, 1]	[0, 0.0711]

Comparing to the first two input information, further observation is added to the probability boxes in the imprecise nodes *Cohesion* and *Friction Angle*. As it is shown in Table 6, the probability bound of failure slope became dramatically tighter after introducing P-boxes. The range of the failure result is from 0 to 7.11%, and the upper bound of the failure slope reveals a steep decrease. Besides, the precise result with 2.74% is included in this range. It illustrates the actual slope failure can be estimated with the consideration of the reasonable application of parametric p-boxes in the CN model.

In Table 7, the possibility of slope failure under drained conditions shows a greatly reduced tendency, and even the risk of failure can decrease to 0 in contrast to the state of no drainage. That is because if water is away, the percolation forces disappear and the resistant forces also increase, as a result of the increase in the normal force and, therefore, the friction component of the strength also increases.

Table 7 Failure probabilities with the state of Drainage

Different information	Scenario 1	Scenario 2	Scenario 1
P(SF|D = True)	0	[0, 1]	0
P(SF|D = False)	0.0514	[0, 1]	[0, 0.0153]

The result with the interval [0, 1] based on the very poor information cannot give further information for decision-makers, but the probability bound of *Slope Failure* with the evidence *Drainage* makes sense by ways of p-boxes. Specifically, if drainage is not implemented, the failure result of the residual soil slope with [0, 1.53%] is much wider than the one with drainage.

5 Conclusion

This study presents applications of the advanced BNs methods to estimate the failure probabilities of the slope subjected to drainage state. To characterize the effect of the induced-factors on the slope failure, new observations are made in some continuous nodes to update the model. The proposed methods proved to be useful and with a reduced cost of computation providing real-time information for the decision makers. Also, the model presents the capability of integrating different events.

Enhanced BNs and CNs are applied to rely on input information availability. Enhanced BNs consist of two types of nodes, continuous and discrete nodes, where an integration of BNs and structural reliability analysis is applied to make the inference in this precise model, while CNs, especially for the scenario that there is no enough abundant information to get the precise CPDs for each of nodes. Additionally, discrete variables, random variables, interval variables and p-boxes are presented in the model. The bounds of results provide a rough estimation of the slope failure. The permission of the application of p-boxes in the model contributes to the reduction of the uncertainty in output. Moreover, a discretization process is applied when new evidence enters the continuous nodes. These capabilities ensure the wide flexibility of the model in analysing the slope failure.

The two examples demonstrate that the approach has interesting possibilities for analyzing the failure of slopes. The exact failure probabilities of soil slopes in the first example indicate a low failure, and according to the analysis of updating the information in the specific nodes, the conclusion can be made that the failure of the slope can be significantly reduced with drainage. Although interval-value is a suitable way for representing the non-probabilistic information, the interval results of the residual soil slope may fail to acquire the usable range of real value. In this case, p-boxes involved obviously narrow the bound of failure probability. All in all, both of eBNs and CNs are effective and feasible means to make failure analysis of one or more slopes.

Acknowledgement

This work presented in this article is supported by the Chinese Scholarship Council.

References

[1] Ma. J, H. Tang, X. Hu, A. Bobet, M. Zhang, T. Zhu and Y. Song, A. M. Mutasim "Identification of Causal Factors for the Majiagou Landslide Using Modern Data Mining Methods", Landslides, 14(1), pp. 311–322, 2017.

[2] Kristo, C, H. Rahardjo and A. Satyanaga. "Effect of Variations in Rainfall Intensity on Slope Stability in Singapore", International Soil and Water Conservation Research, 4(1), pp. 258-264, 2017.

[3] Rahardjo, H, S. N. Alfrendo, C. L. Eng and Y. S. Ng. "Effects of Groundwater Table Position and Soil Properties on Stability of Slope During Rainfall", Geotechnical and Geoenvironmental Engineering, 136(10), pp. 1555-1564, 2010.

[4] Tsai, F, J.S. Lai and T.H. Lin. "Analysis of Topographic and Vegetative Factors with Data Mining for Landslide Verification", Ecological Engineering, 61, pp. 669-677, 2013.

[5] Oka. Y and T. H. Wu. "System Reliability of Slope Stability", Geotechnical Engineering, 116(8), pp. 1185–1189, 1990.

[6] Liu. Y, W. Zhang, L. Zhang, Z. Zhu, J. Hu and H. Wei. "Probabilistic Stability Analyses of Undrained Slopes by 3D Random Fields and Finite Element Methods", Geoscience Frontiers, 9, pp. 1657–1664, 2018.

[7] El-Ramly. H, N. R. Morgenstern, and D. M. Cruden. "Probabilistic Slope Stability Analysis for Practice", Canadian Geotechnical Journal, 39(3), pp. 665–683, 2002.

[8] Metya. S, T. Mukhopadhyay, S. Adhikari, G. Bhattacharya. "System Reliability Analysis of Soil Slopes with General Slip Surfaces Using Multivariate Adaptive Regression Splines", Computers and Geotechnics, 87(7), pp. 212–228, 2017.

[9] Ching, J, K. K. Phoon and Y. G. Hu. "Efficient Evaluation of Reliability for Slopes with Circular Slip Surfaces Using Importance Sampling", Geotechnical and Geoenvironmental Engineering, 135(6), pp. 768-777, 2009.

[10] Wang. Y, Z. J. Cao and S. K. Au. "Practical Reliability Analysis of Slope Stability by Advanced Monte Carlo Simulations in a Spreadsheet", Geotechnical and Geoenvironmental Engineering, 48(1), pp. 162-172, 2011.

[11] Sakellariou, M. G and M. D. Ferentinou. "A Study of Slope Stability Prediction Using Neural Networks", Geotechnical and Geological Engineering, 23, pp. 419-445, 2005.

[12] Chakraborty, A and D. Goswami. "Slope Stability Prediction of Slope Stability Using Multiple Linear Regression and Artificial Neural Network", Arabian Journal of Geosciences, 10(17), pp. 385-395, 2017.

[13] Pearl, J. "Probabilistic Reasoning in Intelligent Systems", In: Morgan Kaufmann Publisher, San Mateo, California, 1988.

[14] Straub, D. "Natural Hazards Risk Assessment Using Bayesian Networks", In: Proceedings of the 9th International Conference on Structural Safety and Reliability(ICOSSAR), 2005.

[15] Pearl, J and S. Russell. "Causality: Models, Reasoning, and Inference", In: Cambridge University Press, 2000.

[16] Song, Y.Q, J. H. Gong and S. Gao. "Susceptibility Assessment of Earthquake-induced Landslides Using Bayesian Network: A Case Study in Beichuan, China ", Computers and Geosciences, 42, pp. 189-199, 2012.

[17] Liu, W.S, S. B. Li and R. Tang. "Slope Stability Evaluation in Open Pit Based on Bayesian Networks ", In: Proceedings of the ICCEAE, 2013, pp. 1227-1231.

[18] Peng, M, X. Y. Li and D. Q. Li. "Slope Safety Evaluation by Integrating Multi-source Monitoring Information", Structural Safety, 49(SI), pp. 65-74, 2013.

[19] Weber, P, G. Medina-Oliva, C. Simon and B. Iung. "Overview on Bayesian Networks Applications for Dependability, Risk Analysis and Maintenance Areas", Engineering Applications of Artificial Intelligence, 25(4), pp. 671-682, 2012.

[20] Daly, R, Q. Shen and Q. S. Aitken. "Learning Bayesian Networks: Approaches and Issues", The Knowledge Engineering Review, 26(2), pp. 99-157, 2011.

[21] Langseth, H, T. D. Nielsen, R. Rumi and A. Salmerón. "Inference in Hybrid Bayesian Networks", Reliability Engineering System Safety, 94, pp. 1499-1509, 2009.

[22] Straub, D and Der Kiureghian, A. "Combining Bayesian networks with Structural Reliability Methods: Application", Journal of Engineering Mechanics, 136(10), pp. 1259-1270, 2010.

[23] Straub, D and Der Kiureghian, A. "Combining Bayesian networks with Structural Reliability Methods: Methodology", Journal of Engineering Mechanics, 136(10), pp. 1248-1258, 2010.

[24] Dougherty, J, R. Kohavi and M. Sahami. "Supervised and Unsupervised Discretization of Continuous Features", In: Proceedings of the 12th International Conference on Machine Learning (ICML), Morgan Kaufmann Publishers, San Francisco, 1995.

[25] Kurgan, L and K.J. Cios. "Discretization Algorithm that Uses Class-attribute Interdependence Maximization", In: Proceedings of the International Conference on Artificial Intelligence (IC-AI), Las Vegas, Nevada, 2001, pp. 980-987.

[26] Chen, Y. C., T. A. Wheeler and M. J. Kochenderfer. "Learning Discrete Bayesian Networks from Continuous Data", Journal of Artificial Intelligence Research, 59, pp. 103-132, 2017.

[27] Fagiuoli, E and M. Zaffalon. "2U: An Exact Interval Propagation Algorithm for Polytrees with Binary Variables", Artificial Intelligence, 106(1), pp. 77-107, 1988.

[28] Cozman, F. G. "Credal Networks", Artificial Intelligence, 120, pp. 199-233, 2000.

[29] Tessem, B. "Interval Probability Propagation", International Journal of Approximate Reasoning, 7, pp. 95-120, 1992.

[30] Ferreira, R, J. C. and F. G. Cozman, " Inference with Separately Specified Sets of Probabilities in Credal Networks", In: Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence, San Francisco, 2002, pp. 430-437.

[31] Antonucci, A., De Campos, C.P. and D. Huber, and M. Zaffalon. "Approximate Credal Network Updating by Linear Programming with Applications to Decision Making", International Journal of Approximate Reasoning, 58(IS), pp. 25-38, 2015.

[32] Tolo, S., E. Patelli and M. Beer. "An Open Toolbox for the Reduction, Inference Computation and Sensitivity Analysis of Credal Networks", Advances in Engineering Software, 115, pp. 126-148, 2018.

[33] De Angelis, M., E. Patelli and M. Beer. "Advanced Line Sampling for Efficient Robust Reliability Analysis", Structural Safety, 52(SI), pp. 170-182, 2014.

[34] Terzaghi, K. "Theoretical Soil Mechanics", John Wiley Sons, New York, 1943.

[35] Tolo, S., E. Patelli and M. Beer. "An Inference Method for Bayesian Networks with Probability Intervals", In: Proceedings of the joint ICVRAM-ISUMA Conference, Florianópolis, SC, Brazil, 2017.

[36] Rocscience. RS2 software. In reocscience, version 7.0, computer program, 2011.

[37] Morgenstern, N. R and V. Price. "The Analysis of the Stability of General Slip Surface", Geotechnique, 15(1), pp. 79-93, 1965.

[38] Niu, W.J. "Determination of Slope Safety Factor with Analytical Solution and Searching Critical Slip Surface with Genetic-traversal Random Method", Scientific World Journal, 2014.

[39] Acharya, G., F. De Smedt, and N. T. Long "Assessing Landslide Hazard in GIS: A Case Study from Rasuwa, Nepal", Bulletin of Engineering Geology and the Environment, 65(1), pp. 99-107, 2006.

[40] Pinheiro Branco, L., A. Topa Gomes, A. Silva Cardoso and C. Santos Pereira. "Natural Variability of Shear Strength in a Granite Residual Soil from Porto", 32(4), pp. 911-922, 2014.

[41] Rubio, E., J. W. Hall and M. G. Anderson. "Uncertainty Analysis in a Slope Hydrology and Stability Model Using Probabilistic and Imprecise Information", 31, pp. 529-536, 2004.

[42] Viana da Fonseca, A and R. Coutinho. "Natural Variability of Shear Strength in a Granite Residual Soil from Porto", ", In: Proceedings of the 3rd International Conference on Site Characterization, 2008, pp. 195-248.

[43] Topa Gomes, A. "Elliptical Shafts by the Sequential Excavation Method. The case of Metro do Porto, (in Portuguese)", pp. 195-248, 2009.

[44] Edoardo P., M. Broggi, S. Tolo and J. Sadeghi. "COSSAN software a multidisciplinary and collaborative software for uncertainty quantification", pp. 212-224, 2017.

[45] Edoardo P. " Handbook of Uncertainty Quantification COSSAN: A Multidisciplinary Software Suite for Uncertainty Quantification and Risk Management", pp. 1-69, 2016.