# Advanced Bayesian Networks for Reliability and Risk Analysis in Geotechnical Engineering

Von der Fakultät für Bauingenieurwesen und Geodäsie
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades

Doktor der Ingenieurwissenschaften
Dr.-Ing.
genehmigte Dissertation
von

Dipl.-Ing. Longxue He

2020

Referent: Prof. Dr.-Ing. Michael Beer

Korreferent: Prof. António Topa Gomes

Tag der Promotion: 20.02.2020

# Erklärung zum Promotionsgesuch

Ich erkläre, dass ich die Regeln der geltenden Promotionsordnung kenne und eingehalten habe und mit einer Prüfung nach den Bestimmungen der Promotionsordnung einverstanden bin. Ich habe die Dissertation selbst verfasst, keine Textabschnitte von Dritten oder eigener Prüfungsarbeiten ohne Kennzeichnung übernommen und alle benutzten Hilfsmittel und Quellen in der Arbeit angegeben. Weiterhin habe ich Dritten weder unmittelbar noch mittelbar geldwerte Leistungen für Vermittlungstätigkeiten oder für die inhaltliche Ausarbeitung der Dissertation erbracht und die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht. Die vorliegende Arbeit wurde ebenso nicht bei einer anderen Hochschule als Dissertation eingereicht. Ich bin damit einverstanden, dass die Dissertation auch zum Zwecke der Überprüfung der Einhaltung allgemein geltender wissenschaftlicher Standards genutzt wird, insbesondere auch unter Verwendung elektronischer Datenverarbeitungsprogramme.

Hannover, am ______________________ ______________________

Datum Unterschrift

# Abstract

The stability and deformation problems of soil have been a research topic of great concern since the past decades. The potential catastrophic events are induced by various complex factors, such as uncertain geotechnical conditions, external environment, and anthropogenic influence, etc. To prevent the occurrence of disasters in geotechnical engineering, the main purpose of this study is to enhance the Bayesian networks (BNs) model for quantifying the uncertainty and predicting the risk level in solving the geotechnical problems. The advanced BNs model is effective for analyzing the geotechnical problems in the poor data environment. The advanced BNs approach proposed in this study is applied to solve the stability of soil slopes problem associated with the specific-site data. When probabilistic models for soil properties are adopted, enhanced BNs approach was adopted to cope with continuous input parameters. On the other hand, Credal networks (CNs), developed on the basis of BNs, are specially used for incomplete input information. In addition, the probabilities of slope failure are also investigated for different evidences. A discretization approach for the enhanced BNs is applied in the case of evidence entering into the continuous nodes. Two examples implemented are to demonstrate the feasibility and predictive effectiveness of the BNs model. The results indicate the enhanced BNs show a precisely low risk for the slope studied. Unlike the BNs, the results of CNs are presented with bounds. The comparison of three different input information reveals the more imprecision in input, the more uncertainty in output. Both of them can provide the useful disaster-induced information for decision-makers. According to the information updating in the models, the position of the water table shows a significant role in the slope failure, which is controlled by the drainage states. Also, it discusses how the different types of BNs contribute to assessing the reliability and risk of real slopes, and how new information could be introduced in the analysis. The proposed models in this study illustrate the advanced BN model is a good diagnosis tool for estimating the risk level of the slope failure. In a follow-up study, the BNs model is developed based on its potential capability for the information updating and importance measure. To reduce the influence of uncertainty, with the proposed BN model, the soil parameters are updated accurately during the excavation process, and besides, the contribution of epistemic uncertainty from geotechnical parameters to the potential disaster can be characterized based on the developed BN model. The results of this study indicate the BNs model is an effective and flexible tool for risk analysis and decision making support in geotechnical engineering.

Keywords: Bayesian networks, Imprecise probability, Uncertainty, Stochastic model updating, Sensitivity analysis

# Kurzfassung

Die Stabilitäts und Verformungsprobleme von Böden sind seit Jahrzehnten ein großes Forschungsthema. Die möglicherweise katastrophalen Ereignisse werden durch verschiedene komplexe Faktoren wie unsichere geotechnische Bedingungen, äußere Umgebung und anthropogenen Einfluss ausgelöst. Um das Auftreten von Katastrophen in der Geotechnik zu verhindern , liegt der Hauptzweck dieser Studie auf der Erweiterung des Bayes'schen Netzwerkmodells (BNs) zur Quantifizierung der Unsicherheit und zur Vorhersage des Risikoniveaus bei der Lösung geotechnischer Probleme. Das erweiterte BNs-Modell ist für die Analyse geotechnischer Probleme in einer schlechten Datenumgebung geeignet. Der in dieser Studie vorgeschlagene fortgeschrittene BNs Ansatz wird verwendet, um das Problem der Stabilität von Bodensteigungen zu lösen, das mit den standortspezifischen Daten verbunden ist. Bei der Anwendung probabilistischer Modelle für Bodeneigenschaften wird ein erweiterter BNs-Ansatz angewendet, um mit kontinuierlichen Eingabeparametern zu arbeiten. Andererseits werden Credal-Netzwerke (CNs), die auf der Basis von BNs entwickelt wurden, speziell für unvollständige Eingabeinformationen verwendet. Darüber hinaus werden die Wahrscheinlichkeiten des Hangversagens für verschiedene Nachweise untersucht. Ein Diskretisierungsansatz für die erweiterten BNs wird angewendet, wenn Beweise in die kontinuierlichen Knoten eintreten. Zwei Beispiele sollen die Machbarkeit und prognostische Wirksamkeit der Modelle demonstrieren. Die Ergebnisse zeigen, dass erhöhte BNs ein genau niedriges Risiko für die untersuchte Steigung aufweisen. Im Gegensatz zu den BNs sind die Ergebnisse von CNs mit Grenzen dargestellt. Der Vergleich von drei verschiedenen Eingabeinformationen zeigt, je ungenauer die Eingabe ist, desto größer ist die Unsicherheit bei der Ausgabe. Beide können den Entscheidungsträgern nützliche katastrophenbedingte Informationen liefern. Nach der Aktualisierung der Informationen in den Modellen spielt die Position des Grundwasserspiegels eine wichtige Rolle beim Hangversagen, das durch die Entwässerungszustände gesteuert wird. Außerdem wird erörtert, wie die verschiedenen Arten von BNs zur Bewertung der Zuverlässigkeit und des Risikos von realen Gefällen beitragen und wie neue Informationen in die Analyse aufgenommen werden können. Die in dieser Studie vorgeschlagenen Modelle veranschaulichen, dass das fortgeschrittene BN-Modell ein gutes Diagnosewerkzeug für die Abschätzung des Risikograds des Hangversagens ist. In einer Folgestudie konzentrierte sich das entwickelte BNs Modell auf seine potenzielle Fähigkeit zur Informationsaktualisierung und Wichtigkeitsmessung. Um den Einfluss der Unsicherheit zu verringern, wurden mit dem vorgeschlagenen BN Modell die Bodenparameter während des Aushubprozesses genau aktualisiert. und außerdem konnte der Beitrag der epistemischen Unsicherheit von geotechnischen Parametern zur möglichen Katastrophe auf der Grundlage des entwickelten BN-Modells charakterisiert werden. Die Ergebnisse dieser Studie zeigen, dass der BN-Modal ein effektives und flexibles Instrument zur Risikoanalyse und Entscheidungsunterstützung in der Geotechnik ist.

Schlagwörter: Bayesianische Netzwerke, Ungenaue Wahrscheinlichkeit, Unsicherheit,

Stochastische Modellaktualisierung, Sensitivitätsanalyse

# Contents

# 1 Introduction

## 1.1 Background

Soil is the foundation of the infrastructure asset, such as highways, railways, channels, etc. In geotechnical problems, the strength and stiffness of the ground contribute to the stability problems and deformation problems. Risk arises where the potential ruptures of soil may incur adverse consequences, and even catastrophic phenomenon, such as landslides and debris flows, which are the large-scale movement of soil accompanied with economic loss and fatalities all over the world. According to the investigation (Dilley et al., 2005), due to the geological, geomorphological, and climate variations, nearly 300 million people in the world are living in areas of potential landslide risk. In Europe, average economic loss per year from landslides is approximately 4.7 billion Euros, and Haque et al. (2016) found fatal landslides still showed an increasing trend. Although the majority of slopes failure lead to shallow landslides, the catastrophic slope failures can occur within the long-term transformation. In May 2010, the catastrophic long-runout landslide happened in the Girová Mountain. The study indicated that it might evolve into a long-runout landslide due to the very humid Atlantic chronozone, where the May 2010 rainfall event induced the collapse of the Girová Mountain slope (Panek et al., 2011). For the purpose of reducing the risk, in the past decades, a number of methods have been studied to analyze the slope safety. Based on the strength reduction technique, Dawson et al. (1999) presented advantages of the limit analysis method to analyze the slope stability for an embankment. To quantify the effect of vegetation on the stability of a natural slope in Tuscany, Italy, Schwarza et al. (2010) employed the limit equilibrium method to calculate the slope stability related to the root reinforcement. Also, Griffiths and Lane (1999), Jiang et al. (2014), Dyson and Tolooiyan(2019) illustrated that the finite element method was an alternative approach for estimating the deformations and mechanisms of the slope failure. However, due to the unforeseen geotechnical conditions, the real-time information updating support is crucial and necessary for decision-makers in risk mitigation for the stability and deformation problems. Up to now, the stability of slopes, natural or artificial, is still a classical and important research topic in geotechnical engineering.

In view of the source of threats, some of fatal geo-hazards are mainly affected by natural hazards, such as rainfall, earthquake, storm and so on, while, from the anthropogenic point of view, the way of the influence of anthropic activities on the failure events are various, such as the construction of engineering, illegal mining and hill-cutting, ageing of infrastructure, etc. Soil plays a vital role in the research for problems of the safety. In the rapid promotion of urbanization, the occurrence of a mass of construction projects makes residents suffer from undesirable accidents. The damage and injury from the failure of construction are unexpected results. From 1994 to 2005, the collapse of construction in the tunnelling projects led to the economic loss of over 570 million (Cardenas et al., 2014). In Do et al. (2016), the collapse of a deep excavation in Nicoll highway, Singapore, resulted in four casualties due to the inability of the excavation support system to redistribute loads. Except that the reason of the extreme weather, some failure could occur in the case of the overestimation of the stability of constructions (Hsieh et al., 2008). According to the investigation of nine cases in the anthropic engineering activity-induced hazards of underground construction, Elbaz et al. (2016) found that underground construction collapses were induced mainly by anthropic activities: the over-excavation, poor quality of structural detailing and poor design, etc. From the database of accidents in the process of tunnel construction (Sousa, 2010), a database of 204 tunnel construction accidents was assembled in order to better understand the causes of accidents. To prevent humans' life and properties from the geotechnical hazard, it is necessary to systematically assess and manage the risks associated with geotechnical activities. The studies for the stability and deformation problems of geotechnical structures are always an important research topic.

The behavior of soils in geotechnical engineering and its properties are complicated, uncertain and incompletely understood. The characterization of geotechnical variability has been of concern over the past few decades (Phoon and Kulhawy, 1999; Elkateb et al., 2003; Lacasse et al., 2017). Uncertainties from the inherent soil variability have a great influence on the stability of constructions, and a better recognition of the variation of soil properties can be beneficial for the geotechnical design and risk reduction. Usually, appropriate soil parameters are selected as indicators of soil properties to evaluate the stability and deformation of soil. With the study of geotechnical parameters, Orense et al. (2014) studied the stability of a sandy slope affected by rainfall, and the results showed the temporal development of deformation could be examined associated with the influence of various slope parameters.

Thanks to the parametric studies (Vucetic and Dobry, 1991; Finno et al., 2007; Oliveira et al., 2011), 2-dimensional and 3-dimensional effects of excavations and slopes are modelled to analyze the stability and deformation problems. Based on the aforementioned methods, the response values of any geotechnical structure such as a factor of safety of slope or the movement of a retaining wall in the excavation can be obtained with some input parameters in a geotechnical site. It is common that uncertainties exist in the geotechincal engineering. The uncertainty quantification contributes to the robust design of engineering and the accurate prediction of the failure risk. Numerous studies have been in recent years to develop probability methods to deal with uncertainties in a systemic way, and the probability of failure in the geotechnical engineering are often computed as an identification of the potential risk. In the studies of Christian et al.

(1994) and Chen et al. (1997), they proposed reliability analysis methods to quantify the uncertainties in geotechnical problems. Furthermore, simulation approaches such as the Monte Carlo simulation, Importance sampling and Subset simulation, have been applied to estimate the stability or deformation problems of soil as well (Morgenstern et al., 2002; Miro et al., 2015; Li et al., 2017). Results from numerical researches indicated that uncertainties from natural origin of soil and lack of knowledge from project staffs or researchers, which have large influence on the calculation of the failure probability. Especially, with the limitation of knowledge, some uncertain parameters or induced-factors are quite difficult to be considered into a mathematical model simultaneously. Therefore, the ability for integrating the different information and multiple variables in a model is needed.

Although many attempts have been made for simulating, analyzing and predicting the complex failure mechanism in the geotechnical engineering, the accurate estimation of failure events are still very challenging in the real engineering design practice. An appropriate model in geotechnical engineering plays a pivotal role in analyzing the behavior of soil and making rational decisions. Artificial intelligence (AI) has been already widely applied for modelling the complex behavior of geotechnical materials. Artificial neural networks (ANNs) have been already used as a predictive model in geotechnical engineering (Goh et al., 1995; Janet et al., 2002). Chua and Goh (2005) used a neural network algorithm to model the soil-structure interaction behavior of deep excavations, and the model enabled to estimate the maximum wall deflection for preliminary design. Neural networks also were used to evaluate slope stability (Sakellariou and Ferentinoue, 2005; Choobbasti et al., 2009). The input layer in ANNs consists of a vector of critical factors, and often the data sets of input factors are based on finite element methods (FEMs). As the limitation of the FEMs in computer software, some factors are limited to be considered as model inputs. Sometimes, the contribution of some events or induced-factors to the output cannot be involved in a finite element model. In light of this, a capability of integration is needed for a model in a global way. Song et al. (2012) analyzed the influence of environmental factors and predicted the probability of the landslide occurrence based on traditional BNs, and a case study for probabilistic assessment of tunnel excavation processes was studied using dynamic BNs (Spackova et al., 2013).

Geotechnical problems are frequently associated with sparse information (Beer et al., 2013). In the geotechnical practice, data from site investigation is scarce, uncertain, sparse, and monetizing (Phoon, 2017), which generates a big 'site challenge' as the model inputs (Phoon, 2018). Phoon (2019) pointed out that data is 'new oil' for us to predict the potential risk and learn decision strategies. However, the considerable data in a specific/regional scales is in demand for quantitative risk assessment in geotechnical engineering. 'Big data' and precise information are valuable and usually unachievable for researchers, and thus the big challenge in the current work is how to efficiently and effectively build a robust model for risk analysis and management with the limited information. In the context of this, we proposed advanced Bayesian networks for solving the geotechnical problems. Though a few studies have attempted to consider the network parameters with the imprecise information (Antonucci et al., 2004; Cardenas et al., 2014b; Zhang et al., 2016), there is still a lack of a detail framework for geotechnical risk assessment, providing the evaluation of the failure probability, uncertainty quantification and identification of sensitive parameters leading to the damage, and even, the support of the real-time information updating.

## 1.2 Motivation and Objectives of the work

As mentioned in the previous section, data from geotechnical engineering is more uncertain than those associated with structural or other engineering where materials are man-made and subject to stringent quality control. Geotechnical materials such as soils and rocks are natural. The volume of geomaterials is also larger. The typical size of a project site is perhaps one or several football fields in size. It is not economically possible to conduct tests over a dense grid for such large areas. Phoon (2017) referred to this situation as the “curse of small sample size”. In addition, Phoon (2020) pointed out that it is not sufficient to think of “uncertainty” in geotechnical engineering data. The author proposed that the features of geotechnical site data can be succinctly described as MUSIC: Multivariate, Uncertain and Unique, Sparse, InComplete, and potentially Corrupted. The “unique” and “incomplete” features have not received the attention they deserve in the literature, although they are surely present to different degrees in geotechnical databases. Site “uniqueness” is already well known in practice, but there is no quantitative method of dealing with this aspect. In fact, the importance of site-specific data is recognized in almost all national building regulations that mandate by law the minimum site investigation (say number of boreholes must be greater than 3) that has to be conducted at each project site.

There are two important ramifications arising from this “unique” site condition. It is useful to clarify here that when a site is said to be “unique”, it does not mean it is completely distinct from an adjacent site. It simply means it is not identical to other sites due to natural origin of geomaterials. The degree of uniqueness varies according to geology, but it is there. The first ramification is that site-specific (or local) data is more important than data from other sites. This also means that the “uncertain” feature is even more important. It is not easy to combine local data with other data to increase the sample size so that we can characterize uncertainty with conventional probabilistic models more reliably. It is clear that a conventional probability distribution and its parameters (mean, coefficient of variation, etc.) cannot be defined using just 10 data points for example. It will be possible to treat the mean, coefficient of variation, and other statistical parameters as random variables and account for the effect of small sample size using statistical uncertainty. Ching et al. (2016, 2017) and Ching and Phoon (2017) adopted this approach for random field parameters. It is also possible to characterize the statistical uncertainty for MUSIC (Ching and Phoon 2019 a, b) and MUSIC-X (Ching and Phoon, 2019c) data. MUSIC-X refers to multivariate cross-correlated MUSIC data that are also varying in space ("X" dimension). However, it is unclear how to quantify the statistical uncertainty in the shape of the probability distribution function. This thesis approaches the problem of small sample size at a single site using imprecise probability, which is probably a better fit to the limited data compared to the conventional probability approach with crisp models.

The second ramification is that since local data is so limited and precious, the question of how to optimize data collection becomes very important. There are numerous studies on how to optimize the sampling locations without a geotechnical structure in mind (Hight and Georgnnou, 1995; Goldsworthy et al., 2007; Cao et al., 2016; Wang et al., 2019) and with a geotechnical structure such as a slope in mind (Gong et al., 2014; Jiang et al., 2017; Li et al., 2019; Park et al., 2019; Yang et al., 2019;). This thesis does not consider spatial variation. Hence, it does not focus on sampling locations. However, it is also possible to optimize based on the type of input parameters one should spend more efforts in data collection. This is a sensitivity problem. The response of any geotechnical structure such as the factor of safety of a slope or the deflection of a retaining wall is sensitive to some input parameters. Clearly, if the budget is very limited, the engineer will want to know how to allocate the budget to better characterize these sensitive parameters. This thesis clarifies how to perform sensitivity analysis in the framework of imprecise probability. Some past studies have been conducted (Bi et al., 2019; Wei et al., 2019), but they have not been applied to geotechnical engineering problems.

Therefore, the purpose of this work is to present a robust model to fit the site-specific data as well as solve the stability and deformation problems in geotechnical engineering. A risk assessment methodology based on advanced Bayesian networks (BNs) is proposed to provide a framework for computing the failure probability of stability and deformation problems in geotechnical engineering, identifying the key risk factors and the real-time information updating, even in the case of incomplete information. The main objective is to develop a flexible and low-cost model for integrating the various risk factors and related events so that we can evaluate the probability of failure occurrence as well as capture the uncertainty from the geotechnical parameters, and besides, the real-time results can support the risk mitigation and management for the decision-makers.

The measurement of some geotechnical parameters is represented as continuous random variables and even interval variables when scarce data and limited knowledge on our problems can be achieved. However, because of the limitation of traditional BNs, the model requires the input nodes associated with discrete probabilities. Thereof, the model should enable to be extended to meet the requirement of the variety of probabilistic property. Considering the complexity of the failure mechanism of a slope in the unforeseen geotechnical condition, the methods enble to analyze the stability of slopes according to the different purposes. With the potential slip surface, a large number of methods based on limited equilibrium methods can be available to analyse the stability of slopes related to various shapes of the failure surface. As an alternative method, FEMs for the slope stability analysis provide the fewer priori assumption, the failure surface of a slope occurs naturally without the knowledge of the exact slip surface. The model should allow the slope stability analysis, including the different analysis methods according to the real-time need of research in the geotechnical practice. On the other hand, water plays an important role in soil stability. The influence of water on slope safety is also necessary to be considered with rainfall and groundwater level. Therefore, the powerful capability of integration in BNs is suitable for handling the unrelated and related factors or events into a model.

The uncertainty from risk factors or induced events in the geotechnical problem, has a great influence on the failure of the target event such as the variability of geotechnical parameters affects the failure event, and due to the spatial variability, soil in the field cannot be ideally homogeneous along the depth in practice, even if there is only one soil type. In view of this, field measurements are proposed to observe the variety in the practical geotechnical engineering. This requires the model updating approach involved in the network in order to reduce the uncertainty by updating the soil parameters. The uncertainty can be divided into aleatory and epistemic uncertainty, and the latter can be reduced by the better understanding of geotechnical parameters. In light of this, the identification of key factors is required for reducing the uncertainty associated with limited information. Generally, attempts were made in the thesis to enhance the robustness and flexibility of the BNs model, and further, extend the application range of the model in the geotechnical problems with the low computation cost and real-time way.

## 1.3 Original Contribution

The main work of this study is to provide a complete model for quantitative risk assessment with regard to site-specific data in geotechnical engineering. We investigated how to capture the uncertainty and evaluate the failure probability by selecting different models (BNs, enhanced BNs or Credal networks) according to the knowledge and/or data available. Further, in the absence of site-specific data, we integrated the advanced BNs model with the proposed sensitivity analysis, which the advanced model is feasible for quantifying the statistical uncertainty contribution of soil parameters to the risk consequence. In this work, we only considered how to identify the influence of input parameters, rather than spatial variability in each influential parameter. Then the information about measurement of uncertainty importance can support the decision-making in practice for risk management.

## 1.4 Structure of the Thesis

The thesis consists of six chapters, where a summarising essay appended with three research articles is presented. Specifically, the thesis is organized as follows:

Following this introduction, Chapter 2 introduces the theoretical support. The proposed methods are presented in detail, and finally, a summary on how to enhance the ability of BNs with these methods is presented before the cumulative part of the thesis. The work starts with preliminary research based on the stability problem of soil slopes, in the first research paper, an attempt is made for the application of advanced BNs in solving the slope problems. Two failure models of slopes are analyzed to predict the real-time probabilities of failure occurrence, and a proposed discretization approach for BNs is used for new observation on continuous parameters in the network. Further, three scenarios associated with different available information are discussed, where CNs are considered to capture the uncertainty propagation in the case of imperfect information.

In second research paper, we propose an updating framework for the BN model of braced excavation. The updating method proposed can effective to update the distribution parameters of soil parameters with monitor data. The accuracy of the estimation of geotechnical parameters affects the deformation of retaining wall and ground surface settlement most. Therefore, with the enhanced BNs model, the predication of failure risk in the construction process can be effectively reduced.

The quantification of the importance of the querying node to the target node in the network is especially useful for risk mitigation and remedy. On the basis of tradition BNs, the characteristic of variation of causal factors is captured by querying. However, for the BN with imprecise probabilities, this method is unreasonable to be used to check their sensitivity degree. In light of this, a new importance measure approach is proposed in the third research paper for the sensitivity analysis in the framework of imprecise probabilities. The method proposed can be used to quantify the effect of epistemic uncertainty from input network parameters on the target event. The results can support the decision-makers for the risk reduction.

Finally, conclusions from the studies in the former chapters are drawn in Chapter 6. The proposed model shows a better fit for the site-specific data, and can provide a new view for analyzing the risk from the stability and deformation problems in the geotechnical engineering. It satisfies the requirement of performing the model including generous amounts of a wide variety of variables and small probability events. The BN model is enhanced with the ability of identification and updating, making the estimation of stability and deformation problems more accuracy and powerful. The flexibility and low-cost computation of the proposed model shows an effective information support for the decision-making.

# 2 Theoretical background

## 2.1 Geotechnical risk assessment

Risk assessment is the foundation of risk mitigation and decision-making support. The main purpose of the risk assessment process is to analyze the cause and consequences of hazards in order to avoid the occurrence of hazards or minimize the level of the risk. As the definition in ISO 2394 (2015) states, the risk is the effect of uncertainties on objectives. Then how to assign a numerical value to risk is an important issue and attracted a great concern. Nowadays, from a quantitative point of view, two ways can be mainly used to calculate the risk. Generally, the overall risk is assigned a quantitative value using the following formula:

$$Risk = P_f \times Consequence \tag{2.1}$$

where $P_f$ represents the probability of failure, and the consequence of failure reflects the economic loss and/or casualties, which is computed as a deterministic value. The equation in Eq. (2.1) is often considered in the risk assessment of natural hazards, such as landslides (Remondo et al., 2008; Huang et al., 2013), where risk assessment consists of two modules, one needs to assess the slope safety and one is to assess the consequence.

The quantity of consequences is related to catastrophic scales, public property and the safety of human life, etc. In the study of Zhang and Huang (2016), they assessed the risk of a slope and considered the consequences of a slope failure with the volume of the sliding mass. Further, in the case of the various sources of risk from the landslides, more detail components of consequences (Dai et al., 2002) need to be divided to evaluate, respectively. For most of natural hazards, there are standard components of consequences used to evaluate the value of consequences. However, when the available information or/and data is quite limited, it is complex and challenging to analyze and evaluate all the consequences of a potential hazard. So, it is impossible to quantify the probabilities of failure, $P_f$ and consequences separately. This situation is very common in the most of human-made constructions, the consequence of hazards mainly depends on the degree of loss in economy and life, which the level of potential damage is hard to evaluate in the uncertain circumstance. Hence, instead of quantifying the consequence, a risk indicator, such as reliability index or the probability of failure, is usually used for the evaluation of risk as a whole. Afterwards, the evaluated risk is contrasted with an allowable risk (Fell, 1993, Duzgun and Lacasse, 2005) in order to identify the level of risk. Nowadays, a set of the technical norms has been established for the design of construction works. In this context, the evaluation of safety became a classical and the most significant part of the research of risk assessment. The study, in this thesis, mainly paid attention to the safety modules in the risk assessment, and further, the definition of consequences classes for the structural design of buildings and civil engineering works in European standard EN 1990 (2002) can be taken into account as a reference for the identification of the level of risk.

In the early works of risk evaluation (Alonso, 1976), the measure of safety in geotechnical profession was attempted to be conducted in the probabilistic terms, and later, it was widely used in solving the stability and deformation problems. Owing to the uncertainty involved in assessing the risk of geotechnical problems, probabilistic risk assessment has increasingly attracted great concern in geotechnical engineering. These works were conducted to estimate the probability of a slope/construction failure based on the various simulation methods. Ramly et al. (2005) proposed a probabilistic slope analysis methodology based on Monte Carlo simulation, where the reliability index and failure probability of the slope were estimated. Additionally, in the estimation of a slope/construction safety, various attempts have been made to quantify the uncertainty involved in risk (Finlay et al., 1999; Silva et al., 2008; Xia et al., 2017), where numerous methodologies such as deterministic analysis, probabilistic analysis, and statistical analysis are adopted in their works. Also, reliability analysis was introduced to analyze the geotechnical problems many decades ago (Whitman, 1984), and the method has been further employed to evaluate the failure probabilities in the geotechnical engineering (Cassidy et al., 2008; Zhang and Huang, 2016).

To avoid or reduce a hazard, it is important to know the failure mechanism of a hazard (e.g. why and where the potential failure occurs) and its risk source. Nevertheless, the unpredicted failure mechanisms prior to the research and the scarce data and/or knowledge for the causal factors of the risk, affecting the accuracy of estimating the failure probability. As for the stability and deformation problems of soil in the geotechnical engineering, it can be normally considered by researchers as a system that many failure modes coexist. The application of system reliability analysis in the process of risk assessment provide new insight into the effect of uncertainties in the geotechnical problems, such as the slope stability problem where reliability analysis is a popular tool to be used to evaluate the safety in a systemic term (Oka and Wu, 1990; Miro et al., 2015). For example, with system reliability analysis, multiple scenario failure events can be evaluated with modified FORMs (Cho, 2013). Furthermore, limit equilibrium (LE) analysis and finite element (FE) analysis, as two basic methods, are usually applied to model the behavior of geotechnical materials in the stability and deformation problems. The main purpose is to bring the slope/a construction to a state of limit failure mode. From a number of possible failure surfaces in the slope, the limit equilibrium methods are usually used to consider the occurrence of slope failure along a known slip surface.

By building the limited equilibrium function, the values of factors of safety can be computed to indicate the safe state of a slope (Li et al., 2005). Finite element methods, without assuming the failure shape or location of a slope or construction, mainly concentrate on the stability or deformation problems based on the overall shear failure. The failure occurs 'natually' when a critical state of the shear strength against failure reaches. This method allows a better understanding of the failure mechanism. FE techniques has been increasing for solving the slope stability problem for many years (Griffiths and Lane, 1999; Griffiths et al., 2011). Rather than compute factors of safety, the shear strength reduction approach is used to perform the FE slope analysis approach. This method is also widely applied in the stability analysis of tunnel constructions. Do et al.(2015) investigated four failure mechanisms of excavations and reasonably estimated the stability of the excavation by the FE method.

In the past decades, quantitative risk assessment methods in the geotechnical engineering have been a popular research topic (Lacasse and Nadim, 1988; Silva et al., 2008; Zhang and Huang, 2016). In the quantitative risk assessment, most of the works mainly focus on how to efficiently estimate the failure, capture the uncertainty involved and the identification and quantitative evaluation of the factors contributing to risk. The efficient risk assessment of the slope failure is a crucial precondition for making the rational strategy against disasters. Nowadays, quantitative risk assessment for the slope problems is mainly facing two key challenges: one is the spatial variability of soil properties, and the other one is multiple failure modes. The soil properties have an inherent variation in the different spatial areas, even in the homogeneous materials. Uncertainty from the spatial variability brings the challenge for evaluating the soil slope failure accurately. To model this spatial fluctuations of soil properties, random field theory has been a prevalent method (Vanmarcke, 1977; El-Ramly et al., 2002; Griffiths and Fenton, 2004; Cho, 2007), which allows soil variables to occur randomly in space. The concept of random fields combining with the LE method was employed in the risk assessment of slope failure (Jiang et al., 2017). Also, it was developed with the FE method to model the spatial variability of soil properties and called the random FE method (Huang et al., 2013). In the quantitative risk assessment of soil slope failure, Liu et al. (2019) considered the variation of soil properties in space by the random field theory, which was further incorporated into the FM method. Furthermore, the field observation based on monitoring program is another significant mean to record the response of soil parameters during construction and evaluate the state of slope safety. Thereof, the integration of multiple sources of data into the risk assessment in the geotechnical problems is necessary under the condition of available monitoring information.

On the other hand, the existence of multi-risk for a geotechnical problem is quite common, posing a range of challenges to capture the cause of different characteristics of hazards, and even when the risk is very low, then it is not possible to quantify the actual risk accurately and fully by analysis methods alone. Therefore, the single method is not sufficient for analyzing the risk in geotechnical problems. As stated above, reliability analysis of a slope or construction is an efficient method to evaluate the multiple failure modes, but a model also is needed to evaluate the multiple failure types for one or more slopes or constructions. In the risk assessment, each of failure associated with the corresponding consequences. Therefore, the equation of risk should be expressed as:

$$Risk = \sum_{1}^{n} P_{fi} \times Consequence_i \tag{2.2}$$

where $i = 1, ..., n$ and $n$ represents the number of identifiable failures.

Moreover, except the natural hazards that tend to be caused by extreme events, like heavy rainfall, storm or earthquake, many hazards are also induced by the other induced-factors. The sources of risk in geotechnical engineering are various. In light of this, the identification of critical induced-factors plays an important role in the risk assessment. In this context, a framework of risk evaluation is in demand to integrate the quantitative and qualitative information.

Bayesian machine learning is a promising tool to manage the uncertainties in geotechnical engineering (Phoon, 2019). As the risk of geotechnical problems can be regarded as system failure, event-tree analysis in geotechnical engineering was studied to provide a framework for the evaluation of the multi-risk concerning the uncertainty (Whitman, 2000). Recently, many studies have always analyzed the system failure with mapping the event-tree to the corresponding Bayesian network (BN), which is very popular in artificial intelligence (AI) field, and the conditional probabilities of an event in the network can be replaced by expert's 'beliefs'. The BN tool has much common with fault tree analysis (FTA), event tree analysis (ETA), and barrier analysis for accident analysis, but its powerful capability in the risk assessment shows much more advantages over these analysis approaches. The BN tool allows various data, functional relationship, expert knowledge and even incomplete information into the structure of the network, and the inference with the exact and approximate terms can be widely applied for the forward or backward reasoning. The real-time updating ability of BNs tool makes the risk assessment more robust and flexible to achieve and identify the failure state of a slope or constructions. The way of inserting the additional evidence into the network contributes to the structure learning as well as parameters learning, which they are also important parts in the machine learning. Also, a dynamic BN can provide the dynamic risk analysis considering the dynamic environment and time series.

In recent years, the BNs method has been applied to evaluate the risk in geotechnical engineering (Straub, 2005; Schubert et al., 2012; Wang et al., 2018). Straub (2005) demonstrated the potential and advantages of BNs for the application in risk assessment of natural hazards. In his work for investigating a rock-fall hazard, BNs showed a large potential of the detailed evaluation of the joint influence of the different indicators on the risk. For the diagnosis of embankment dam distress (Zhang et al., 2011a; Zhang et al., 2011b), they found the BNs tool could be efficient in both local and global consideration, and also, allowed the identification of the most essential distress causes with the project-specific evidence. Also, as a meta-modelling approach, BNs showed an excellent capability to capture the uncertainties in slope stability analysis (Cardenas, 2019). Besides, for assessing the risk for construction works, it was verified that the BN model was a powerful tool for synthesizing multiple sources to evaluate the risk, with which planners and engineers could systematically assess and mitigate the inherent risk (Sousa and Einstein, 2012; Cardenas et al., 2012; Spackova, 2013; Wu et al., 2015). The application of BNs in geotechnical risk assessment shows an increasing tendency.

However, some limitations of BNs hinder the application of the BNs approach for solving the geotechnical problems. In light of this, this work mainly focuses on coping with the obstacle in the implementation of the BN model for estimating the risk in geotechnical engineering. Contributions of our work and detail introduction for the enhanced BN tool can be found in three research articles. In general, we propose a framework for analyzing the risk in the geotechnical problems with advanced BNs. In our works, some problems of geotechnical engineering associated with imperfect information are analyzed with the proposed model, and the same time, the ability of BNs model is enhanced with the integration of other methods. Additionally, the enhanced BNs model is illustrated that the advanced model risk assessment can efficiently address the uncertainty, and provide useful information for the decision-makers. A brief introduction of the applied methods is presented in the following section.

## 2.2 Bayesian network

BNs also called Bayesian belief networks or Bayes nets, was first proposed by Pearl (1988). It becomes a popular tool with combing the powerful probability theory, graphical theory and computer science. With a series of sophisticated exact and approximate inference algorithms, the model is used widely in the field research as reasoning, diagnostics, and decision-making tool under uncertainty and time series prediction. The application of the BN model in the risk analysis can be traced to the year of 2001. Hudson et al. (2001) investigated the decision support system in military installations with BNs. Then risk analysis using BNs were widely applied in the environment (Sperotto et al., 2017), economics (Cornalba and Giudici, 2004) and engineering professions (Morales-Napoles et al., 2014), etc. In terms of probabilistic risk assessment, Smith (2006) analyzed the risk of a dam with the discrete BN model. Sousa and Einstein (2012) made the risk assessment during tunnel construction, and a BN model was built to combine the domain knowledge. Mitra et al. (2018) used the BN approach to assess the influences of the factors in landslide risk assessment. Most of the researches were implemented using the mature method of BNs in risk assessment. In real practice, the limitations of traditional BN can hinder the construction of BN model. To efficiently quantifying uncertainties in geotechnical risk assessment, it is not enough to only rely on the existing

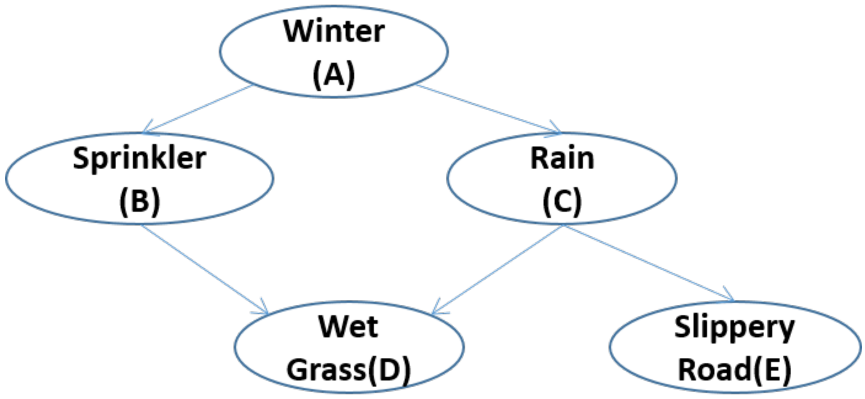



Figure 2.1: An example of the Bayesian network

BN approach. According to risk analysis in the problems of slope stability and the deformation of pit excavation, we sufficiently applied the potential ability for solving the geotechnical problems and as well, we developed the ability of the BN model for coping with the uncertainties in geotechnical engineering in the framework of imprecise probabilities.

The mature method of the developed BN model: traditional BNs and the developed BNs with imprecise probability is introduced briefly with a simple example in the next section.

The Bayesian network (BN) is a probabilistic multivariate model that can be used to build the model from data and/or expert opinions. The structure of a BN is defined by two parts: a set of nodes and a set of directed arrows. The arrow from node $a$ to $b$ (in Figure 1) represents the conditional dependence among the two nodes. Figure 1 shows a causal graphical network, where the relationship of linked nodes is termed as parents-children, such as the node $D$ is the child of nodes $B$ and $C$. Assume that all the nodes have tow states: True ($T$) and Failure ($F$). For the quantification part of the model, the nodes are described in a manner of the probability distribution. Initially, the nodes in a BN are described as discrete random variables, and therefore the network also called discrete BNs. The quantification of each node is represented by a conditional probability table (CPT), denoted by $\Phi$ in the model. A node is conditionally independent of its ancestors given its parents, such as the CPT of node $B$ in Figure 1 are $\Phi_B$: $P(B|A)$. In addition, a CPT is composed of some point probabilities related to the states of a node, where all the possible states of each variable conditionally depend on its parents in the network. Further, a factorisation of the joint probability distribution over the set of variables in a BN can be written by

$$P(A, B, C, D, E) = \bigcap_{i=1}^{5} \Phi_i = P(A)P(B|A)P(C|A)P(D|B, C)P(E|C) \tag{2.3}$$

The efficient inference and learning the network can be conducted by querying. If we assume the evidence: $A = T$, $C = F$ in the network, the result of interest is the probability of occurrence of a slippery road. According to Bayes' rule for variables,

$$P(E|A = T, C = F) = \frac{P(E, A = T, C = F)}{P(A = T, C = F)} \tag{2.4}$$

where

$$P(E, A = T, C = F) = \sum_{B,D} P(A = T)P(B|A = T)P(C = F|A = T)P(D|B, C = F)P(E|C = F) \tag{2.5}$$

and

$$P(A = T, C = F) = \sum_{B,D,E} P(A = T)P(B|A = T)P(C = F|A = T)P(D|B, C = F)P(E|C = F) \tag{2.6}$$

Note that the applicability of BNs would be enormously limited if one would only consider discrete variables. A significant development of discrete BNs is to make both continuous and discrete variables coexist in the network, called hybrid BNs. Lauritzen (1992) proposed Linear conditional Gaussian BNs, but the method limited the structure of BNs about the order among the variables, which each continuous variable must follow a linear Gaussian distribution conditional on the configuration of its discrete parent variables. The attempts of using the mixtures of truncated exponential (MTE) distributions (Moral et al., 2001; Cobba and Shenoyb, 2006) and mixture of polynomials (MOP) (Shenoy and West, 2011) approximations was made for building the hybrid BNs in order to avoid the restriction of structure. The way of approximating the conditional probability distribution functions (PDFs) of continuous variables by MTE or MOP leads to the unavoidable loss of accuracy. In light of this, an enhanced BNs approach was proposed by Straub and Der Kiureghian (2010). The model combines with the concept of structural reliability methods (SRMs). The model proposed enables practical computation of continuous nodes with stochastic distributions and marginal computation without approximation. The main concept of this enhanced BN method is to simplify the hybrid BN by removing all the continuous nodes from the original model by means of SRMs. Precisely, SRMs erase the links between continuous nodes and their discrete children (the so-called deterministic nodes). Thus, they become barren nodes (a barren node has neither evidence nor children), enabling them to be removed without altering the CPTs of their offspring.

Another important development of BNs is to allow the model including the theory of imprecise probability. Credal networks (CNs), as a novel class of imprecise probability graphical models, are first introduced by Cozman (2005). The model is an extension of BNs, which represents a generalization of Bayesian networks to handle imprecise and incomplete information. Instead of the definition with the single probability distribution for a random variable $X_i$ in the network, the node is denoted with a credal set, which is defined by a set of probability densities $P(X_i)$ $(i = 1, 2, ..., n)$, indicated by $K(X_i)$,

$$K(X_i) = CH\{P(X_i)|P(X_i) = \bigcap_{i=1}^{n} P(X_i|pa(X_i))\} \tag{2.7}$$

where $CH$ denotes the convex hull operator and $pa(x_i)$ is the assignment to the parents of $X_i$

Then the joint mass function $P(X) = P(X_1, X_2, ...X_n)$ over all the variables in the Credal network is given by

$$P(X) = \bigcap_{i=1}^{n} K(X_i|pa(X_i))\} \tag{2.8}$$

The calculation of the marginal probabilities of variables is based on the joint credal set definition to calculate the bounds of each node. The inference for CNs is more complicated than for BNs, still being in its infancy stage of development. Currently, some exact and approximate inference algorithms (Fagiuoli and Zaffalon 1988; Cozman 2005; Alessandro and Zaffalon 2008; Maua et al. 2014; Tolo et al. 2018) have been studied for the reasoning of CNs. These developed BNs are uniformly called advanced BNs in this thesis.

Based on the advanced BNs, we enhanced the capability of the model mainly on the sensitivity and information updating in the framework of imprecise probabilities. To be specific, we combine the BNs approach with the proposed methods, which they are introduced as the following sections shows.

## 2.3 Bayesian model updating

Model updating with dynamic test data becomes extremely imperative for reasons of the discrepancy between a theoretical model and the practical behaviour of the real system. As the uncertainty of model parameters affects the accuracy of the response prediction of the model output, the updated model could reduce the uncertainty of the predicted response. Generally speaking, model updating is a process of calibrating the model parameters based on the actually observed data of the real system, which is also inverse problems.

Due to the limited test data, measurement error and/or lack of knowledge, uncertainty in model updating is a non-negligible issue. Over the past several decades, quantification (UQ) implemented in model updating, has been studied using Bayesian statistical framework (Beck and Katafygiotis, 1998; Papadimitriou et al., 2001; Beck, 2010; Simoen et al., 2015; Bi et al., 2018). Model updating with Bayesian theory, namely Bayesian model updating (Beck and Au 2002; Ching and Chen 2007), addresses the difficulties from the ill-conditioned inverse problems (Beck and Katafygiotis 1998), and enables to consider a class of models with the uncertainty quantification.

Let $D$ be the measured data of real system, and $\theta$ be the uncertain parameters of model inputs. Based on Bayes' theorem, the updated probability distributions of model parameters are given by

$$P(\theta|D) = \frac{P_L(D|\theta)P(\theta)}{P(D)} \tag{2.9}$$

where $P_L(D|\theta)$ is the likelihood of obtaining the data $D$ given the model parameters $\theta$, and $P(\theta)$ is the prior distribution of model parameters. $P(D)$ is the probabilities of observing $D$ independently, which is termed as the normalizing constant, ensuring that the posterior PDF integrates to one: $\sum_\theta P(D|\theta)P(\theta)$.

The challenging components of the Bayesian model updating are the normalizing constant and the likelihood, as it is difficult to evaluate them directly. In light of this, stochastic simulation methods are used to tackle this problem, and the goal of using Simulation-based methods is to obtain samples from the posterior distribution based on the given data and the prior distribution. Among simulation methods, the Markov Chain Monte Carlo (MCMC) approach have been developed and successfully applied in the Bayesian model updating (Beck and Au, 2002; Wu and Chen, 2009; Straub and Papaioannou, 2015) owing to the ability to compute the high-dimensional integrals. In addition, The Transitional Markov Chain Monte Carlo (TMCMC ) along with Metropolis-Hasting algorithm is used as an effective updating tool (Ching and Chen, 2007). Comparing with the MCMC simulation, the TMCMC approach with the adaptive Metropolis-Hasting (AMH) algorithm implemented has the advantageous of coping with the complex PDFs, such as multimodal PDFs and very peaked or flat PDFs. The method allows sampling from a series of intermediate PDFs till converge to the target PDF, instead of direct sampling from the difficult target PDF in the MCMC simulation. For the MCMC-based Bayesian updating, the sampling algorithm of posterior distribution is considered with the following equation,

$$P(\theta|D) \propto P_L(D|\theta) \cdot P(\theta) \tag{2.10}$$

As stated, sometimes, sampling from $P_L(D|\theta)$ can be difficult in some sense. Therefore, a series of intermediate PDFs are constructed from the prior PDF $P(\theta)$ up to the convergence:

$$P_j \propto P_L(D|\theta)^{\alpha_j} \cdot P(\theta) \tag{2.11}$$

where the exponent of the likelihood $\alpha_j \in [0,1]$, and $j$ denotes the step number of convergence. Generally speaking, the TMCMC algorithm consists of a set of resampling stages, and the sampling from $P_L(D|\theta)$ and the estimation $P(D)$ are accomplished in the process of the TMCMC simulation with the AMH algorithm (Beck and Au, 2002). A brief summary of the simulation procedure is introduced as follows,

(1) At first stage $j = 0$ of the convergence, $N_s$ samples: $\theta_{(0,1)}$,...$\theta_{(0,s)}$ draws from the prior distribution $P(\theta)$.
(2) Find $q_{j+1}$ by the samples of the previous level:

$$q_{j+1} = argmin(|COV(q_j) - V_{threshold}|) \tag{2.12}$$

where $COV(q_j)$ denotes the coefficient of variation of the plausibility weights, and $V_{threshold}$ is a prescribed threshold, and is usually defined as 100%. Afterwards, for all the samples, computing the weighting coefficient $w(\theta_{j,k})$,

$$w(\theta_{j,k}) = P_L(D|\theta_{j,k})^{\alpha_{j+1}-\alpha_j} \tag{2.13}$$

where the weights $w(\theta_{j,k})$, for $k = 1, 2, ..s$, attached to each Markov chain and its value denotes how close $q_{j+1}$ to $q_j$, and compute its mean,

$$S_j = \frac{1}{N_s} \sum_{k=1}^{N_s} w(\theta_{j,k}) \tag{2.14}$$

(3) Do the $N_j$ sampling based on MCMC using the AMH algorithm (Beck and Au, 2002) until the target numbers of samples is reached, where $S = \sum_{j=0}^{n} S_j$ is asymptotically unbiased for the evidence $P(D)$.

On the other hand, the likelihood function $P_L(D|\theta)$ is normally defined based on observations given the certain parameters $\theta$,

$$P_L(D|\theta) = \prod_{i=1}^{N_{obs}} P(d_k|\theta) \tag{2.15}$$

where $P(d_k|\theta)$ is the probability density value at $d_k$ under the given parameters $\theta$.

However, sometimes the computation of the likelihood function is quite expensive or even impossible. In this context, approximate Bayesian computation (ABC) (Beaumont et al., 2002; Beaumont, 2010; Prangle, 2017) has been studied and applied in the Bayesian updating framework since the last decades. The ABC method is approximate inference methods which replace the computation of the likelihood function $P_L(D|\theta)$ with a simulation of the model, where the distance between observed data and simulated data plays an important role in the computation. The basic ABC algorithm is

1. Sample $\theta^*$ from the prior distribution $P(\theta)$.
2. Simulate a data set $D*$.

3. Compute the discrepancy between the observed data and simulated data: $d(D^*, D)$.
4. Accept $\theta^*$ if $d(D*, D) \leq \epsilon$;

$\epsilon$ denotes the threshold of the discrepancy, and the smaller $\epsilon$ the closer simulations are to reality. The Euclidean distance becomes particular popular with the advantages of efficiency and simpleness, is widely employed as the distance metrics in the ABC (Toni et al., 2009; Prangle, 2017). Bi et al., (2018) proposed a stochastic model updating framework using the Bhattacharyya distance as the distance metric in the ABC. The choice depends on the specific situation.

## 2.4 Global sensitivity analysis

Initially, sensitivity analysis (SA) is known as a local approach, which the effect of small variation from inputs on the model output is studied. Since the late 1980s, the methods of SA have been developed with studying the wider range of the variation of inputs. Instead of the limited in calculating or estimating the partial derivatives of the model at a specific point, the developed SA allows taking into consideration the entire input distribution, which is so-called Global SA.

In the context of risk and reliability analysis, a variety of computational models are built to simulate the practical problems. Uncertainties in model inputs gain a great concern, which the impact on the variation of model output affects the results of decision analysis problems. Epistemic uncertainty in model inputs, such as because of the poor knowledge from the real-world or scarce data, can be reduced. Therefore, many researchers use the global SA approach to measure the impact of uncertainty of variation in model input on the model output. Saltelli et al. (2004) defined the sensitivity analysis in the uncertainty view of point: The study of how uncertainty in the output of a model (numerical or otherwise) could be apportioned to different sources of uncertainty in the model input. There are a family of indicators of global SA (Sobol, 1993; Helton, 2000; Wei et al., 2013) being studied to determine which of the input variables influence the modal output most in their whole uncertainty ranges.

### 2.4.1 Variance-based sensitivity analysis

In this thesis, we refer to two methods of global SA. One is the variance-based sensitivity analysis (Saltelli et al. 2010), which the sensitivity analysis is based on the sole variance of the model output. Cukier et al.(1973) were first proposed to compute the sensitivity indices by the first-order effects, and then extended to the higher-orders effects by the decomposition functions (Iman and Hora, 1990; Saltelli, 2002).

Given a model $Y = f(X)$, the decomposition of the model output can be written as (Sobol, 1993; Saltelli et al., 2010):

$$f(X) = f_0 + \sum_i f_i(X_i) + \sum_i \sum_{j>i} f_{ij}(X_i, X_j) + ... + f_{1,2,...,n}(X_1, ..., X_n) \tag{2.16}$$

where $f_0$ is constant, $f_i(X_i)$ denote the first-order functions, and $f_{ij}(X_i, X_j)$ are the second-order functions, etc. Further, each component can be computed in terms of conditional expectations of the model output. That is,

$$f_0 = E(Y) f_i(X_i) = E(Y|X_i) - E(Y) f_{ij}(X_i, X_j) = E(Y|X_i, X_j) - f_i(X_i) - f_j(X_j) - f_0 \tag{2.17}$$

and likewise for the higher orders. The variance of model output $Y$ can be decomposed in the following (Sobol, 1993; Borgonovo, 2007),

$$V(Y) = \sum_i^n V(V_i) + \sum_{i<j} V_{i,j} + ... + V_{1,2,...n} \tag{2.18}$$

where $V_i$ represents the contribution of single variable $X_i$ to the output $Y$, and $V_{i,j}$ is the interaction effect of $X_i$ and $X_j$ on the model output. Then we can obtain the associated sensitivity measure, and the first-order sensitivity index can be written as:

$$S_i = \frac{V(E(Y|X_i))}{V(Y)} \tag{2.19}$$

the first-order index $S_i$ denotes the main contribution of each inputs to the variance of the output.

Another popular sensitivity index is the second-order effect,

$$S_{ij} = \frac{V_{i,j}}{V(Y)} \tag{2.20}$$

where $V_{i,j}$ represents the joint effect of $X_i$ and $X_j$, and In views of aforementioned, its decomposition is

$$V_{i,j} = V(f_{ij}(X_i, X_j)) = V(E(Y|X_i, X_j)) - V(E(Y|X_i)) - E(Y|X_j) \tag{2.21}$$

and similarly for the higher orders effect.

### 2.4.2 PDF-based sensitivity analysis

Another important global SA method: Moment-independent approaches (Park and Ahn, 1994; Chun et al., 2000; Borgonovo, 2007; Borgonovo et al., 2011; Wei et al., 2013), regarded as the PDF-based SA. The sensitivity measure depends on the variation of the entire distribution of the model output. Given $f_Y(y)$ is the unconditional probability density of the output $Y$, $f_{Y|X_i}(y)$ denotes the conditional probability density of $Y$ given one of input variable $X_i$ is fixed at a value $x_i^*$. Following the definition of sensitivity indicator $\delta_i$ (Borgonovo, 2007),

$$\delta_i = \frac{1}{2}E(s(X_i)) \tag{2.22}$$

where $\delta_i$ represents the normalized expect the shift from an unconditional probability distribution to the conditional probability distribution, which the shift is measured by the area $s(X_i)$,

$$s(X_i) = \int |f_Y(y) - f_{Y|X_i}(y)| \tag{2.23}$$

where the values of variable $X_i$: $\mathbf{x} = (x_1, x_2, ..., x_n)$ cover the whole range, and the expect shift can be written as

$$E(s(X_i)) = \int f_{X_i}(x_i)(\int |f_Y(y) - f_{Y|X_i}(y)|)dx_i \tag{2.24}$$

where the marginal PDF of input variable $X_i$ can be computed by

$$f_{X_i}(x_i) = \int ... \int f_X(\mathbf{x}) \prod_{m=1, m\neq i}^{n} dx_m \tag{2.25}$$

and the sensitivity measure of a group of input variables is computed in the same way.

## 2.5 Imprecise stochastic model

In practice, the available data and our knowledge for the real-world problems are often quite limited and imperfect. It will be insufficient with poor information to build the precise model. Therefore, the input parameters of a model appear the uncertain variation due to the vague, scarce or linguistic information, and so on. In the case of this, the input parameters with imprecise information should be quantified with the theory of imprecise probabilities (Beer et al., 2013), such as evidence theory (Shafer, 1976), interval probabilities (Weichselberger, 2000), and p-boxes (Scott et al., 2003) and so on. Moreover, a series of sampling methods has been developed for imprecise probability applications. Monte Carlo simulation (Fetz and Oberguggenberger, 2016), advanced line sampling (Angelis et al., 2014), Subset simulation (Alvarez et al., 2018) are enhanced to be applied in estimating the lower and upper bounds on the probability of small failure associated to the imprecise inputs. Among these simulation methods, an extended Monte Carlo simulation (EMCS) procedures proposed by Wei et al. (2014), can be employed for any black-box models in the parametric global SA. This so-called global EMCS, is the extension of EMCS, combining with the method of the Random Sampling-high dimensional model representation (RS-HDMR) (Li et al., 2017), and is applied to global SA in the imprecise probability models. The procedure of this non-intrusive imprecise stochastic simulation (Wei et al., 2019) is presented in the next section.

### 2.5.1 Global extended Monte Carlo simulation

Assume a model response function : $y = g(\mathbf{x})$, and $\mathbf{x}$ represent a group of input variables. For an parameterized imprecise model, the distribution parameters of input variables $\mathbf{x}$ are defined as $\mathbf{x}$. Let the failure domain F=$\mathbf{x}$: g($\mathbf{x}$<0), and $I_F(\theta)$ denotes the indicator of the failure domain. Then we have

$$\begin{cases} E(y|\theta) = \int g(\mathbf{x}) f(\mathbf{x}|\theta) dx \\ V(y|\theta) = \int g^2(\mathbf{x}) f(\mathbf{x}|\theta) dx \\ P_f(y|\theta) = \int I_F(\mathbf{x}) f(\mathbf{x}|\theta) dx \end{cases} \tag{2.26}$$

where $E(y|\theta)$, $V(y|\theta)$ and $P_f(y|\theta)$ are functions of expectation, variance and failure probability for the output $y$ given the uncertain parameters $\theta$, respectively. According to the derivation in the literature (Wei et al., 2014; Wei et al., 2019), the unbiased global EMCS estimators for the above functions can be expressed as,

$$\begin{cases} \hat{E}(y|\theta) = \frac{1}{n} \sum_{s=1}^{n} g(\mathbf{x}^s) \frac{f(\mathbf{x}^s|\theta)}{f(\mathbf{x}^s|\theta^s)} \\ \hat{V}(y|\theta) = \frac{1}{n} \sum_{s=1}^{n} g^2 \frac{f(\mathbf{x}^s|\theta)}{f(\mathbf{x}^s|\theta^s)} \\ \hat{P}_f(y|\theta) = \frac{1}{n} \sum_{s=1}^{n} I_F(\mathbf{x}^s) \frac{f(\mathbf{x}^s|\theta)}{f(\mathbf{x}^s|\theta^s)} \end{cases} \tag{2.27}$$

where $(\mathbf{x}^s, \theta^s)$ for $s = 1, 2, ..., n$, are a joint sample set generating from the joint PDF $f(\mathbf{x}, \theta)$. In a global sense, the sample points of $\theta$ spread over their entire range. Furthermore, to overcome the high dimension problems, the RS-HDMR approach is considered to improve the performance.

### 2.5.2 Random sampling HDMR method

In this section, the RS-HDMR method is presented to improve the performance of the GEMCS estimators in the high dimension problems.

Take the first-order model response function $E(y|\theta)$ for instance, the decomposition can be written as,

$$E(y|\theta) = E_0 + \sum_{i=1}^{d} E_i(y|\theta_i) + \sum_{i<j} E_{ij}(y|\theta_{ij}) + ... + E_{12...d}(y|\theta) \tag{2.28}$$

where subscript of $\theta$ denotes the dimension vector.

The unbiased GEMCS estimater of the RS-HDMR component functions can be expressed as follow (Wei et al., 2019):

$$\begin{cases} \hat{E}_0(y|\theta) = \frac{1}{n}\sum_{s=1}^{n} g(\mathbf{x}^s) \\ \hat{E}_i(y|\theta) = \frac{1}{n}\sum_{s=1}^{n} g(\mathbf{x}^s) r_i(\mathbf{x}^s|\theta_i, \theta^s) \\ \hat{E}_{ij}(y|\theta) = \frac{1}{n}\sum_{s=1}^{n} g(\mathbf{x}^s) r_{ij}(\mathbf{x}^s|\theta_{ij}, \theta^s) \end{cases} \tag{2.29}$$

where

$$\begin{cases} r_i(\mathbf{x}^s|\theta_i, \theta^s) = \frac{f(\mathbf{x}^s|\theta_i, \theta^s_{\sim i})}{(\mathbf{x}^s|\theta^s)} \\ r_{ij}(\mathbf{x}^s|\theta_{ij}, \theta^s) = \frac{f(\mathbf{x}^s|\theta_{ij}, \theta^s_{\sim ij})}{f(\mathbf{x}^s|\theta^s)} - \frac{f(\mathbf{x}^s|\theta_i, \theta^s_{\sim i})}{(\mathbf{x}^s|\theta^s)} - \frac{f(\mathbf{x}^s|\theta_j, \theta^s_{\sim j})}{(\mathbf{x}^s|\theta^s)} + 1 \end{cases} \tag{2.30}$$

and similarly for the second-order function and failure probability function.

## 2.6 The theoretical summary for the research articles

Based on the aforementioned methods, we enhanced the BN model and successfully applied the proposed model in solving the geotechnical problems. The achievements of this work have already been partly published in the scientific journals, which are presented in the following chapters.

Generally speaking, in **Research article I**, we cope with the prediction of slope failure by using the advanced BNs. The influence of internal and external triggering-factors are considered into the model. A integration of an infinite slope and a finite slope are proposed with an enhanced BN model. Two methods of slope stability analysis are discussed with consideration of the failure mechanism. With the model, the failure probabilities of slopes are estimated, and the real-time information updating is achieved as well. For observation being given on the continuous variables, a discretization approach is introduced in the hybrid BN. In addition, Credal networks are present for solving the slope problem in the case of the limited information, and a comparison is conducted with the information availability.

**Research article II** proposes a BN model for evaluating the safety state based on the monitored data information during the process of braced excavation, where a novel distanced-based Bayesian model updating method is applied contributing to improving the updating capability of BNs. With the enhanced model, the distribution parameters of soil parameters are effectively updated by using the data from the field observation. The response values of geotechnical structure due to input soil parameters are quantified with the method of 3-Dimensional finite element analysis, which is conducted in finite element package ABAQUS 6.13. Afterwards, a black-box is built with the ANNs, which can be embedded in the BN model, inducing the BN model from the databases.

**Research article III** copes with the sensitivity measurement for the advanced BNs and a case study of the slope stability is analyzed with the proposed method. The sensitivity propagation among the variables for a BN model is explored only in the directed path. The sensitivity analysis for the BNs model in the framework of both precise information and imprecise information is studied. For the precise BNs, we apply the PDF-based sensitivity analysis method to quantify the importance of causal nodes to their linked node, showing their dependency relationship in the network. On the other hand, considering the BN model with imperfect information, the non-intrusive imprecise stochastic simulation method combing with the global sensitivity analysis is used to measure the effect of epistemic uncertainty of input parameters on the slope stability.

# 3 Research article I: Failure Analysis of Soil Slopes with Advanced Bayesian Networks

## 3.1 Declaration of my contribution

I carried out all the related BNs model computations and analyses included in this article. I wrote the manuscript of this paper, and the idea of this manuscript came from the discussion with the co-authors, and then the analysis of slope problems with advanced BNs approach was done by myself. The case studies of slope problems were provided by co-author: Pro. António Topa Gomes from University of Porto, whose research background is geotechnical engineering. Further, the article was improved after the review process.

## 3.2 Published article

This paper is published in Periodica Polytechnica Civil Engineering.

Received: 27 March 2019, Accepted: 28 May 2019, Published online: 20 June 2019.

**Longxue He, António Topa Gomes, Matteo Broggi, and Michael Beer**. Failure Analysis of Soil Slopes with Advanced Bayesian Networks. *Periodica Polytechnica Civil Engineering*. doi: 10.3311/PPci.14092.

## 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 University Hannover, Callinstr.34, D-30167, Hannover, Germany
[2]CONSTRUCT, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, s/n 4200-465, Porto, Portugal
[3]Institute for Risk and Uncertainty, University of Liverpool, Peach Street, Chadwick Building, Liverpool, L69 7ZF, UK
[4]CERSM, Tongji University, 1239 Siping Road, 200092 Shanghai, China
*Corresponding author, e-mail: beer@irz.uni-hannover.de
Received: 27 March 2019, Accepted: 28 May 2019, Published online: 20 June 2019

## 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 (Ma et al., 2017). 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. (2017) 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 (Rahardjob et al., 2010; Tsai et al., 2013). 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 (Oka and Wu, 1990; Liu et al., 2018). 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) (El-Ramly et al., 2002) and simulation approaches, such as Monte Carlo Simulation (Metya et al., 2017), Importance Sampling (Metya et al., 2009) and Subset Simulation (Wang et al., 2011). 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 (Sakellariou and Ferentinou, 2005; Chakraborty and Goswami, 2017). 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 (1988). 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 (2005).

- 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 (Pearl and Russell, 2000).
- 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 (Song et al., 2012; Liu et al., 2013; Peng et al., 2013). 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 (Weber et al., 2012).

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,

$$P(X_i) = \prod_{i}^{n} P(X_i | pa(X_i)) \tag{1}$$

where $X_i = X_1, ..., X_n$ denote the nodes of the BN, $pa(X_i)$ are the set of parents of $X_i$, and $P(X_i|pa(X_i))$ 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 $X_j = e$, this information will propagate through the prior probabilities to the posterior probabilities as follows,

$$P(X_i|e) = \frac{P(X_i, e)}{P(e)} = \frac{\prod_i^n P(X_i|pa(X_i), e)}{\sum_{X_i \backslash X_j} P(X_i, e)} \tag{2}$$

note that the joint distribution $P(X_i, e)$, can be 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 $X_j$. 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.

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-

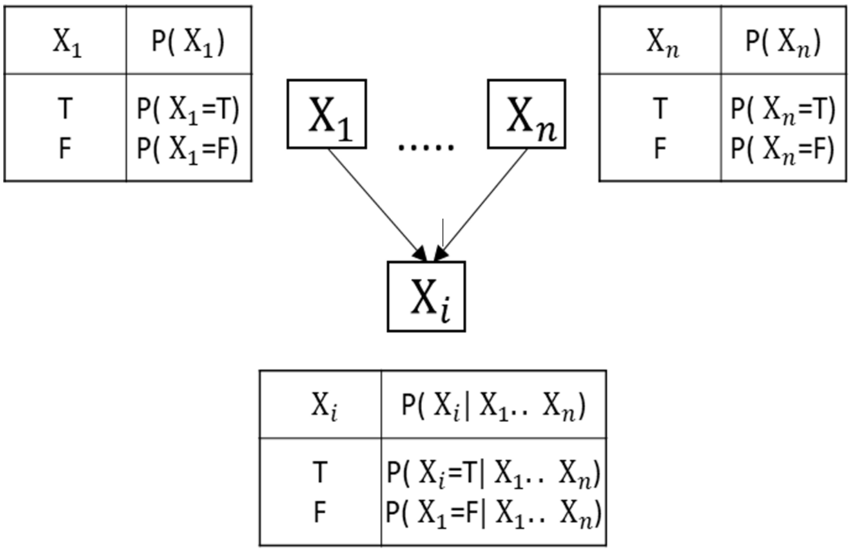



Fig.1 A simple graph of a general BN (T=True; F=False)

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 (Langseth et al., 2009). The final methodology is eBNs, implemented by joining BNs with SRMs, and was successfully applied in risk and reliability analysis by Straub and Kiureghian (2010a). An introduction to this method is given in detail in the following section.

## 2.2 Enhanced Bayesian Networks

Enhanced BNs approach (Straub and Kiureghian, 2010b) 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 (Straub and Kiureghian, 2010b). 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. 1. From of Eq. (1), the joint probability of all the nodes for the eBN can be written as

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_4|X_2, X_3)f(X_3)f(X_2|X_1) \tag{3}$$

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 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:

$$P(X_1, X_4) = P(X_1) \int\int_{\Omega_{X_4}(X_2,X_3)} f(X_3)f(X_2|X_1)\, dX_2\, dX_3 \tag{4}$$

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 Equation (**??**) are in line with the definition of structural reliability problems, and hence can be estimated by means of SRMs.

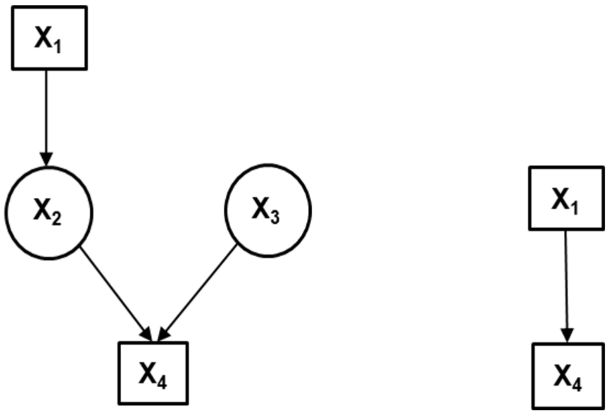



Fig.2 An example of reduction of an eBN into BN (circle represents continuous node and rectangle represents discrete node)

## 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 (Dougherty et al., 1995; Kurgan and Cios, 2001; Chen et al., 2017). 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 (Straub and Kiureghian, 2010b) 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 $X_{3continuous}$.

$X_{3\,discrete}$ has $i$ states that are defined by the outcome space of $X_3$ with conditional cumulative distribution function $F_{X_3}[x_3]$, 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 $X_{3\,discrete}$ given the state $i$ can be achieved as,

$$P(X^i_{3\,discrete}) = F_{X_3}[\overline{x_{3i}}] - F_{X_3}[\underline{x_{3i}}] \tag{5}$$

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

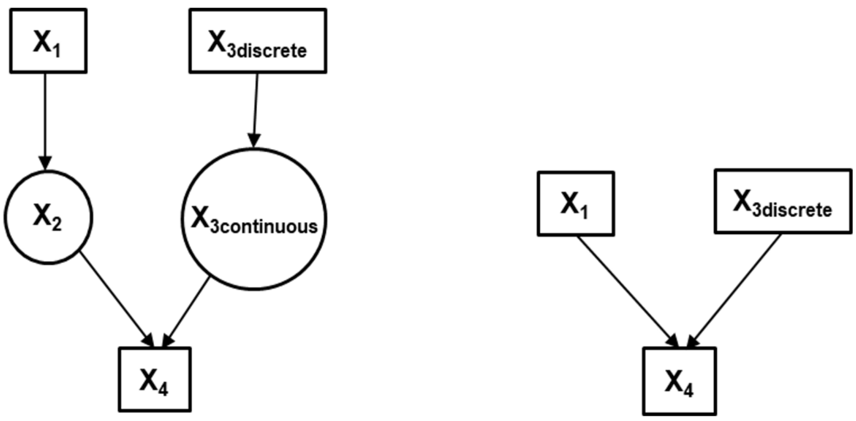



Fig.3 An example of the discretization procedure

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 number of the sub-domains on the observed continuous node is defined with the same length (Straub and Kiureghian, 2010b).

## 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 (1988) 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 (Cozman 2000).

The inference for CNs is more complex than for BNs, still being in its infancy stage of development (Tessem, 1992; Ferreira da Rocha and Cozman, 2002; Antonucci et al., 2015). 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 (Tolo et al., 2018) 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 (De Angelis et al., 2014) is considered herein. It is a recently developed advanced Monte Carlo methods, based on line sampling (Koutsourelakis, 2004) 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

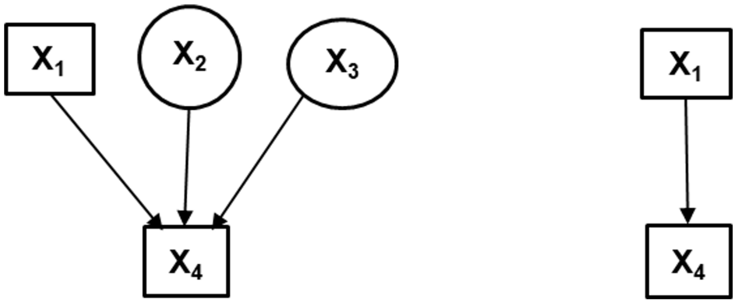



Fig. 4 An example of elimination procedure in a CN (Circle, rectangle, ellipse denotes continuous, discrete and imprecise node, respectively)

probabilities. Afterwards, exact inference for BNs such as the variable elimination algorithm (Pearl and Russell, 2000), 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,

$$P(X_1, \underline{\overline{X}}_4) = P(X_1)P(\underline{\overline{X}}_4|X_1) \tag{6}$$

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,

$$\begin{aligned} P(\overline{x}_{11})_{exact} &= \max\left(\textstyle\sum_{\underline{\overline{X}}_4} P(X_1)P\left(\underline{\overline{X}}_4|X_1\right)\right) \\ &= \max\begin{bmatrix} P(x_{11})P(\underline{x}_{41}|x_{11}) + P(x_{11})P(\overline{x}_{42}|x_{11}) \\ P(x_{11})P(\overline{x}_{41}|x_{11}) + P(x_{11})P(\underline{x}_{42}|x_{11}) \end{bmatrix} \end{aligned} \tag{7}$$

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.

A novel algorithm has been introduced to avoid this combinatorial explosion encountered by exact inference (Tolo et al., 2018). 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 (Tolo et al., 2018).



# 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 geotechnical 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. Furthermore, the nodes *Unsaturated Unit Weight*,

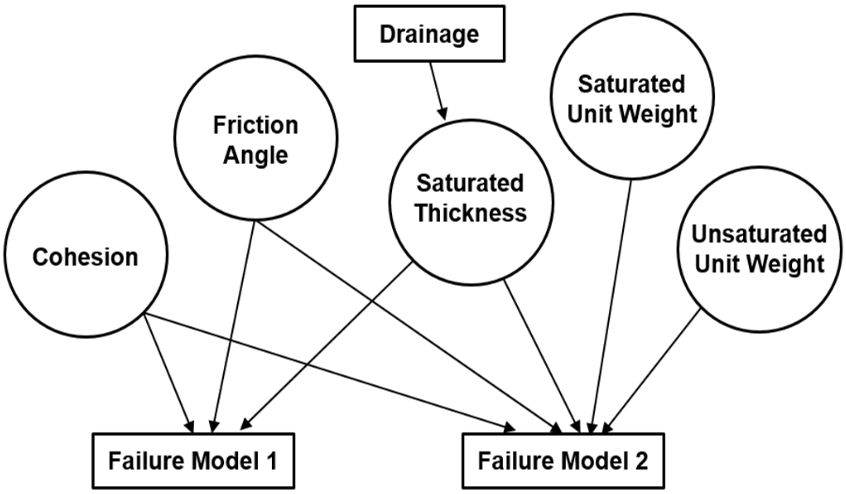



Fig. 5 The hybrid BN model of an infinite slope

*Saturated Unit Weight* and *Saturated Thickness* are selected in the slope model according to effective stress principle (Terzaghi, 1943), 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 (Morgenstern and Price, 1965; Niu, 2014) that can be selected to obtain the

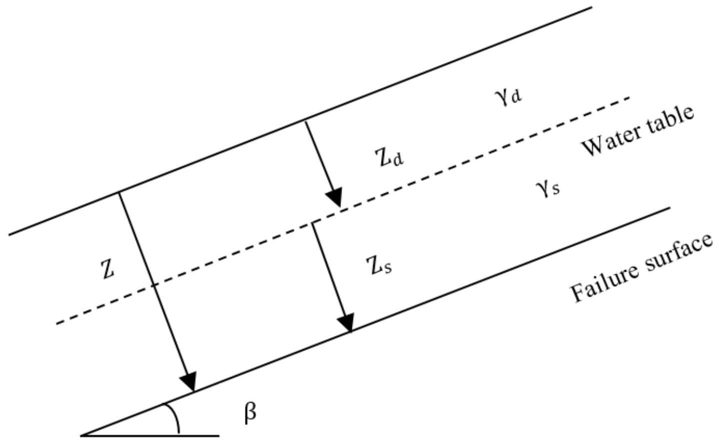



Fig. 6 The slope with translational slip

factor of safety according to different conditions. These conditions also depend on the type of failure surface and its extension.

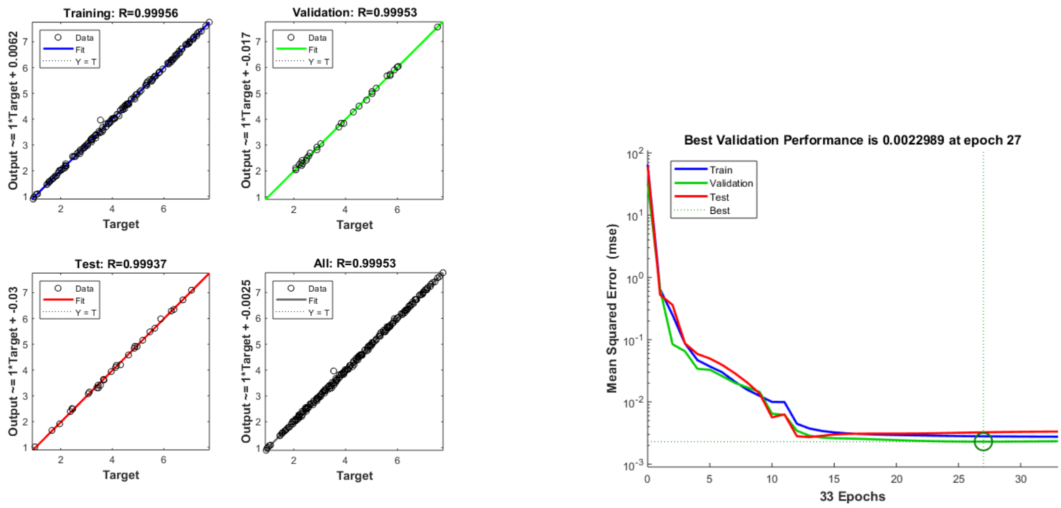
In the analysis of a given failure surface, as Fig. 6 shows, the equation for the factor of safety in terms of effective stress analysis is given by

$$FOS = \frac{c + (\gamma_d Z_d + \gamma_s Z_s - \gamma_w Z_s) \cdot \cos\beta \cdot \tan\phi}{(\gamma_d Z_d + \gamma_s Z_s) \cdot \sin\beta} \tag{8}$$

here, the drained parameters of cohesion ($c$ ) and friction angle ($\phi$) 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 ($Z$). $\beta$ is the slope inclination and $\gamma_w$ is the unit weight of water, 9.81 $kN/m^3$. For the layer above and below water table, soil unit weight should be split into two parts: dry unit soil weight ($\gamma_d$) and saturated unit soil weight ($\gamma_s$). This analysis has been completed using the equilibrium of an infinite (Acharya et al., 2006). Moreover, $FOS \leq 1$ 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 $\phi$, 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 $c$, ($\phi$) and $Z_s$ is 5, 5, 8, respectively and the results of FOS are carried out by the experimental runs on $c$, ($\phi$) 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 ($c$, ($\phi$) 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 (Pinheiro Branco et al., 2014) is adopted to describe the correlation between $\gamma_d$ and $\gamma_s$. The probabilities of two failure models are computed by the limit state function $G(X)$,

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

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:

$$P(FM) = \begin{cases} P_f, & G(X) - 1 < 0 \\ P_s, & G(X) - 1 \geq 0 \end{cases} \tag{10}$$

herein $P_f$ denotes the failure probability of the slope while the safe probability is $P_s$. To characterize the relationship between slope stability and its influence factors, one

Table 1 Input parameters of the soil slope.

| Parameters | Variable type | CPD* |
|---|---|---|
| Cohesion ($kPa$) | Continuous | logN(22, 10) |
| Friction angle (°) | Continuous | N(35, 3) |
| Unsaturated unit weight ($kN/m^3$) | Continuous | N(17, 0.4) |
| Saturated unit weight ($kN/m^3$) | Continuous | N(19, 0.5) |
| Saturated thickness of soil ($m$) | Continuous | U(0, 4) or 0 |
| Drainage ($D$) | Discrete | [0.5, 0.5] |
| Failure model ($FM$) | Discrete | [$P_f$,$P_s$] |
| Young's modulus($MPa$) | Constant | 50 |
| Poison's ratio | Constant | 0.3 |

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

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 $c$, $\phi$, $Z_s$, $\gamma_d$ and $\gamma_s$ 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.

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

Table 2 The effect of Drainage on slope safety.

| State | $P(FM)$ | $P(FM \mid D = false)$ | $P(FM \mid D = true)$ |
|---|---|---|---|
| $FM1$ | 7.77e-02 | 8.01e-02 | <1.00e-08 |
| $FM2$ | 7.21e-02 | 7.50e-02 | 6.00e-04 |

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 Slope failure probability updated with new information.

| Parameter<br>Evidence | $c$<br>$[0, 100]$ | $\phi$<br>$[20, 50]$ | $Z_s$<br>$[0, 4]$ | $\gamma_d$<br>$[16, 19]$ | $\gamma_s$<br>$[18, 21]$ |
|---|---|---|---|---|---|
| $P(FM1)$ | 7.93e-02 | 7.94e-02 | 7.91e-02 | - | - |
| $P(FM2)$ | 7.37e-02 | 7.27e-02 | 7.28e-02 | 7.38e-02 | 7.25e-02 |

Table 4 Slope failure probability updated with further information

| Parameter<br>Evidence | $c$<br>$[25, 75]$ | $\phi$<br>$[30, 40]$ | $Z_s$<br>$[1, 3]$ | $\gamma_d$<br>$[17, 18]$ | $\gamma_s$<br>$[19, 20]$ |
|---|---|---|---|---|---|
| $P(FM1)$ | 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 |

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 $\phi$ 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 $\phi$, $\gamma_d$ and $\gamma_s$. 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.

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.

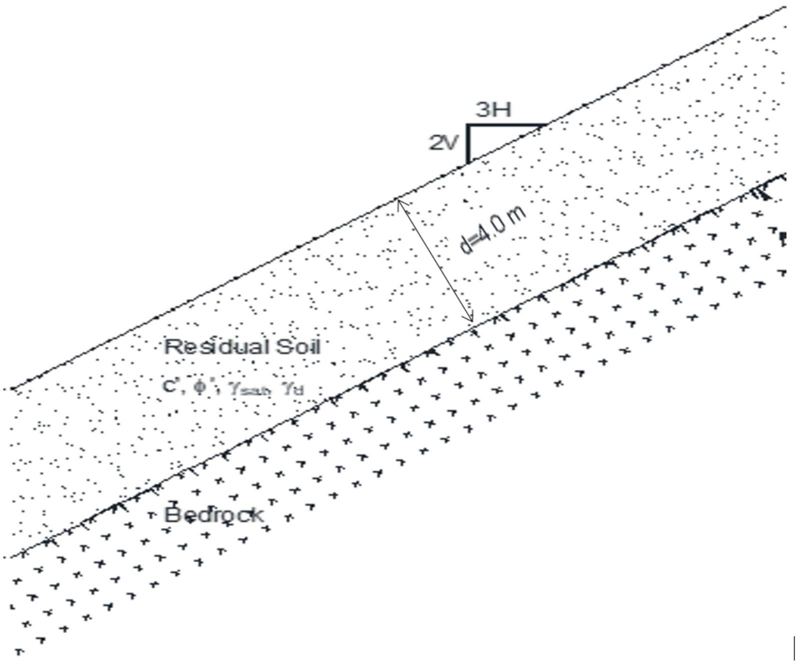



Fig. 8 A residual soil slope

# 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 (Viana da Fonseca and Coutinho, 2008; Topa Gomes, 2009; Pinheiro Branco et al., 2014). 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.

Three different situations of information available in $c$, $\phi$, $\gamma_d$ and $\gamma_s$, 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.

Table 5 Input parameters of the residual soil

| Nodes | Scenario 1 | Scenario 2 min | Scenario 2 max | Scenario 3* |
|---|---|---|---|---|
| $c$ | logN(20, 4) | 0 | 70 | logN($\mu_c$, 4), $\mu_c \in [16, 22]$ |
| $\phi$ | N(37,1.85) | 25 | 47 | N($\mu_f$, 1.85), $\mu_f \in [36, 38.5]$ |
| $\gamma_d$ | N(18.5, 0.51) | 17 | 20 | [17, 20] |
| $\gamma_s$ | N(20, 0.6) | 18 | 22 | [18, 22] |

*$\mu$ indicates the mean of the distributions. The notes of Table **??** also apply here.

## 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 $c$,$\phi$, $\gamma_d$ and $\gamma_s$ 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.

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

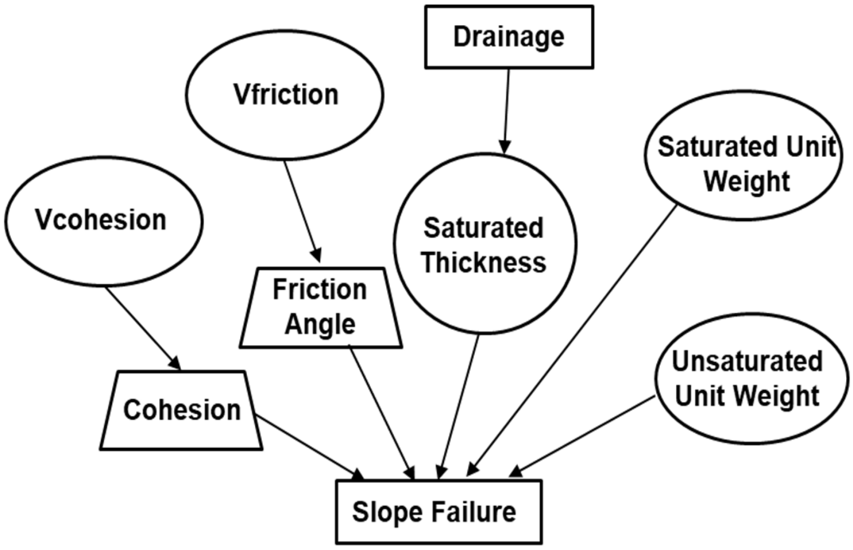



Fig. 9 The CN model of an infinite slope

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 (Patelli, 2016; Patelli et al., 2017). 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 (De Angelis et al., 2014) 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. 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

Table 6 Slope failure probability

| Different information | Scenario 1 | Scenario 2 | Scenario 3 |
|---|---|---|---|
| $P(FS)$ | 0.0274 | [0, 1] | [0, 0.0711] |

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.

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

Table 7 Failure probabilities with two states of Drainage

| Different information | Scenario 1 | Scenario 2 | Scenario 3 |
|---|---|---|---|
| $P(FS \mid D = true)$ | 0 | [0, 1] | 0 |
| $P(FS \mid D = false)$ | 0.0514 | [0, 1] | [0, 0.0153] |

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

Acharya, G., F. De Smedt, and N. T. Long. 2006. "Assessing Landslide Hazard in GIS: A CaseStudy from Rasuwa, Nepal."Bulletin of Engineering Geology and the Environment65 (1):99–107.

Antonucci, A., C.P. de Campos, D. Huber, and M. Zaffalon. 2015. "Approximate Credal Net-work Updating by Linear Programming with Applications to Decision Making."Interna-tional Journal of Approximate Reasoning58 (IS): 25–38.

Branco, L. Pinheiro, A. Topa Gomes, A. Silva Cardoso, and C. Santos Pereira. 2014. "NaturalVariability of Shear Strength in a Granite Residual Soil from Porto."Geotechnical andGeological Engineering32 (4): 911–922.

Chakraborty, A., and D. Goswami. 2017. "Slope Stability Prediction of Slope Stability Us-ing Multiple Linear Regression (MLR) and Artificial Neural Network."Arabian Journal ofGeosciences10 (17): 385–395.

Chen, Y. C., T. A. Wheeler, and M. J. Kochenderfer. 2017. "Learning Discrete BayesianNet-works from Continuous Data."Journal of Artificial Intelligence Research59: 103–132.

Cozman, F. G. 2000. "Credal Networks."Artificial Intelligence120: 199–233.da Fonseca, A. Viana, and R. Coutinho. 2008. "Characterization of Residual Soils." InPro-ceedings of the 3rd International Conference on Site Characterization,, 195–248.

Daly, R., Q. Shen, and Q. S. Aitken. 2011. "Learning Bayesian Networks: Approaches andIssues."The Knowledge Engineering Review26 (2): 99–157.

De Angelis, M., E. Patelli, and M. Beer. 2014. "Advanced Line Sampling for Efficient RobustReliability Analysis."Structural Safety52 (SI): 170–182.

Dougherty, J., R. Kohavi, and M. Sahami. 1995. "Supervised and Unsupervised Dis-cretizationof Continuous Features."In: Proceedings of the 12th International Conference on MachineLearning (ICML).

El-Ramly, H., N. R. Morgenstern, and D. M. Cruden. 2002. "Probabilistic Slope Stability-Analysis for Practice."Canadian Geotechnical Journal39 (3): 665–683.

Fagiuoli, E., and M. Zaffalon. 1988. "2U: An Exact Interval Propagation Algorithm for Poly-trees with Binary Variables."Artificial Intelligence106 (1): 77–107.

Ferreira da Rocha, J. C., and F. G. Cozman. 2002. "Inference with Separately Specified Setsof Probabilities in Credal Networks." InProceedings of the 18th Conference on Uncertaintyin Artificial Intelligence, San Francisco, 430–437.

Gomes, A. Topa. 2009. "Elliptical Shafts by the Sequential Excavation Method. The case ofMetro do Porto, (in Portuguese)."PhD Thesis, University of Porto195–248.

Koutsourelakis, P. S. 2004. "Reliability of Structures in High Dimensions.Part ii. Theo-reticalValidation."Probability Engineering Mechnique19 (4): 19–23.

Kristo, C., H. Rahardjo, and A. Satyanaga. 2017. "Effect of Variations in Rainfall Intensity onSlope Stability in Singapore."International Soil and Water Conservation Research4 (1):258–264.

Kurgan, L., and K.J. Cios. 2001. "Discretization Algorithm that Uses Class-attribute Inter-dependence Maximization." InProceedings of the International Conference on ArtificialIntelligence (IC-AI), Las Vegas, Nevada, 980–987.

Langseth, H., T. D. Nielsen, R. Rumi, and A. Salmern. 2009. "Inference in Hybrid BayesianNetworks."Reliability Engineering System Safety94: 1499–1509.

Liu, W. S., S. B. Li, and R. Tang. 2013. "Slope Stability Evaluation in Open Pit Based onBayesian Networks."Proceedings of the ICCEAE2012181: 1227–1231.Ma, J., H.

Liu, Y., W. Zhang, L. Zhang, Z. Zhu, J. Hu, and H. Wei. 2018. "Probabilistic stability analysesof undrained slopes by 3D random fields and finite element methods."Geoscience Frontiers9 (6): 16571664.

Ma, J., H. Tang, X. Hu, A. Bobet, M. Zhang, T. Zhu, Y. Song, and A. M. Mutasim. 2017."Identification of Causal Factors for the Majiagou Landslide Using Modern Data MiningMethods."Landslides14 (1): 311–322.

Metya, S., T. Mukhopadhyay, S. Adhikari, and G. Bhattacharya. 2009. "Efficient Evaluation ofReliability for Slopes with Circular Slip Surfaces Using Importance Sampling."Geotechnicaland Geoenvironmental Engineering135 (6): 768–777.

Metya, S., T. Mukhopadhyay, S. Adhikari, and G. Bhattacharya. 2017. "System ReliabilityAnalysis of Soil Slopes with General Slip Surfaces Using Multivariate Adaptive RegressionSplines."Computers and Geotechnics87 (7): 212–228.

Morgenstern, N. R., and V. Price. 1965. "The Analysis of the Stability of General Slip Surface."Geotechnique15 (1): 79–93.

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

Oka, Y., and T. H. Wu. 1990. "System Reliability of Slope Stability."Geotechnical Engineering116 (8): 1185–1189.

Patelli, E. 2016. "A Multidisciplinary Software Suite for Uncertainty Quantification and RiskManagement."PIn: Ghanem R., Higdon D., Owhadi H. (eds.) Handbook of UncertaintyQuantification, Springer, Cham, Switzerland.

Patelli, E., M. Broggi, S. Tolo, and J. Sadeghi. 2017. "COSSAN software: A multidisciplinaryand collaborative software for uncertainty quantification."Proceedings of 2nd ECCOMASThematic Conference on Uncertainty Quantification in Computational Sciences and Engineering, Rhodes, Greece212224.

Pearl, J. 1988.Probabilistic Reasoning in Intelligent Systems. San Mateo, California: MorganKaufmann Publishers.

Pearl, J., and S. Russell. 2000.Causality: Models, Reasoning, and Inference. Cambridge Uni-versity Press.Peng, M., X. Y. Li, and D. Q. Li. 2013. "Slope Safety Evaluation by Integrating Multi-sourceMonitoring Information."Structural Safety49 (SI): 65–74.

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

Rocscience. 2011. "RS2 Software."In reocscience, version 7.0 [computer program].

Sakellariou, M. G., and M. D. Ferentinou. 2005. "A Study of Slope Stability Prediction UsingNeural Networks."Geotechnical and Geological Engineering23: 419–445.

Song, Y. Q., J. H. Gong, and S. Gao. 2012. "Susceptibility Assessment of Earthquake-inducedLandslides Using Bayesian Network: A Case Study in Beichuan, China."Computers andGeosciences42: 189–199.

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

Straub, D., and A. Der Kiureghian. 2010. "Combining Bayesian networks with StructuralReliability Methods: Methodology."Journal of Engineering Mechanics136 (10): 1248–1258.

Straub, D., and A. Der Kiureghian. 2010. "Bayesian Network Enhanced with Structural Reli-ability Methods: Methodology."Journal of Engineering Mechanics136 (10): 12481258.

Terzaghi, K. 1943.Theoretical Soil Mechanics. New York: John Wiley  Sons.Tessem, B. 1992. "Interval Probability Propagation."International Journal of ApproximateReasoning7: 95–120.

Tolo, S., E. Patelli, and M. Beer. 2018a. "An Inference Method for Bayesian Networks withProbability Intervals." InProceedings of the joint ICVRAM-ISUMA UNCERTAINTIES Conference, Florianpolis, SC, Brazil.

Tolo, S., E. Patelli, and M. Beer. 2018b. "An Open Toolbox for the Reduction, Inference Computation and Sensitivity Analysis of Credal Networks."Advances in Engineering Software115: 126–148.

Tsai, F., J.S. Lai, and T.H. Lin. 2013. "Analysis of Topographic and Vegetative Factors withData Mining for Landslide Verification."Ecological Engineering61: 669–677.

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

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

# 4 Research article II: Estimation of failure probability in braced excavation using Bayesian networks with integrated model updating

## 4.1 Declaration of my contribution

Most of the related computations and analyses included in this article were implemented by myself. The manuscript benefited from the discussion with the co-authors. The related data and finite element model of the braced excavation were provided by associate Prof. Wang Li and Prof. Yong Li, and the research field of them are both from geotechnical engineering. Besides, the concept of the model updating method based on distance metrics was provided by Dr. Bi, which was described in the articled. Based on the provided information, I enhanced the capability of BNs and its application for solving the excavation problem. Finally, the article was improved by comments and suggestions from the reviewers during the review process.

## 4.2 Published article

This paper has been accepted for publication in Underground Space.

Submitted: 4 April 2019. Accepted: 4 July 2019. Available online 16 September 2019 In Press.

**Longxue He, Yong Liu, Sifeng Bi, Li Wang, Matteo Broggi, Michael Beer**. Estimation of failure probability in a braced excavation by Bayesian networks integrating with model updating approaches.

# Estimation of failure probability in braced excavation using Bayesian networks with integrated model updating

Longxue He[1], Yong Liu [2], Sifeng Bi[1], Li Wang[3], Matteo Broggi [1], Michael Beer [1,4,5*],

[1]Institute for Risk and Reliability, Leibniz University Hannover, Germany
[2]State Key Laboratory of Water Resources and Hydropower Engineering Science, Institute of Engineering Risk and Disaster Prevention, Wuhan University, China
[3]School of Civil and Safety Engineering, Dalian Jiaotong University, China
[4]Institute for Risk and Uncertainty, University of Liverpool, UK
[5]International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, China
*Corresponding author, e-mail: beer@irz.uni-hannover.de
Accepted: 4 July 2019

## Abstract

A probabilistic model is proposed that uses observation data to estimate failure probabilities during excavations. The model integrates a Bayesian network and distanced-based Bayesian model updating. In the network, the movement of a retaining wall is selected as the indicator of failure, and the observed ground surface settlement is used to update the soil parameters. The responses of wall deflection and ground surface settlement are accurately predicted using finite element analysis. An artificial neural network is employed to construct the response surface relationship using the aforementioned input factors. The proposed model effectively estimates the uncertainty of influential factors. A case study of a braced excavation is presented to demonstrate the feasibility of the proposed approach. The update results facilitate accurate estimates according to the target value, from which the corresponding probabilities of failure are obtained. The proposed model enables failure probabilities to be determined with real-time result updating.

## keyword

Failure probability; Braced excavation; Bayesian networks; Stochastic model updating; Sensitivity analysis

# 1. Introduction

As urban construction activities increase, so does foundation pit excavation, as this is the first step of most construction projects. However, this activity often has unfavorable consequences in urban areas. During the excavation process, deformation of the diaphragm wall and ground surface elevation can occur, which can cause collapse of the adjacent building structures and sometimes results in human casualties. Many numerical models have been proposed to compute maximum wall displacement and maximum ground surface settlement. Do et al. (2016) studied the failure mechanism of excavation in soft clay using finite element (FE) analysis. They showed that the FE method could effectively estimate excavation stability. Kung et al. (2007) proposed a simplified semi-empirical model, named the KJHH model, to estimate the deformation behavior of a braced excavation. The KJHH model incorporated three models that assessed wall deflection, ground surface settlement, and deformation rate. These studies provide empirical or semi-empirical methods to predict the response values of deformation and failure threshold using various input parameters.

Furthermore, FE analysis is a popular approach for addressing problems associated with sophisticated excavations, wherein two-dimensional (2D) plane strain problems are utilized to predict the stability of excavation with the purpose of simplification. However, three-dimensional (3D) effects are more suitable and more accurate for the analysis of certain geotechnical problems (such as tunnel excavations) in practical situations. Several 3D FE analyses have been studied (OU et al., 1996; Zdravkovic et al., 2005; Lee et al., 2011). Additionally, Janin et al. (2015) compared the relative ability of both 2D and 3D approaches. They found that the 3D approach enabled representation of the reinforcements, ground reaction, and the 3D phenomenon of tunnel excavations, while the 2D simulation failed to represent these complex effects fully. Therefore, given the case studied in this paper, the 3D FE method was applied using the software package ABAQUS 6.13. This software package was selected because it has been proven effective for 3D FE analyses (Lee et al., 2011; Liu et al., 2015; Li and Gang, 2018).

Owing to the unavoidable model errors associated with insufficient knowledge of the reality of a situation and its complex excavation conditions, discrepancies between design parameters and observation parameters are quantified using an updating method. In situations in which field measurements are provided, it is common practice to update geotechnical parameters with back analysis or inverse analysis based on modeling functions (Ledesma and Alonso, 1996; Finno and Calvello, 2005; Calvello and Finno, 2004; Hashash et al., 2010; Zhang et al., 2011). For trial-and-error calibration methods, inverse model algorithms are initially considered. For supported excavations, Finno and Calvello (2005) handled design prediction updates using UCODE (a computer code) to assign identification numbers to physical objects. They were able to minimize model errors using their proposed numerical procedure. Juang et al. (2013), attempted to present the maximum likelihood method for updating soil parameters in a stage-by-stage manner. The likelihood function was obtained according to the bias factor of the KJHH model, wherein the bias factor was often assumed to have a normal or lognormal distribution based on expert knowledge. It is considered impractical to implement updating for wall and ground surface deformation prediction using back analysis by way of implicit probability distribution, mass functions of the model, or black-box methods.

Approximate Bayesian computation (ABC) (Beaumont et al., 2002; Turner and Zandt, 2012) can reveal discrepancies by checking the distance, rather than likelihood function. Bi et al. (2018) developed an ABC model updating framework that considered both Euclidian and Bhattacharyya in quantifying uncertainty. The study revealed an efficient and capable metric for stochastic model updating. In geotechnical updating problems, when the likelihood function is intractable or cannot be approached in a closed form (as a likelihood-free method), an ABC approach is typically used. Nonetheless, there remains scant research on updating geotechnical materials using distance-based ABC approaches. Accordingly, the method presented in this paper incorporates distance metrics for material parameter updating in the supported excavation.

Regardless of whether it is from the perspective of design considerations or risk management, prediction of excavation stability enables the implementation of crucial pre-failure controls. The risks and causes of potential failures in the excavation process are complicated, various, and interactional. Numerous dynamic factors such as soil structure and strength, excavation width, and workmanship affect surface settlement and the movement of braced walls. As a diagnostic tool, Bayesian networks possess the powerful capability of being able to analyze multiple causal failures. The flexibility of network structures contributes to their application in many fields for risk analysis, risk management, and decision analysis.

Zhang et al. (2013) presented a decision support Bayesian network (BN) model to predict ground settlement for safety control. Influential factors in this network were all defined by discrete nodes with three states. A dynamic BN model (Spackova and Straub, 2013) was utilized to assess the risk of human factors and other external events in the tunnel construction process. Zhou et al. (2018) used a BN for the analysis of risk classification for diaphragm wall deflection based on field data. Their model combined the field data of the diaphragm wall with other data as evidence input to validate the predicted results. In these application of BNs, the prior probabilities of each node were highly dependent on collected data and expert opinion.

The primary objective of this study is to present a real-time probabilistic model that will use updated information to predict the possibility of collapse during the excavation process. To accomplish this, we focus on capturing the uncertainty of material parameters, and on characterizing their effect on a diaphragm wall using a BN model. To overcome the difficulty of monitoring wall deformation in tandem with ground surface settlement, ground settlement is accounted for as a field observation. Hence, it is incorporated into the BN model as an input used to update the material parameters. The proposed method combines Bayesian networks with a model updating approach. Not only does the proposed model incorporate information based on expert judgment and limited data from direct and indirect factors, it also captures the propagation of uncertainty throughout the network components. Thus, the proposed model characterizes the relationship between uncertain parameters and the safety states of the excavation process, and identifies the influence of induced-factors on the stability of the excavation.



# 2. The Bayesian network of excavation evaluation by integrating field observation

## 2.1 Structure of Bayesian networks

A BN is a graphical statistical model involving the that realizes powerful probability theory. In this directed acyclic graph, a visible cause-effect relationship can be shown with a set of variables being linked by an arc that, showing their conditional dependency, while the indirect edges indicate the independent conditional relations among the nodes. A detailed overview of BNs can be found in (Pearl, 1988; Jensen, 2001). Bayesian updating for braced excavation is generally usually conducted in stages. By The useing of a BN model, it is able to addresses the complexities handle the difficulties of stepped excavations for multiple layers while providing and estimates of the real-time failure state of the excavation.

Additionally, BNs are well suited to capturing uncertainty propagation. Because of this, a general network for braced excavations by steps is constructed by considering the material parameters $X_i$ (as root nodes in the model) and the parent nodes of deformation parameters Di of the different layers. The discrete nodes $Y_i$ present the states of safety or failure for diverse materials of each layer, while $Y_{total}$ represents the final failure events. Additionally, monitor parameters $M_i : (i = 1, 2, \ldots, n)$ in the excavation process are integrated into the network. Per Fig. 1, a BN is built considering two material layers and one monitor parameter.

According to the chain rule, the factorization of partial variables in this network is written by

$$f(D_2, X_2, X_3, X_4, M_1) = f(X_2)f(X_3)f(X_4)(D_2|X_2, X_3, X_4)f(M_1|X_3, X_4) \quad (1)$$

where $f(D_2, X_2, X_3, X_4, M_1)$ is a joint probability distribution, and the probability of node $D_2$ can be obtained with the marginal computation of Eq. (1). The probability of

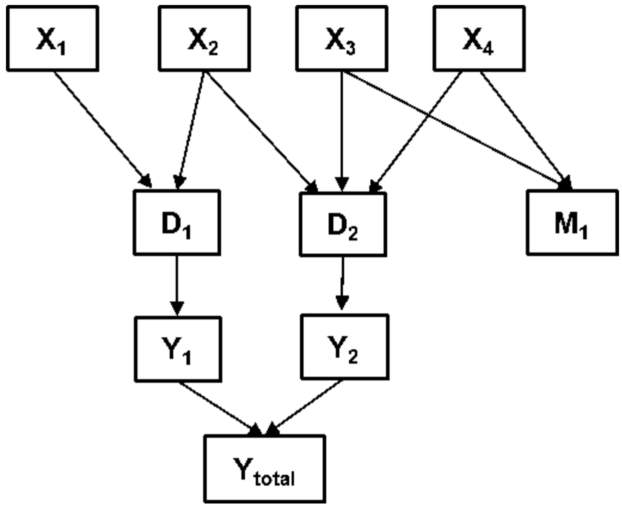



Fig. 1: A BN involving the field measurement.

excavation failure $Y_2$ is given by

$$Pr(Y_2|D_2) = \int_{\Omega_{D_2}} f(D_2|X_2, X_3, X_4)d(D_2) \tag{2}$$

where the domain $\Omega_{D_2}$ of variable $Y_2$ is divided by safe and failure domain. Likewise, the conditional probability distribution (CPD) of node $D_1$ and the conditional probability of node $Y_1$ also can be computed. Furthermore, the events of $Y_1$ and $Y_2$ are independent and both exist in a binary state. The event $Y_{total}$ is the joint event of $Y_1$ and $Y_2$.

## 2.2 Quantitative component of a Bayesian network

In a BN, each node should be defined by the corresponding prior probabilities. In this work, we employ a neural network to quantify the inter-relationship among the main parameters. Artificial neural networks (ANN) are an efficient tool to simulate the response of output associated with input variables (Anjum et al., 1997; Yuan and Bai, 2011; Mia and Dhar, 2016). Hashash et al. (2004) demonstrated complex stress-strain behavior of engineering materials could be effectively captured using an ANN.

A neural network is designed using three layers: input, hidden, and output. Two response relationships comprise the required input to provide a single output. The displacement of the wall is taken as an indicator parameter for detecting failure in the excavation process. As previously mentioned, the response of ground surface settlement is selected as observation data. Furthermore, the response of wall deflection has four material variable inputs, while two variable inputs represent ground surface settlement. For the training and test data, we adopted the simulated values from a dataset calculated using the FE method in the software package ABAQUS 6.13. The amount of training was defined by full factorial designs. The ANN computation was run in the MATLAB R2017 'nnstart' toolbox, and the network was trained using Bayesian regularization (MacKay, 1992). The performance of the training was evaluated by mean square error (MSE) as given in Eq. (3).

$$MSE = \frac{1}{2}\sum_{1}^{n}(Actual - predicted)^2 \tag{3}$$

Note that the response surface built via an ANN is a black-box. In the next step, we integrate this black-box with the BN model. From this point forward in the process, information updating and sensitivity analysis are executed based on this integrated model.

## 2.3 Bayesian updating

Soil parameters will vary as the excavation is conducted, which makes direct measurement intractable in current practice. Therefore, soil parameters are generally updated with monitor parameters. This article considers ABC as a likelihood-free method. The recently developed ABC updating framework utilizing various statistical distances (Bi et al., 2018) is employed to update the soil parameters in the proposed approach. Given the problem addressed by this paper, this model updating approach is applied to update the key soil parameters: a Bayesian updating framework is presented with the distance-based ABC approach, including Euclidian and Bhattacharyya distances. The entire update process will be briefly presented in the next section.

Let $\theta$ be the uncertain parameters and $D$ be the observed data. Then from Bayes' theorem, the computation of the posterior probability density function (PDF): $f(\theta|D)$ improves the accuracy of the predictions of the model given the data $D$; and is expressed as

$$f(\theta|D) = \frac{f(D|\theta)f(\theta)}{\int f(D|\theta)f(\theta)d(\theta)} \tag{4}$$

where $f(\theta)$ is the priori distribution and $f(D|\theta)$ is equal to the likelihood distribution. $\int f(D|\theta)f(\theta)d(\theta)$ represents normalization and should be constant, but can be intractable to compute because of the high-dimension of the parameter space or multimodal distribution. In this regard, the transitional Markov Chain Monte Carlo (TMCMC) method (Ching and Chen, 2007) is proposed to overcome the difficulty in evaluating the target PDF. Briefly, the TMCMC method is an effective simulation to support sampling from a set of the intermediate PDFs and converge to the target PDF. Generally, the sampling from the posterior distribution with the TMCMC method is estimated based on the following Eq. (5).

$$f(\theta|D) \propto f(D|\theta) \times f(\theta) \tag{5}$$

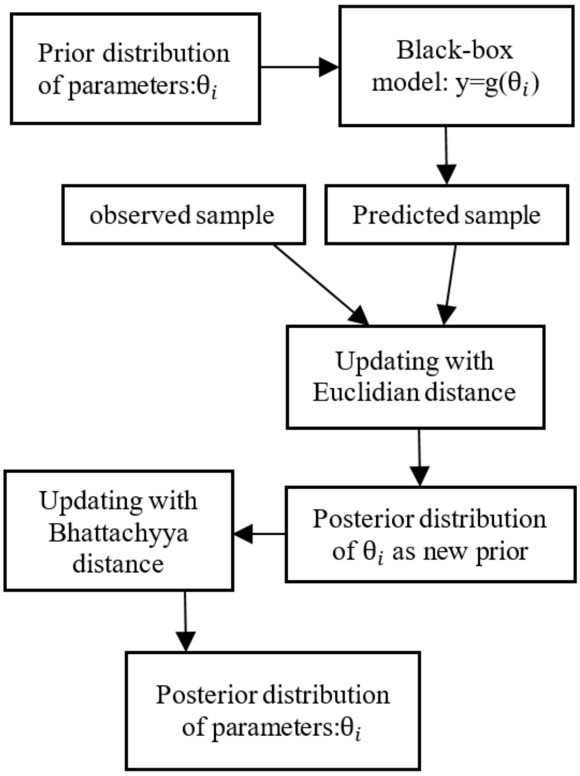



Fig. 2: Updating framework with distance-based ABC method.

Accordingly, these intermediate PDFs are constructed:

$$f_j \propto f(D|\theta)^{P_j} \times f(\theta) \tag{6}$$

where $j$ is the stage number and $P_j$ denotes the exponent of the likelihood, ranging from 0 to 1 (Ching and Chen,2007). In the updating framework, $f(D|\theta)$ is estimated by approximate distance-based likelihood based on the Gaussian function,

$$f(D|\theta) \propto e^{(-\frac{d^2}{\varepsilon^2})} \tag{7}$$

where $d$ is the distance metric, which can be either the Euclidian distance or the Bhattacharyya distance. $\varepsilon$ is the width factor in the range between 0.001 and 0.1. The smaller the value of $\varepsilon$ is, the more likely that the result converges to the true value, but the increasing likelihood brings a corresponding requirement for more calculation (Bi et al., 2018). The formula for calculating the Euclidian distance between two n-dimension vectors **x** (predicted data) and **y** (observed data) is delineated in Eq. (8).

$$d = \sum_{i=1}^{n}(x_i - y_i)^2 \tag{8}$$

The Bhattachyya distance is defined in Eq. (9).

$$d = -log[\int_n p_{pre}(x)p_{obs}(x)dx] \tag{9}$$

In the Bhattacharyya distance, $p_{pre}(x)$ and $p_{obs}(x)$ are the PDF of the prediction and observation samples, respectively. This stochastic distance metric is especially suitable for measuring the overlap of the sample set. However, without an overlap situation, it is insensitive to the center of mass of the sets. Therefore, as described in the framework depicted in Fig. 2, we first conduct the updating with Euclidian distance to measure the likelihood and then use the results as new prior distribution of $\theta$ and execute the next updating with the Bhattacharyya distance-based ABC.

## 2.4 Moment-independent sensitivity analysis using Monte Carlo simulation

The sensitivity analysis of a BN contributes to the process of building the BN from data, and is especially useful for large and complex networks. Several types of methods have been studied to determine how causal nodes influence the target node in traditional BNs (Laskey, 1995; Chan and Darwiche, 2004). In these approaches, the evidence is inserted by querying the different states of each variable. The characteristic of the sensitivity is then estimated based on the changes of the posterior probabilities for the target node. However, this is not suitable for complex networks. Given the problem addressed in this paper, a sensitivity approach for identifying the critical inputs before network updating is proposed.

The variance-based sensitivity analysis method is a summary measure of sensitivity that studies how the variance of the output changes when an input variable is fixed. Li and Mahadevan (2017) applied the first-order Sobol' index to BNs to analyze the sensitivity and proposed an approximated algorithm to reduce the computational cost. For problems of risk analysis, the robust sensitivity measure for a BN should enable to capture the entire distribution of an output node, rather than only a single moment.

Moment-independent sensitivity analysis enables capture of the entire distribution of output referring to varying input parameters. Therefore, in this paper, we employ the PDF-based sensitivity approach to measure the contribution of the uncertain parameters in the deformation of a retaining wall. Using this method, we attempt to select the key factors to update. Generally, the moment-independent sensitivity index $\delta_i$ is evaluated to identify the effect of any of an input $X_i$ on the PDF of model output $Y$. According to the definition of the delta index (Borgonovo, 2007), the formulation of the computation is written by:

$$\begin{aligned}\delta_i &= \frac{1}{2}E_{X_i}[s(X_i)] \\ &= \frac{1}{2}\int s(X_i) f_{X_i}(x_i) dx_i\end{aligned} \tag{10}$$

where $s(X_i)$ denotes the area difference between the unconditional PDF of output $Y$ and its conditional distribution given the individual input $X_i$,

$$s(X_i) = \int |f_Y(y) - f_{Y|X_i}(y)| dy \tag{11}$$

where $f(\cdot)$ denotes the PDF. Note that $\delta_i \in [0, 1]$, where 0 means input $X_i$ has no effect on the PDF of Y and the contribution of all the inputs are the same to the PDF of $Y$ when $\delta_i = 1$. Additionally, the computation of the delta index with the single-loop Monte Carlo simulation (Wei et al., 2013) is employed. Based on Eq. (11), the form of $\delta_i$ can be further considered by,

$$\begin{aligned}\delta_i &= \frac{1}{2}\int\int |f_{X_i}(x_i) f_Y(y) - f_{Y,X_i}(y, x_i)| dy\, dx_i \\ &= \frac{1}{2}E_{Y,X_i}\left(\left|\frac{f_{X_i}(x_i) f_Y(y)}{f_{Y,X_i}(y, x_i)} - 1\right|\right)\end{aligned} \tag{12}$$

So, for a group of input parameters $R = (X_{i1}, X_{i2}, \ldots, X_{in})$, similarly, we can obtain

$$\delta_i = \frac{1}{2}E_{Y,R}\left(\left|\frac{f_R(x_{i1}, x_{i2}, ..., x_{in}) f_Y(y)}{f_{Y,R}(y, x_{i1}, x_{i2}, ..., x_{in})} - 1\right|\right) \tag{13}$$

In this paper, the PDF of output $f_Y(y)$ is estimated with the kernel density estimator (KDE) method (Botev et al., 2010), while a bivariate KDE toolbox (Botev,2010) is used to achieve the joint PDF of $Y$ and $R$. Thus, the sensitivity index $\delta_i$ is easily calculated by means of Monte Carlo simulation.

## 3. Example application

The braced excavation for the tunnel was conducted in the Marina Bay area of Singapore. The pit excavation, depicted in Fig. 3(a), consisted of three layers of material: sand fill, marine clay, and Old Alluvium. The marine clay was modelled using the Cam Clay model, while the sand fill and Old Alluvium were modelled as Mohr-Coulomb materials with effective stress parameters. The model in Fig. 3(b) is 24 m width along

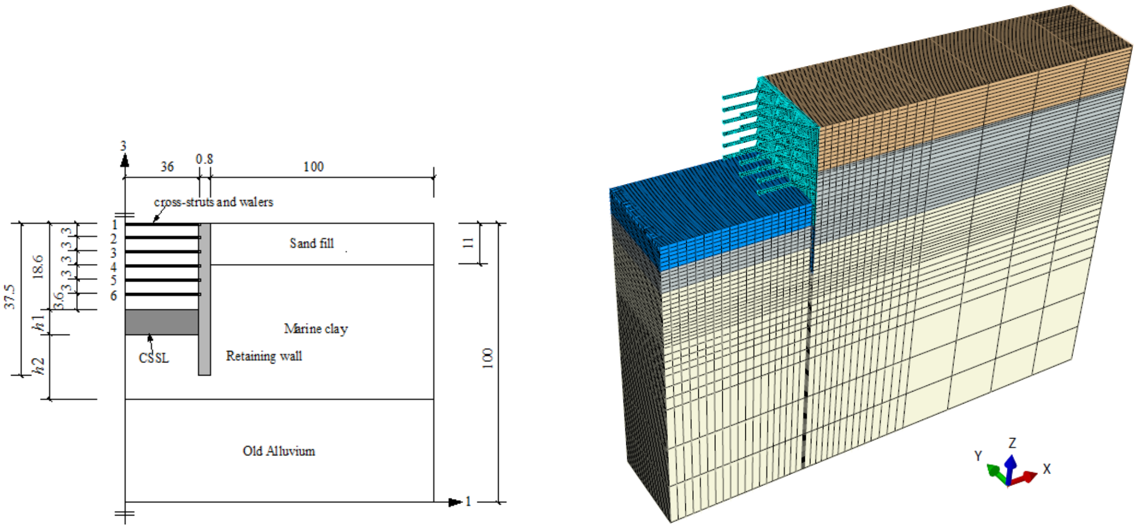
Table 1: Properties of Soils and CSSL (Lee et al., 2011).

| Soil type | Marine Clay | Sand Fill | Old Alluvium | CSSL |
|---|---|---|---|---|
| Bulk unit weight ($kN/m^3$) | 16 | 19 | 20 | 16 |
| Isotropic swelling index | 0.093 | - | - | - |
| Isotropic compression index | 0.27 | - | - | - |
| Critical state friction coefficient | 0.87 | - | - | - |
| Effective poisson's ratio | 0.3 | 0.3 | 0.3 | 0.2 |
| Earth pressure coefficient | 0.7 | 0.5 | 1.0 | 0.7 |
| Coefficient of permeability | $1 \times 10^{-9}$ | $1 \times 10^{-6}$ | $1 \times 10^{-7}$ | $1 \times 10^{-10}$ |
| Void ratio | 1.9 | - | - | - |
| Friction angle ($^\circ$) | - | 30 | 37 | 41 |
| Effective Young's modulus | - | 10 | 130 | 272 |
| Angle of dilation ($^\circ$) | - | 0 | 10 | 0 |
| Effective cohesion | - | 2 | 20 | 400 |

the $Y$ direction. At the two surfaces (i.e. $Y = 0$ and $Y = 24$), vertical rollers are set, and only vertical movement is allowed.

In this study, the cement-treated soil was modelled as a Mohr-Coulomb material. The properties of the soil layers and cement stabilized soil layer (CSSL) used in this analysis are listed in Table 1. The height and width of the pit excavation were 100 m. The excavation had a total of six stages using the top-down construction method The final excavation depth was 18.6 m. A retaining wall with a thickness of 0.8 m was supported by cross-struts and walers. The maximum movement of the retaining wall was 162.2 mm. The diaphragm wall was modelled as an elastic material with an equivalent Young's modulus of 10.5 GPa and Poisson's ratio of 0.2. The cross-struts and walers supporting the retaining wall were simplified to a rectangular section (400×400 $mm^2$), with equivalent bending stiffness. This equivalent was derived by equating the product of Young's modulus and cross-sectional area for struts. The Young's modulus and Poisson's ratio for the walers were 47.5 GPa and 0.2, respectively. The groundwater table was assumed stable and located 1 m below the ground surface. The magnitude of the wall movement, as well as ground surface settlement was measured in the FE package ABAQUS 6.13.

According to the conditions described above, a BN model, shown in Fig. 4, was constructed to analyze the failure state, i.e., safe or failure, during the excavation process. In this paper, we only considered the potential risk of the excavation process in the layer of marine clay, as this is where potential failure is most likely to occur. On this basis, we selected three key material parameters: $k$, $M$ and $\lambda$, where $\lambda$ is the logarithmic hardening constant defined for the clay plasticity material behavior; $k$ is the logarithmic bulk modulus of the material defined for the porous elastic material behavior, and $M$ is the ratio of the shear stress. These factors exert substantial influence on the displacement




(a) schematic drawing.

(b) finite element model.

Fig. 3: Lateral cross section of basement excavation in Singapore.

Table 2: Inputs of the parameters for the BN model.

| Parameter | Prior distribution* | Target value | Sensitivity index: $\delta_i$ |
|---|---|---|---|
| $\lambda$ | Normal, $\mu$=0.27, $\sigma$=0.04 | - | 0.1398 |
| $M$ | Normal, $\mu$=0.87, $\sigma$=0.07 | - | 0.1694 |
| $k$ | Normal, $\mu \in [0, 0.3]$, $\sigma \in [0, 0.05]$ | $\mu = 0.1, \sigma = 0.03$ | 0.3286 |
| $H$ | Uniform, $H \in [11, 18.6]$ | - | 0.3575 |

*$\mu$ and $\sigma$ denotes the distribution parameter mean and standard deviation, respectively.

of the wall. The proposed network also considered the effect of different excavation depths. In the BN model, node $H$ represented the varying height of the pit excavation from the ground surface, mainly from stages 4 to 6 of excavation in the marine clay, approximately 11.0 to 18.6 m, as this is a crucial factor in the stability of excavation. Given the difficulty of monitoring the wall deformation, ground surface settlement was incorporated as a field observation used to update the material parameters. Furthermore, the sensitivity analysis in the BN was used to identify the key objects to update.

This paper regards the whole process of excavation as a continuous process. We assumed that node $H$ followed the uniform distribution between [11, 18.6] and other parameters were also defined with the known distributions, per Table 2. Based on the sensitivity analysis, $k$ and $H$ have the greatest effect on the deformation of retaining wall. So, only parameter $k$ was selected as the update target used herein to demonstrate the proposed method. Moreover, the initial values of mean $\mu_k$ and standard variation $\sigma_k$ for variable $k$ were initially estimated with an interval from the limited information, though the true

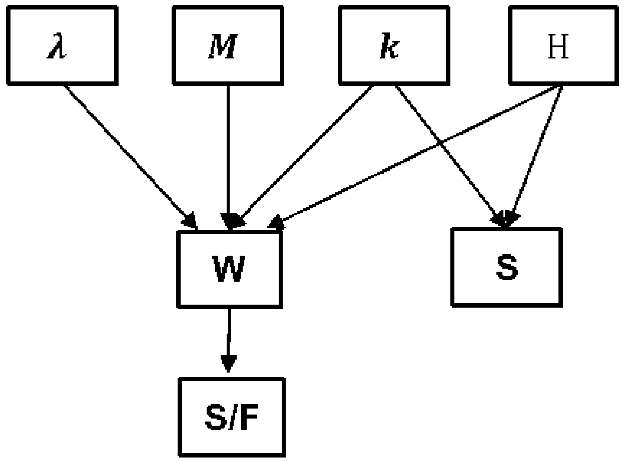



Fig. 4: The Bayesian network for braced excavation.

value is provided as a reference to validate the credibility of the outcome. Hence, they were assumed to be known with the exact values of 0.1 and 0.03, respectively, and were then compared with the updated results of $\mu_k$ and $\sigma_k$.

Given the preceding factors, the observed samples of ground subsidence were generated by Monte Carlo sampling from the input variables $H$ and $k$ in the built input-output model. The distribution parameters of variable $k$ used the target mean and standard deviation listed in the 3rd column of Table 2. The size of the observation sample was set to 100.

## 4. Results

Before updating, the sensitivity indices of input parameters on of the wall movement were computed to identify the key factors. As shown in the last column of Table 2 shows, the ranking of importance was $H > k > M > \lambda$. Given their uncertain influence on wall movement, we mainly consider $k$ and $H$ for the purpose of parameter updating. The dependency relation of nodes $W$ and $S$ with the causal nodes were quantified with the FE method, respectively. Then, the response of wall movement and ground subsidence on uncertain input parameters were predicted using the ANN method.

Figure 5 shows the regression plots of the retaining wall and ground surface settlement, and the results of the ANN prediction are detailed in Table 3. We note that both exhibit strong linear relationships with the input parameters. The associated mean square error for each parameter was also calculated. The model results indicate that the ANN models can be employed to predict wall defection and ground surface settlement. The next step used the input-output black-boxes, and then estimated the uncertainty quantification in the BN.

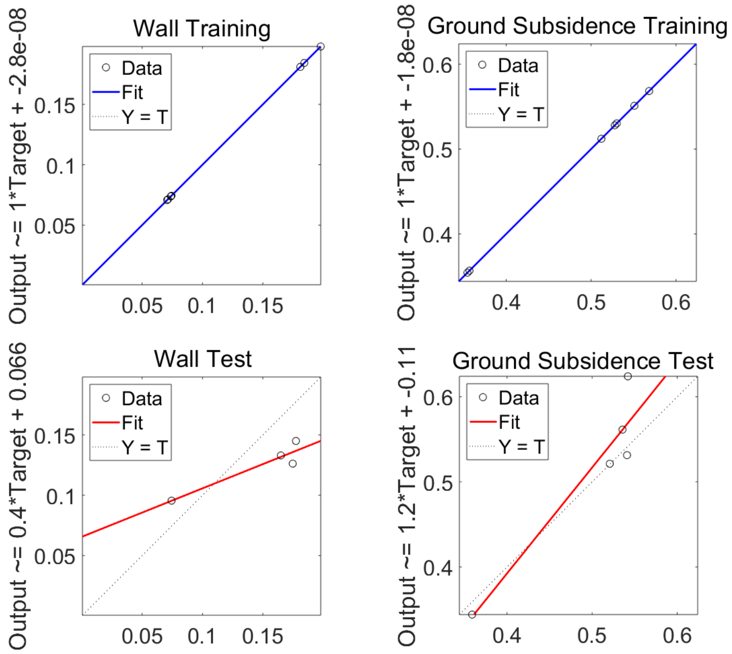
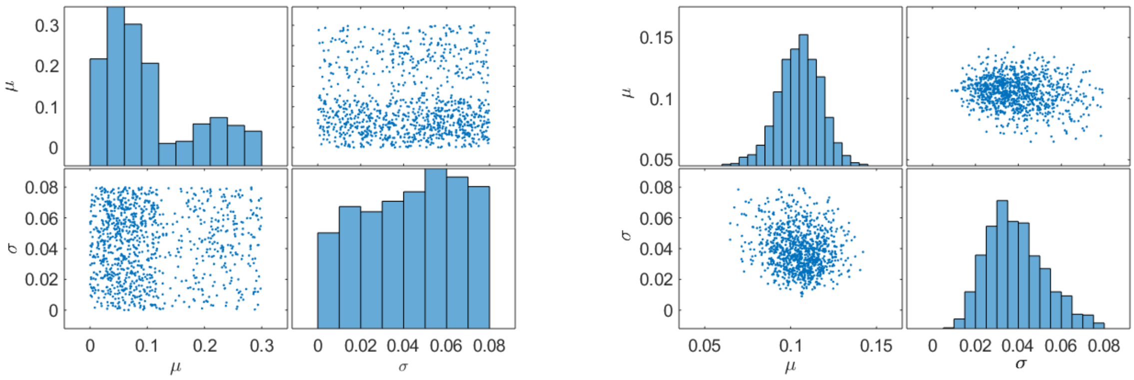



Fig. 5: Neural network training regression

Table 2: The results of the prediction with the ANN.

| Parameter | MSE | R-square(%) |
|---|---|---|
| $W_{training}$ | 3.84e-15 | 99.99 |
| $W_{testing}$ | 1.45e-3 | 93.85 |
| $S_{training}$ | 2.71e-15 | 99.99 |
| $S_{testing}$ | 1.54e-3 | 94.32 |

As the depth of excavation $H$ is a random variable with aleatory uncertainty, only two statistic parameters of $k$ as inputs were used in the distance-based updating procedure. The observed values of variable $S$ were defined as previously described, and used as monitoring data for variable $S$. In this example, the prior distributions of distribution parameters: $\mu_k$ and $\sigma_k$ were set with an interval, following the uniform distributions. Based on each distribution parameter, a group of prior values of variable $S$ can be obtained by the model evaluation and comprise the that is predicted samples seen in Fig. 2. Subsequently, the distance metric for ABC updating can be computed based on Eq. (8) and Eq. (9). After executing the first step of updating with using the Euclidian distance metric, the posterior distributions of variable $k$'s mean and standard deviation are shown in Fig. 6(c), where a proper width coefficient $\varepsilon$ was set as 0.0015, and 13 TMCMC iterations are executed to reach the convergence. Table 4 shows the mean values of their posterior distributions. Comparing with the prior uniform distribution, the posterior distribution of $\mu_k$ accurately converged to the target value, where the updated value of $\mu_k$ was 0.1067.




(c) Results with the Euclidian distance. (d) Results with the Bhattacharyya distance.

Fig. 6: The posterior distribution of distribution $k$ parameters in variablek after updating with distance-based ABC.

From the histogram plot of $\sigma_k$ in Fig. 6(c), we observe that the posterior distribution of $\sigma_k$ was still nearly uniform. That means the standard deviation of variable $k$ was incapable of updating using the Euclidian distance metric. Therefore, further updating based on Bhattacharyya distance must be conducted.

In the second update step, six TMCMC iterations were implemented with the width coefficient of $\varepsilon = 0.08$ in the ABC update process. As seen in Fig. 6(d), the posterior distributions of $\mu_k$ and $\sigma_k$ are dramatically more peaked than those in Fig. 6(c), while both remained close to their respective targets ($\mu_k$ =0.1 and $\sigma_k$=0.003). In the last row of Table 4, the updated values $\mu_k$ and $\sigma_k$ are 0.1052 and 0.0381, respectively. Thus, the second update step reduced the discrepancy between the initial sample and the observation sample. This also demonstrated that the Bhattacharyya distance improved prediction accuracy, where the update results from the Euclidian distance were used as the prior distribution of input in the second update step. Thus, the distributions of $\mu_k$ and $\sigma_k$ were distinctly more centralized to the target values. Using the soil parameter update results, the corresponding failure probability of the excavation was then obtained accordingly in the network. The failure probability of the braced excavation can be obtained in real-time within a few seconds using the BN software. In this paper, OpenCossan (Edoardo et al., 2017) was used to execute the computation. The failure probability of the examined excavation was 95.02%. The updated failure state was provided for the appropriate decision-makers.

Table 3: Values of posterior distributions.

| Parameter | $\mu_k$ | $\sigma_k$ |
|---|---|---|
| Target value | 0.1 | 0.03 |
| Updated with Euclidian distance | 0.1067 | - |
| Updated with Bhattacharyya distance | 0.1052 | 0.0381 |

# 5. Summary

This paper introduced a novel framework for incorporating a Bayesian network with a distance-based ABC real-time update method. The proposed framework estimates real-time failure probabilities using ground surface settlement as observation data collected during the excavation process. The distanced-based ABC approach updates the soil parameters for the model using an input-output black-box. This update method overcomes the limitation of the complex likelihood function and proved effective in reducing the discrepancy between the updated soil parameters and pre-design. Both Euclidian and Bhattacharyya distance-based ABC update methods are computed in the example. The results in the example demonstrate that the mean value of the distribution of a soil parameter approximates the true value using the distance-based ABC updating. However, for the variance of this soil parameter, only the updated value that used Bhattacharyya distance-based ABC was close to the target value. That being said, the prior distribution is also a key factor in the accuracy of this update approach, and must be defined reasonably.

Given the reduced number of calculation points, moment-independent sensitivity analysis applied prior to updating can provide the vital information about which factors exert the greatest influence on the safety or failure state of the excavation. Furthermore, the sensitivity analysis approach captures the dependency relations amongst the nodes in the network. Moreover, it is especially suitable for estimating the variables in the large structure of the BN. With the ranking information, the key parameters can be selected to link with the monitor parameters in the BN model. Based on the updated soil parameters, the uncertainty of the induced-factors is then captured by the BN model. The real-time updated probabilities of the failure in the excavation process can then be estimated, the results of which provide valuable information to support decision-makers.

# Acknowledgment

This work is supported by Chinese Scholarship Council.

## References

Anjum, M., Tasadduq, I., AlSultan, K., 1997. Response surface methodology: A neural network approach. European Journal of Operational Research 101, 65–73.doi:10.1016/S0377-2217(96)00232-9.

Beaumont, M., Zhang, W., Balding, D., 2002. Approximate Bayesian computation in population genetics. Genetics. 162, 2025–2035.

Bi, S., Broggi, M., Beer, M., 2018. The role of the Bhattacharyya distance in stochastic model updating. Mechanical Systems and Signal Processing 117, 437–452.

Borgonovo, E., 2007. A new uncertainty importance measure. Reliability Engineering and System Safety 92, 771–784.

Botev, Z., 2010. Kernel density estimation using matlab .

Botev, Z., Grotowski, J., Kroese, D., 2010. Kernel density estimation via diusion. Annals of Statistics 38, 2916–2957. doi:10.1214/10- AOS799.

Calvello, M., Finno, R., 2004. Selecting parameters to optimize in model calibration by inverse analysis. Computers and Geotechnics 31, 411–425.

Chan, H., Darwiche, A., 2004. Sensitivity analysis in Bayesian networks: From single to multiple parameters. Artificial Intelligence 25, 67–75.

Ching, J., Chen, Y., 2007. Transitional Markov Chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. Journal of Engineering Mechanics 133, 816– 832. doi:10.1061/(ASCE)0733-9399(2007)133:(816).

Do, T., Ou, C., Chen, R., 2016. A study of failure mechanisms of deep excavations in soft clay using the finite element method. Computers and Geotechnics 73, 153–163. doi:10.1016/j.compgeo.2015.12.009.

Edoardo, P., Broggi, M., Tolo, S., Sadeghi, J., 2017. Cossan software a multidisciplinary and collaborative software for uncertainty quantification , 212–224.

Finno, R., Calvello, M., 2005. Supported excavations: Observational method and inverse modeling. Journal of Geotechnical and Geoenvironmental Engineering 131, 826–836. doi:10.1061/(ASCE)1090-0241(2005)131:7(826).

Hashash, Y., Jung, S., Ghaboussi, J., 2004. Numerical implementation of a neural network based material model in finite element analysis. International Journal for Numerical Methods in Engineering 59, 989–1005.

Hashash, Y., Levasseur, S., Osouli, A., Finno, R., Malecot, Y., 2010. Comparison of two inverse analysis techniques for learning deep excavation response. Computers and Geotechnics 37, 323–333. doi:10.1016/j.compgeo.2009.11.005.

Janin, J., Dias, D., Emeriault, F., Kastner, R., Bissonnais, H.L., Guilloux, A., 2015. Numerical back-analysis of the southern toulon tunnel measurements: A comparison of 3d and 2d approaches. Engineering Geology 195, 42–52. doi:10.1016/j.enggeo.2015.04.028.

Jensen, F., 2001. Bayesian networks and decision graphs. Springer-Verlag, Heidelberg Press .

Juang, C., Luo, Z., Atamturktur, S., Huang, H., 2013. Bayesian updating of soil parameters for braced excavations using field observations. Journal of Geotechnical and Geoenvironmental Engineering 139, 395–406. doi:10.1061/(ASCE)GT.1943-5606.0000782.

Kung, G., Juang, C., Hsiao, E., Hashash, Y., 2007. Simplified model for wall deflection and ground-surface settlement caused by braced excavation in clays. Journal of Geotechnical and Geoenvironmental Engineering 133, 731–747. doi:10.1061/(ASCE)1090-0241(2007)133:6(731).

Laskey, K., 1995. Sensitivity analysis for probability assessments in bayesian networks. IEEE Transactions on System 25, 901–909.

Ledesma, A., Alonso, A.G.E., 1996. Three-dimensional finite element analysis of deep excavations. International Journal for Numerical and Analytical Methods in Geomechanics 20, 119–141. doi:10.1002/(SICI)1096-9853(199602)20:2¡119::AIDNAG810 ¿3.3.CO;2-C.

Lee, F., Hong, S., Gu, Q., Zhao, P., 2011. Application of large three-dimensional finite-element analyses to practical problems. International Journal of Geomechanics 11, 529–539. doi:10.1061/(ASCE)GM.1943-5622.0000049.

Li, C., Mahadevan, S., 2017. Sensitivity analysis of a Bayesian network. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering 4, 1–10.

Li, W., Gang, Z., 2018. Protection of cement-soil reinforced regions for adjacent running tunnels during pit excavation. Chinese Journal of Rock Mechanics and Engineering 37. doi:1000- 6915(2018)Supp.1-367312.

Liu, Y., Lee, F., Quek, S., Chen, J., Yi, J., 2015. Effect of spatial variation of strength and modulus on the lateral compression response of cement-admixed clay slab. Geotechnique 65, 851–865.

MacKay, D., 1992. Bayesian interpolation. Neural Computation 4, 415–447.

Mia, M., Dhar, N., 2016. Response surface and neural network based predictive models of cutting temperature in hard turning. Journal of Advanced Research 7, 1035–1044. doi:10.1016/j.jare.2016.05.004.

Ou, C., Chiou, D., Wu, T., 1996. Three-dimensional finite element analysis of deep excavations. JOURNAL OF GEOTECHNICAL ENGINEERING-ASCE 122, 337–345. doi:10.1061/(ASCE)0733- 9410(1996)122:5(337).

Pearl, J., 1988. Probabilistic reasoning in intelligent systems. Morgan Kaufmann Publishers, San Mateo, California.

Spackova, O., Straub, D., 2013. Dynamic bayesian network for probabilistic modeling of tunnel excavation processes. Computer-Aided Civil and Infrastructure Engineering 28, 1–21.

Turner, B., Zandt, TV., 2012. A tutorial on approximate Bayesian computation. Journal of Mathematical Psychology 56, 69–85. doi:10.1016/j.jmp.2012.02.005.

Wei, P., Zheng, Z., Yuan, X., 2013. Monte Carlo simulation for moment-independent sensitivity analysis. Reliability Engineering and System Safety 110, 60–67.

Yuan, R., Bai, G., 2011. New neural network response surface methods for reliability analysis. Chinese Journal of Aeronautics 24, 25–31. doi:10.1016/S1000-9361(11)60004-6.

Zdravkovic, L., Potts, D., John, H.S., 2005. Modelling of a 3d excavation in finite element analysis. Geotechnique 55, 497–513. doi:10.1680/geot.2005.55.7.497.

Zhang, L., Wu, X., Ding, L., Skibniewski, M., Yan, Y., 2013. Decision support analysis for safety control in complex project environments based on bayesian networks. EXPERT SYSTEMS WITH APPLICATIONS 40, 4273–4282.

Zhang, L., Zhang, J., Zhang, L., Tang, W., 2011. Back analysis of slope failure with Markov Chain Monte Carlo simulation. Computers and Geotechnics 37, 905–912. doi: 10.1016/j.compgeo.2010.07.009.

Zhou, Y., Li, C., Zhou, C., Luo, H., 2018. Using Bayesian network for safety risk analysis of diaphragm wall deflection based on field data. Reliability Engineering System Safety 180, 152– 167. doi:10.1016/j.ress.2018.07.014.

4

# 5 Research article III: Sensitivity analysis of prior beliefs in advanced Bayesian networks

5

## 5.1 Declaration of my contribution

The article was written by myself and the idea of this manuscript benefited from the discussions with the co-authors. I carried out all the computation and analysis included in this paper. The article was improved after the comments from the co-authors, and further useful comments and suggestions from the reviewers.

## 5.2 Published article

This paper has been published in the proceeding of the 2019 IEEE Symposium Series on Computational Intelligence, December 6-9, Xiamen, China, and will be published in IEEE Xplore.

## Sensitivity analysis of prior beliefs in advanced Bayesian networks

Longxue He[1], Michael Beer[1,2,3], Matteo Broggi[1], Pengfei Wei[4], António Topa Gomes [5]

[1]Institute for Risk and Reliability, Leibniz University Hannover, Germany
[2]Institute for Risk and Uncertainty, University of Liverpool, Peach Street, Chadwick Building, Liverpool, L69 7ZF, UK
[3]CERSM, Tongji University, 1239 Siping Road, 200092 Shanghai, China
[4]Northwestern Polytechnical University
Xi'an, China
[5]CONSTRUCT, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, s/n

4200-465, Porto, Portugal
*Corresponding author, e-mail: beer@irz.uni-hannover.de
Submitted: July, 2019

## Abstract

Bayesian Network (BN) is an efficient model tool for approximate reasoning based on machine learning. It has been widely used for supporting the decision in many engineering applications such as geotechnical engineering. However, the current studies on BN are mostly on uncertainty quantification and decision-making, while the sensitivity analysis on BN, which may provide much more insights for decision-making, has not received much attention. The current research on sensitivity analysis of BN mainly focuses on local method, and there is a need to develop global sensitivity analysis (GSA) for both forward and backward inferences of BN. We present in this paper GSA analysis for BN within two different settings. For the first setting, it is assumed that the BN nodes, as well as their connection are characterized by precise (conditional) probabilities, and we introduce GSA for both forward and backward analysis. It is shown that, by forward analysis, the GSA indices can effectively identify the nodes which make the most contribution to the end nodes directly related to the reliability; by backward analysis, the GSA indices can inform the most important information needs to be collected for BN model updating. The second setting concerns the incomplete knowledge of nodes and their connections, and it is assumed these quantities are characterized by imprecise probability models. In this setting, the GSA is then introduced, and implemented with the newly developed non-intrusive imprecise stochastic simulation (NISS) method, for learning the most important epistemic uncertainty sources, by reducing which the robustness of the BN inference can be enhanced the most. The above theoretical developments are then applied to an infinite slope reliability analysis problem.

## Index Terms

Global sensitivity analysis, Dependence measure, Advanced Bayesian networks, Imprecise stochastic simulation, Second-order probability model

# I. Introduction

Bayesian network (BN) is a popular graphical tool modeling the dependency relationship among a set of random variables. Due to the flexible structure, it has been widely used in many areas, including machine learning, for statistical inference, supporting decision making, etc. The BN consists of nodes (correspond to the random variables or events) and edges (correspond to conditional probabilities) connecting these nodes. This tool is especially useful for inference based on data when the physical process of interest can be decomposed into many subprocesses. In the past decades, many efficient algorithms (Korb and Nicholson,2010) have been developed for dealing with the inference and uncertainty propagation within the framework of BN interior, and thanks to the development of a variety of BN softwares (Aguilera et al., 2011), the implementation of BN for real-world engineering applications has been a trivial task. However, in contrast to the rapid development of BN model and algorithms, the uncertainty quantification and sensitivity analysis, especially when the data available is incomplete/imprecise, has not attracted too much attention, limiting the robustness of statistical inference and the mining of valuable information embedded in the BN models.

The sensitivity analysis has been applied to the BNs for identifying the sensitivities of nodes (Bednarskia et al., 2004; Chan, 2009). Traditional sensitivity analysis in a BN is concerned by the changes to the nodes of interest in the model, where the changes of the node of interest are observed by querying each of states in causal nodes in the model. Both of single and multiple parameters have been studied for the sensitivity analysis in BNs (Chan and Darwiche, 2004; Coupand der Gaag, 2002), which is restrained in a certain query. Concerning the conditional probability tables (CPTs) associating to each node in the network, the methods of traditional sensitivity analysis for BNs will lead to the high computation cost and time-consuming for the complex networks and especially, insufficient for the model with continuous variables and imprecise variables. In this context, we propose a sensitivity analysis framework for the advanced BNs, to identify the effect of the uncertain nodes on the nodes of interest under the condition of incomplete information.

Among all the GSA techniques, the variance-based method is considered as one of the most appealing methods, which measures the relative importance of each input variable based on its contribution to the model output. Li and Mahadevan (2017) investigates the first-order Sobol' index for the BNs to analyze the sensitivity of nodes. To some extent, this proposed GSA for measuring the sensitivity in the precise network can obtain the influence of the input nodes on the variance of the observed node. However, considering the problems of risk and reliability with the BN approach, the decision-makers could obtain useful information based on analyzing the uncertainty of the target node. The variance-based sensitivity measure summaries the information with a single value, which can lead to the risk of information loss through the uncertainty propagation in the network. Therefore, the sensitivity measure should focus on the entire distribution, rather than the limit on one of its moments.

In this regard, thereof, we propose the moment-independent sensitivity analysis method, simultaneously a dependence measure with copulas also are conducted as a comparison, to measure the sensitivity propagation in the network. This process is conducted with the prior knowledge of variables before updating the precise network. Moreover, a further estimation for the failure domain is investigated with the global reliability SA methods. On the other hand, in the real world, sometimes our knowledge for solving a problem is scarce. To deal with imprecise information situation, the concept of imprecise probability (Beer et al., 2013) is broadly utilized to quantify the vague information. In views of this, BNs are extended to the imprecise networks (Cozman, 2000), termed as Credal networks (CN), which can be regarded as a set of BNs. In this paper, we use a method of the second-order probability model to capture the epistemic uncertainty, where sparse data are only considered.

Generally, an attempt is made to analyze the sensitivity for the precise and imprecise networks. With the proposed approaches, the key induced-factors are identified before information updating for the networks. As well, we could know which factors affect the failure domain most. The paper is divided into four sections in total. In Section II, we introduce how to realise the connection for the sensitivity measure among the indirect nodes. After that, an overview of the GSA methods for the BNs with precise information is presented in III. Also, the proposed GSA framework for the BNs with imprecise information is present in Section IV. A slope stability problem is used to illustrate the proposed methods in Section V, and finally, Section VI provides conclusions.

## II. Exploring the uncertainty propagation in a Bayesian network

In this paper, the uncertainty propagation in the directed path of a BN is studied. For instance, a BN is shown in Fig. 1, the arcs serve to capture the conditional dependencies between two nodes, and there is no directly interplay for the nodes without the link.

Let the nodes be represented by the variables $X_i$ $(i = 1, 2, 3, 4, 5)$, and $\pi(X_i|Pa(X_i))$ denote the corresponding conditional probability distributions (CPDs), where $Pa(X_i)$ denote ancestor nodes of $X_i$. A decision function is essential to model the input-output relationship of the target node with other nodes, and thus, the mapping way should be determined before calculating the sensitivity of the indirect nodes. Regarding the prior CPDs associated with each node, the auxiliary variable method is attested to be feasible for building the bridge between a target node and the indirect nodes (Li and

Mahadevan, 2017). Briefly, the hierarchical propagation: map from $X_4$ to $X_1$ is

$$\begin{cases} X_2 = \pi^{-1}(U_{X_2}|X_4, X_5) \\ X_1 = \pi^{-1}(U_{X_1}|X_2, X_3) \end{cases} \tag{1}$$

where $\pi^{-1}(\cdot)$ is the inverse CPDs of node $X_i$. The auxiliary variable $U_{X_i}$ follows the standard uniform distribution. So, the input-output model for this BN can be expressed as the general form $X_1 = g(U_{X_1}, U_{X_2}, X_3, X_4, X_5)$.



However, in case of continuous variables involving the network, such as the enhanced BN (Straub and Der Kiureghian,2010b), which is the integration of BNs with the structure reliability methods. It is a hybrid BN including both discrete variables and continuous variables, and continuous nodes should be defined in a finite sample space. In this case, the notion of recursion for mapping from $X_4$ to $X_1$ is proposed herein. Assume that the CPDs of the intermediate nodes: $X_1$ and $X_2$, are determined by its parents nodes. Specifically, assuming the so-called decision functions:$X_2 = g(X_4, X_5)$; $X_1 = g(X_2, X_3)$. Then, according to the chain rule of BNs, the connection from $X_4$ and $X_5$ to $X_1$ is bridged as:

$$\begin{cases} \pi(X_2|X_4, X_5) = \int\int_{\Omega_{X_2}} \pi(X_4)\pi(X_5)dX_4\, dX_5 \\ \pi(X_1|X_2, X_3) = \int\int_{\Omega_{X_1}} \pi(X_2|X_4, X_5)\pi(X_3)dX_2\, dX_3 \end{cases} \tag{2}$$

where $\pi(\cdot)$ denotes the CPD of a node. $\Omega_{X_j}$ is the outcome domain of variable $X_j (j = 1, 2)$, which is defined by the performance function $g(\cdot)$.

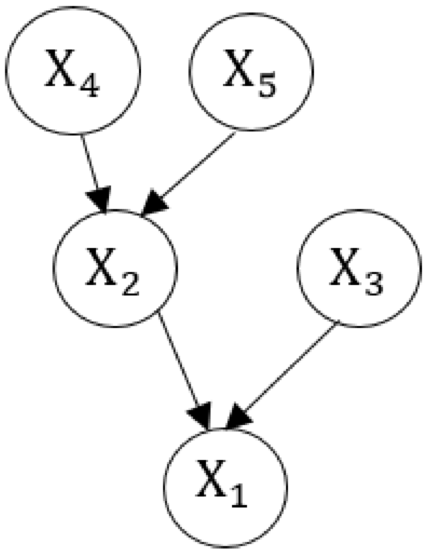



Fig. 1: A directed path in a simple Bayesian network.

# III. Sensitivity analysis in the precise networks

## A. Overview of moment-independent sensitivity analysis

Assume that a computational model: $Y = g(X_i)$, where $i = 1, 2, ..., n$ is any of a model input. The unconditional probability distribution dunction (PDF) of $Y$ is denoted as $f_Y(y)$. The conditional PDF of $f_{Y|X_i}(y)$ denotes the posterior distribution of $Y$ given $X_i$. The PDF-based sensitivity measure can be described based on the discrepancy between the prior distribution and posterior distribution of $Y$:

$$s(X_i) = \int \left| f_Y(y) - f_{Y|X_i}(y) \right| dy. \tag{3}$$

To evaluate the effect of the single input on the PDF of model output, the moment independent sensitivity index (Borgonovo, 2007) can be computed by:

$$\delta_i = \frac{1}{2} E_{X_i}[s(X_i)] = \frac{1}{2} \int f_{X_i}(x_i) \times \left[ \int \left| f_Y(y) - f_{Y|X_i}(y) \right| dy \right] dx_i, \tag{4}$$

where $\delta_i \in [0, 1]$ and 0 means $X_i$ is non-influential while 1 means the importance of all the input parameters for $Y$ is equal. Single-loop Monte Carlo simulation method is utilised herein to estimate the sensitivity indicators. In (Wei et al., 2013), the form of $\delta_i$ is further considered by,

$$\begin{aligned} \delta_i &= \frac{1}{2} \int \int |f_{X_i}(x_i) f_Y(y) - f_{Y,X_i}(y, x_i)| dy\, dx_i \\ &= \frac{1}{2} E_{Y,X_i}\left( \left| \frac{f_{X_i}(x_i) f_Y(y)}{f_{Y,X_i}(y, x_i)} - 1 \right| \right). \end{aligned} \tag{5}$$

For a group of observed parameters $R = (X_{i1}, X_{i2}, \ldots, X_{in})$, likewise the delta indice is given by

$$\delta_i = \frac{1}{2} E_{Y,R}\left( \left| \frac{f_R(x_{i1}, x_{i2}, ..., x_{in}) f_Y(y)}{f_{Y,R}(y, x_{i1}, x_{i2}, ..., x_{in})} - 1 \right| \right), \tag{6}$$

where $f_Y(y)$ is estimated with kernel density estimator (KDE) method, and a bivariate KDE toolbox (Botev, 2010) is used for achieving the joint PDF of $Y$ and $R$. Additionally, Monte Carlo simulation is employed to compute the sensitivity index $\delta_i$.

## B. Review of dependence measure with copula

In Eq. (5), obviously, the role of input and output is exchangeable. That is, the delta index can represent mutual degree of the sensitivity. Thereof, their dependence can be expressed by use of copula approach. The joint CDF of random variables $Y$ and $X_i$ can be given by,

$$F_{Y,X_i} = C(F_{X_i}(x_i), F_Y(y)) = C(\mu, \nu_i), \tag{7}$$

here $C(\cdot,\cdot)$ is the copula function with uniform marginal distribution in I : [0,1], and $u = F_{X_i}(x_i)$, $v_i = F_Y(y)$.



Based on the idea of Eq. (7), the aforementioned delta index can be extended by the following formula (Schweizer and Wolff, 1981):

$$\delta_i^E = 12 \int \int_I (|C(\mu,\nu_i) - \prod(\mu,\nu_i)|)\, d\mu,\, d\nu_i. \tag{8}$$

in which $\prod(u, v_i) = uv_i$. This extended delta index measures how much the variable $Y$ is dependent on the material parameter $X_i$. One way to estimate the extended delta index with the empirical copula-function (Wei et al., 2014) is employed in the following calculation. Briefly, the extended delta index $\delta_i^E$ referring to the empirical copula-function can be written as:

$$\delta_i^E = 12 \sum_{p=1}^{N-1} \sum_{q=1}^{N-1} |C_N(\mu',\nu') - \prod(\mu',\nu')| \times S(p,q)|, \tag{9}$$

where $\mu'=\frac{\mu_{(p)}+\mu_{(p+1)}}{2}$, $\nu'=\frac{\nu_{(q)}+\nu_{(q+1)}}{2}$ and $C_N(\mu',\nu')$ represents the empirical copula value of midpoint of the $(p,q)$th $(p \geq 1, q \leq N-1)$ element. $S(p,q)$ is the area of the $(p,q)$th element.

## C. Global reliability sensitivity analysis for the enhanced BN

The method of global reliability sensitivity analysis is based on the Sobol' indices. Instead of measuring the effect of the uncertainty of input on the variance of output, global reliability sensitivity indices represents both of the main and total effect contribution of input to the variance of the failure domain of output.

An enhanced BN (Straub and Der Kiureghian, 2010a) is a BN combining with structure reliability methods. Considering the structure of enhanced BNs, the global reliability index is suitable to be computed for measuring the importance of nodes in the network. The method can tell us which the uncertainty of nodes affect the failure state of a node most, and the interaction accordingly on the failure probability can be obtained as well.

An enhanced BN for the slope stability analysis in Figure 2 is used as an example, which

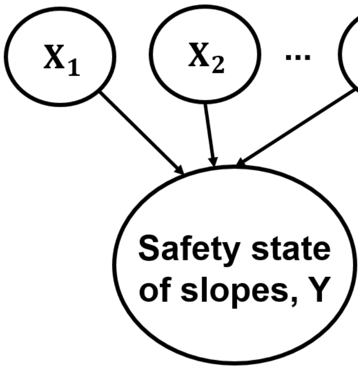



Fig. 2: An enhanced Bayesian network of a slope.

the network is considered with the reference in (Li and Zhang, 2018). The safety state of a slope: failure or safety, is denoted with node $\boldsymbol{Y}$. The causal nodes $\boldsymbol{X}$ ($\boldsymbol{X}$=$(X_1, X_2, ..X_n)$ are a series of geotechnical parameters affecting the node $\boldsymbol{Y}$. According to the chain rule of BNs, the joint probability can defined as:

$$P(\boldsymbol{Y}, \boldsymbol{X}) = \pi(\boldsymbol{Y}|\boldsymbol{X})\pi(\boldsymbol{X}). \tag{10}$$

where $\pi(\cdot)$ denotes the probability measure of nodes. Further, in light of the properties of enhanced BNs, Eq. (10) can be rearranged as:

$$P(\boldsymbol{Y}) = \int_{\boldsymbol{X}\epsilon\Omega_{\boldsymbol{Y}}(\boldsymbol{X})} \pi(\boldsymbol{X})d\boldsymbol{X}. \tag{11}$$

The domain $\Omega_{\boldsymbol{Y}}(\boldsymbol{X})$ of the node $\boldsymbol{Y}$ is defined by the limited state function $G(\boldsymbol{X})$, where the indicator $I(\boldsymbol{X})$ is given by

$$I(\boldsymbol{X}) = \begin{cases} 1, & \text{Failure} : G(\boldsymbol{X}) < 0 \\ 0, & \text{Safe} : G(\boldsymbol{X}) \geq 0 \end{cases} \tag{12}$$

then, the failure probability of node $\boldsymbol{Y}$ can be written as

$$\begin{aligned} P_f &= \int_{\Omega_{\boldsymbol{Y}}(\boldsymbol{X})} I_F(\boldsymbol{X})\pi(\boldsymbol{X})d\boldsymbol{X} \\ &= E(I_F(\boldsymbol{X})), \end{aligned} \tag{13}$$

here $E(\cdot)$ is the expectation operator. According to Wei et al.(2012), the individual effect index is expressed as:

$$S_i = \frac{V(E(I_F|X_i))}{V(I_F)} = \frac{E(E^2(I_F|X_i)) - P_f^2}{P_f - P_f^2}, \tag{14}$$

where $V(\cdot)$ is the variance and $i = 1, 2, .., n$. The total effect index is expressed as:

$$\begin{aligned} S_{Ti} &= 1 - \frac{V(E(I_F|X_{\sim i}))}{V(I_F)} \\ &= 1 - \frac{E(E^2(I_F|X_{\sim i})) - P_f^2}{P_f - P_f^2}. \end{aligned} \tag{15}$$

where $X_{\sim i}$ means all the causal nodes $\boldsymbol{X}$ except $X_i$.

## IV. Sensitivity analysis in the imprecise networks

The above developments on GSA are based on the assumption that the BN nodes as well as their connected lines, and they are all characterized by precise probability models. This is only applicable when the available information on the nodes and their dependencies is sufficiently big. However, in real-world applications such as geotechnical engineering, the data available for both nodes and their dependencies can be incomplete and/or imprecise, making it impossible to generate precise probability models. In this situation, the imprecise probability models such as the p-box model and second-order probability models can be introduced and injected into the BN framework so as to generate robust inference. Here, we firstly review the second-order probability model.

### A. Second-order probability model

Epistemic uncertainty refers to a variable comes from the lack of knowledge in a subjective point of view when statistics of a variable are described with incomplete and vague information. In the case of imprecise information, Bayesian networks are extended to Credal networks (CNs) (Cozman, 2000), where the CPDs of nodes are defined with imprecise probabilities. In this paper, we only consider data uncertainty, where scarce point data and/or interval data are available. Moreover, a likelihood-based probabilistic method (Sankararaman and Mahadevan, 2011) is considered herein to quantify the components in CNs. Let $f(x|\theta)$ denote the probability density function (PDF) of a node, then due to incompleteness/imprecision of the data of $x$, the statistics $\theta$ is uncertain, and characterized by a subject probability model. The probability distribution of $\theta$ can be inferred with Bayesian inference and it is briefly reviewed as follows.

Assume that any of nodes $X$ in a CN is given with poor and sparse data, only distribution type is known with a specific probability distribution $f_X$. Using the likelihood-based method, the distribution parameters $\theta$ of variable $X$ is inferred with limited point data $x_i$ $(i = 1, 2, ..n)$ and/or interval data $[x_{li}, x_{ui}]$. Given the observed data, the likelihood function of parameters $\theta$ can be defined with the conditional PDF of variable $X$ as,

$$L(\theta) \propto \sqcap_{i=1}^{n} f_X(x_i|\theta). \tag{16}$$

Similarly, for the several interval data, the likehood function of parameters $\theta$ are expressed as,

$$L(\theta) \propto \sqcap_{i=1}^{n} \int_{x_{li}}^{x_{ui}} f_X(x_i|\theta)dx. \tag{17}$$

Note that the right of equation can further be calculated with the cumulative distribution function $F_X(x_i|\theta)$. Then in the case of the combination of these two types of data, the likelihood function $L(\theta)$ can be described by the multiplication:

$$L(\theta) \propto (\sqcap_{i=1}^{n} f_X(x_i|\theta)) \times (\sqcap_{i=1}^{n} F_X(x_{ui}|\theta) - F_X(x_{li}|\theta)). \tag{18}$$

To construct the PDFs of parameters $\theta$: $f(\theta)$, the Bayes' theorem is concerned instead of the maximun likelihood estimate, which is more robust with consideration of the entire likelihood function (Sankararaman and Mahadevan, 2011). The expression is

$$f(\theta) = \frac{L(\theta)}{\int L(\theta)d\theta}. \tag{19}$$

With the nodes being characterized by second-order probability models, and their dependencies being characterized by second-order conditional probability models, the BN is said to be an imprecise BN. In this situation, the epistemic uncertainty presented in the distribution parameters and characterized by subjective probability have important effects on the inference results of the BN. With ignoring the effect of the epistemic uncertainty on the results, the prediction as well as decision derived from BN can be misleading. Thus, there is a need to qualitatively measuring the effects of the epistemic uncertainty on the prediction. The process how to identify the effect of uncertain distribution parameters on the target variable in the network are presented in the next section.

## B. Global sensitivity analysis for the imprecise BN

1) The Sobol' index: The Sobol' index is used here (Sobol et al., 2007) to estimate the importance of epistemic uncertainty in a BN with imprecise information. Sobol's consider a model: $Y = g(\boldsymbol{X})$, where the square-integrable function $g(\boldsymbol{X})$ can be decomposed in the following formula:



$$g(\boldsymbol{X}) = g_0 + \sum_i g_i(X_i) + \sum_{j>i} g_{ij}(X_i, X_j) + ... + g_{1,2,...,n}(\boldsymbol{X}), \tag{20}$$

where the expansion of $g(\boldsymbol{X})$, called high-dimensional model representation (HDMR). Each individual term is calculated using the conditional expectations of the model output $Y$.

According to the variance-based sensitivity analysis, the first-order Sobol' index $S_i$ ($i = 1, 2, ...k$) can be expressed as,

$$S_i = \frac{V(g_i(X_i))}{V(Y)} = \frac{V(E(Y|X_i)}{V(Y)}, \tag{21}$$

where $S_i$ means the main effect of each input $X_i$ on the variance of the output $Y$. $V(\cdot)$ means the variance and $E(\cdot)$ means the mean value. For the second-order Sobol's index $S_{i,j}$ is the ratios of the two-way interaction of input variables.

$$S_{i,j} = \frac{V(g_{ij}(X_i, X_j))}{V(Y)} = \frac{V(E(Y|X_i, X_j)) - V_i - V_j)}{V(Y)}, \tag{22}$$

where $V(E(Y|X_i, X_j)$ represent the joint effect of $X_i$ and $X_j$. $V_i$ and $V_j$ are the first-order effect.

2) Imprecise stochastic simulation: In this paper, a method termed as non-intrusive imprecise stochastic simulation (NISS) (Wei et al., 2019) is utilized to capture the uncertainty propagation in the BN with imprecise information. In (Wei et al.,2019), the method is testified efficiently with the imprecise probability models associating with the uncertain distribution parameters, where the global extend Monte Carlo simulation (GEMCS) is combined with the Random Sampling HDMR (RS-HDMR). The use of the RM-HDMR model contributes to reducing the computation cost of high dimension in the performance of the GEMCS estimators.

Based on the above-mentioned method, in this work, we consider the parameterized probability model. The uncertain distribution parameters $\boldsymbol{\theta}$ of input variables $\boldsymbol{X}$ are determined with the approach of the second-order probability model, and we consider the first-order moment and failure probability functions for performing the proposed method, respectively. Taking the response expectation function $E_y(\boldsymbol{\theta})$ of node $y$ as a example, the RS-HDMR decomposition of $E_y(\boldsymbol{\theta})$ can be written as:

$$E_y(\boldsymbol{\theta}) = E_{y_0} + \sum_{i=1}^{k} E_{y_i}(\theta_i) + \sum_{i<j} E_{y_{ij}}(\theta_{ij}) + ... + E_{y_{12...k}}(\boldsymbol{\theta}), \tag{23}$$

where imprecise parameters $\boldsymbol{\theta} = \theta_1, \theta_2, ..., \theta_k$, and $\theta_{ij}$ is the 2-dimension vector from $\theta_i$ and $\theta_j$. Furthermore, each component in the left of Eq. (23) can be computed with the GEMCS-RS-HDMR method. The detail computation process of the RS-HDMR component functions can be obtained can be found in (Wei et al., 2019), which is not presented herein.

Note that in the imprecise probability model, the distribution parameters of variables $\boldsymbol{X}$ are uncertain, so the influence of distribution parameters $\boldsymbol{\theta}$ on the target node $y$ rather than variables $\boldsymbol{X}$ should be analyzed to quantify the epistemic uncertainty propagation in a CN. Therefore, for calculating the Sobol's index, the output here is the expectation $E_y(\boldsymbol{\theta})$ and input variables are the distribution parameters $\boldsymbol{\theta}$. By estimating the variance of $E_y(\boldsymbol{\theta})$ in Eq. (23), the first-order Sobol' sensitivity indexes for the target node can be computed:

$$S_{E_{yi}} = \frac{V(E(E_y(\boldsymbol{\theta})|\theta_i))}{V(E_y(\boldsymbol{\theta}))} \tag{24}$$

and the joint effect can be computed in the same way.

# V. Example for illustration

## A. Problem description

It is common that ubiquitous uncertainties exist in the slope stability analysis. BNs are useful probabilistic models for integrating multi-factors with graphical means (Li and Zhang,2018). It can serve to capture the uncertainty presented in random variables, and also obtain the information of interest by inquiring the network. With these powerful capabilities, this kind of probabilistic networks has been applied to solve the geotechnical problems (Liu et al., 2013; Peng et al., 2014). In this paper, we employ a practical problem about the slope stability analysis from (Phoon, 2008). The stability of this infinite slope (see Fig. 3) is mainly affected by six independent random variables. The uncertain induced-factors determine the variation of the uncertainty on the slope stability, and more detailed description about this slope problem can be found in (Phoon, 2008). The case used here is only to illustrate how to implement the proposed methods for making the sensitivity analysis for the BNs with scarce information. Considering the cause-effect relationship, the BN structure of an infinite slope can be constructed referring to the structural reliability problem as Fig. 4 shows.

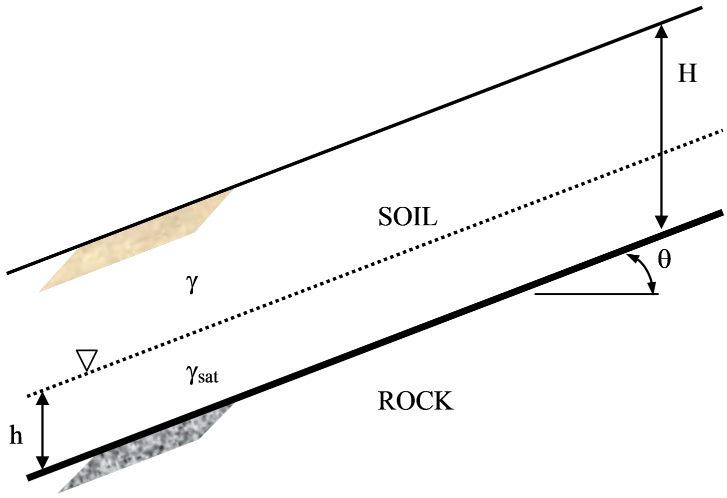



Fig. 3: An infinite slope (Phoon, 2008).

$FS$ represents the factors of safety. Node $S/F$ represents the states of slope stability: safe or failure, and the corresponding conditional probability table can be obtained by the performance function. The definition of slope parameters is presented in Table I, where the distribution of slope parameters are defined with available information. However, sometimes in case of incomplete information, probability distributions of these parameters cannot be obtained. Hence, the likelihood-based probabilistic method is applied herein. In this example, we define two parameters $e$ and $\phi$ with lack of information, where data about variables $e$ and $\phi$ are given with sparse information. To be specific, variables $e$ and $\phi$ follows uniform and lognormal distribution, respectively, and the upper and lower bound of $e$: $b_e$ and $a_e$ and the mean and standard deviation of $\phi$: $\mu_\phi$ and $\sigma_\phi$ are uncertain.

TABLE I: Definitions of parameters

| Variable | Description | Available information* |
|---|---|---|
| $G_s$ | gravity of solids | U(2.5, 2.7) |
| $e$ | void ratio | U($a_e$, $b_e$) |
| $H$ | depth of soil | U(2, 8) |
| $h = H * U$ | height of water table | $U \in$ U(0, 1) |
| $\phi$ | effective stress | log($\mu_\phi$, $\sigma_\phi$) |
| $\theta$ | slop inclination | log(20, 1) |
| $\gamma$ | moist unit weight of soil | $\gamma = \frac{\gamma_w(G_s+0.2e)}{(1+e)}$ |
| $\gamma_{sat}$ | saturated unit weight of soil | $\gamma_{sat} = \frac{\gamma_w(G_s+e)}{(1+e)}$ |

*U(·) and logN(·), represent uniform and lognormal distribution, respectively.

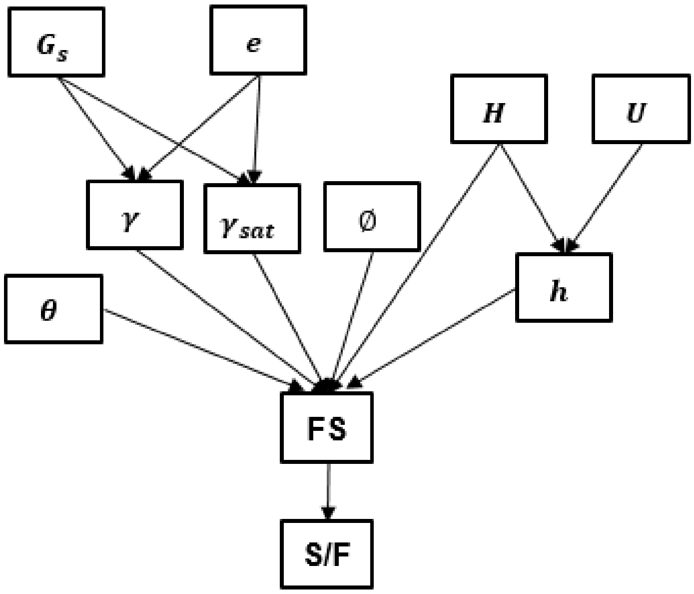



Fig. 4: The BN model of the infinite slope.

## B. Results with the precise network

Here nodes $G_s$, $e$, $U$, $H$, $\phi$ and $\theta$ are selected as six independent inputs and $FS$ is the node of interested. In this paper, we computed the sensitive indexes of these parameters in order to testify the feasibility of the proposed SA approaches. The results of importance analysis can be achieved from Table III. In the 2nd and 3th column, the sensitivity index $\delta_i$ of the moment-independent SA and $\delta_i^E$ of the dependence measure show the same importance ranking: $U > \phi > \theta > e > G_s > H$, and obviously the sensitive values of parameters $U$ and $\phi$ is much larger than other factors. On the ground of this, we can obtain the state of the slope stability is mainly affected by the variation of $U$ and $\phi$ and variable $H$ has the least influence on the change of node $FS$. It also indicates the PDF of the target node $FS$ can be most changed through reducing the uncertainty of important variables. Furthermore, by comparing the two methods for analyzing the sensitivity of BNs, the method of the dependence measure with copula strengthens the results of sensitivity degree to indicate the dependence of the target node on the variation of causal factors in the network. Also, according to the importance ranking, the identified key factors can be selected to update by combining the information with field data for further study.

$S_i$ and $S_{Ti}$ in Table II is computed by global reliability SA, which is used to identify the key input for the failure domain of node $S/F$. The results of sensitivity indices $S_i$ show similar to the ranking of delta indices. Comparing to other factors of the single sensitivity indices $S_i$, the uncertainty of parameters $U$, $\phi$ and $\theta$ contribute to the failure probability of the slope most while other single factors show roughly equal importance.

TABLE II: Sensitivity indices for the precise network

| Parameter | $\delta_i$ | $\delta_i^E$ | $S_i$ | $S_{Ti}$ |
|---|---|---|---|---|
| $G_s$ | 0.0319 | 0.0332 | 0.0802 | 0.3403 |
| $e$ | 0.0393 | 0.0358 | 0.0834 | 0.3624 |
| $U$ | 0.3340 | 0.6235 | 0.3846 | 0.5995 |
| $H$ | 0.0267 | 0.0330 | 0.0812 | 0.6027 |
| $\phi$ | 0.1172 | 0.2020 | 0.1463 | 0.4629 |
| $\theta$ | 0.0658 | 0.0722 | 0.0974 | 0.3746 |



On the other hand, the values of the total effect indexes indicate highly interaction effect of these factors on the slope failure, especially $G_s$, $e$, $H$ shows much interaction effect in comparison with the corresponding sensitivity indices $S_i$. The results reveal the effect of the induced-factors on failure probability is not independent, and they are a joint influence on the change of the failure probability of the infinite slope.

## C. Results with the imprecise network

The proposed framework for computing the sensitivity indices in a Credal network includes the following steps:

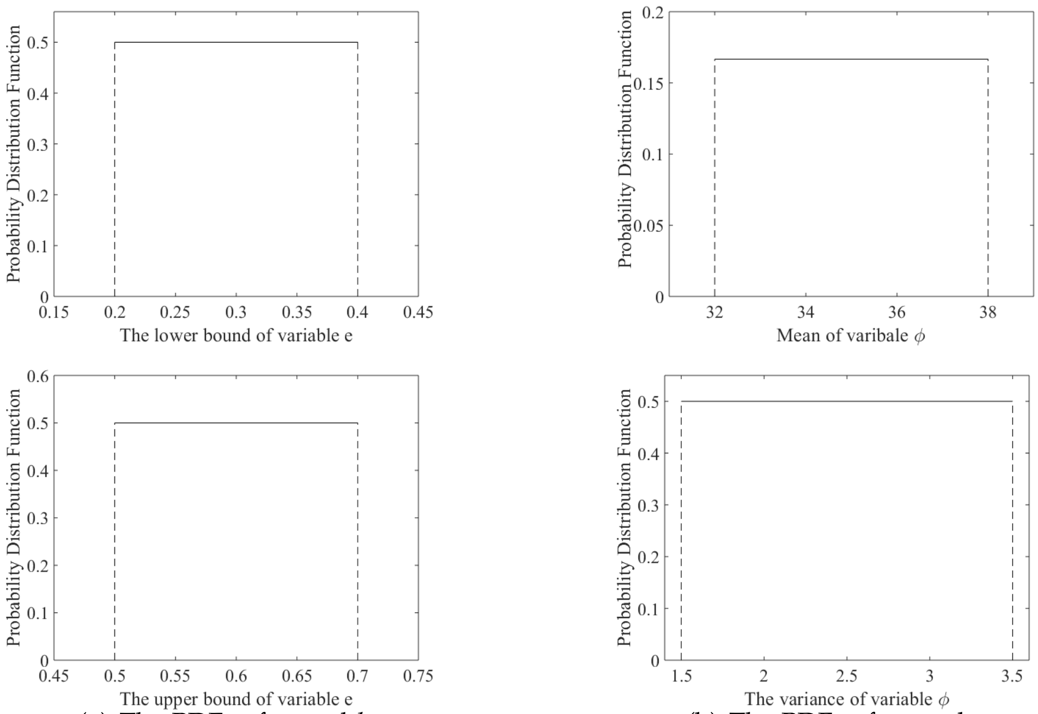
Step 1. According to the given sparse data to compute the PDF of distribution parameters in variables $e$ and $\phi$, the PDFs of their distribution parameter in Fig. 5 can be drawn based on the Eq. (16) - (19).

Step 2. The GEMCS-RS-HDMR procedure is performed for computing the RS-HDMR component functions, where the response expectation function $E_i$ and failure probability function $Pf_i$ of the node of interest $FS$ and $S/F$ are estimated, respectively.

Step 3. Latin-hypercube sampling technique is used to generate the samples. In this example, nodes $FS$ and $S/F$ are observed as the response of epistemic uncertainty, and the samples of nodes $FS$ and $S/F$ are determined through the propagation in the network as mentioned earlier.

Step4. Compute the Sobol' indices to identify which of distribution parameters contribute to the uncertainty of nodes $FS$ and $S/F$ most, where the importance of each component function is quantified by the NISS approach, respectively, as Table III and Table IV shows.

In this computation, 5000 samples are generated for evaluating the sensitivity indices of the RS-HDMR component functions. Table 3 shows the results of the first-order sensitivity indices with NISS approach, For the indices $S_{E_i}$ of the response of node $FS$,




(a) The PDFs of $a_e$ and $b_e$ (b) The PDFs of $\mu_\phi$ and $\sigma_\phi$

Fig. 5: The PDFs of the uncertain parameters.

the importance ranking for the uncertain distribution parameters is $\sigma_\phi < b_e < a_e < \mu_\phi$, while the sensitivity index of parameter $b_e$ is slightly larger than $a_e$ for the response of the failure probability. Generally, parameters $a_e$ and $b_e$ have the roughly identically importance, and the sensitivity indices of parameter $\mu_\phi$ show significant effect most, especially the contribution for the slope failure. The least one is $\sigma_\phi$, actually non-influential. For the sensitivity analysis of the first-order moment function, the influence of $a_e$ and $b_e$ are also significant, however, from the result of the failure probability function, the two distribution parameters are much less influential on the response of $S/F$.

The results of importance measure are different according to the selection of the response variable. The first-order sensitivity indices $S_{Ei}$ show the effect of epistemic uncertainty on the slope stability, can be positive or negative. However, in a different point of view, through the sensitivity analysis of the failure probability, we can obtain which uncertain parameters affect the slope failure. With these results, it can be reasonable to infer the uncertainty of the slope stability in this example can be effective to be reduced by increasing our knowledge of the distribution parameter $\mu_\phi$.

The second-order sensitivity indices in Table IV indicate the interaction of distribution parameters on the change of nodes $FS$ and $S/F$, respectively. All the sensitivity indices of the fist-order moment function$S_{E_i}$ show the interaction of parameters $a_e$ and $\mu_\phi$ are more significant than other pairs. and the joint effect if $(a_e, b_e)$ and $(\mu_\phi, b_e)$ are slight effect. The sensitivity values of $a_e$ and $\sigma_\phi$ in $S_{E_i}$ and $S_{Pf_i}$ are approximately zero. It indicates their interaction on node $FS$ and $S/F$ is non-influential, and can be ignored. Additionally, the influence of the pairs $(a_e, \mu_\phi)$ and $(\mu_\phi, b_e)$ contribute to the slope failure most while other pairs are less influential.



TABLE III: The first-order sensitivity indices

| Indices | $a_e$ | $\mu_\phi$ | $b_e$ | $\sigma_\phi$ |
|---|---|---|---|---|
| $S_{E_i}$ | 0.2250 | 0.4898 | 0.2232 | 0.0001 |
| $S_{Pf_i}$ | 0.0210 | 0.8829 | 0.0250 | 0.0007 |

TABLE IV: The second-order sensitivity indices

| Indices | $S_{E_i}$ | $S_{Pf_i}$ |
|---|---|---|
| $(a_e, \mu_\phi)$ | 0.0338 | 0.0328 |
| $(a_e, b_e)$ | 0.0166 | 0.0021 |
| $(a_e, \sigma_\phi)$ | 0.0001 | 0.0000 |
| $(\mu_\phi, b_e)$ | 0.0100 | 0.0337 |
| $(\mu_\phi, \sigma_\phi)$ | 0.0015 | 0.0018 |
| $(b_e, \sigma_\phi)$ | 0.0000 | 0.0000 |

## VI. Conclusions

This paper presents how to make the sensitivity analysis before the model updating, which can avoid the computation cost. In the proposed framework, the proposed method for transforming an evidence variable to a random variable contributes to the computation of global sensitivity indexes. Further, in case of imperfect information, a method, called non-intrusive imprecise stochastic simulation is employed herein to analyze the importance of nodes in a network, where the combination of GEMCS with RS-HDMR is presented to decompose the expectation and failure probability function of the node of interest. Through the propagation of uncertainty among the nodes in the network, the Sobol' sensitivity indexes of distribution parameters of each influential nodes are computed. Therefore, the effect of the uncertain parameters of causal nodes on the change of the target node can be captured based on the values of sensitivity indexes.

A case with sufficient information and imperfect information is separately is utilised to verify the performance of the proposed method. The ranking sensitivity indexes provide decision-maker with the key induced-factors to integrate with field data, and global reliability sensitivity indices provide the important information of the most influencing factors on the slope failure. Also, the proposed general framework for sensitivity analysis of Bayesian networks with imprecise information is applied for an infinite slope problem, where epistemic uncertainty is quantified with the second-order probabilistic method. After that, in the Bayesian network structure of the slope, the sensitivity indices of imprecise nodes are computed by the non-intrusive imprecise stochastic simulation methods. From the results, we can obtain useful information for the identification of the key uncertain factors. The example indicates the proposed methods is feasible and effective to provide significant information for the decision-makers to reduce the uncertainty of the target event.

## Acknowledgments

This work is supported by Chinese Scholarship Council.

## References

P. Aguilera, A. Fernndez, R. Fernndez, R. Rum, and A. Salmern. Bayesian networks in environmental modelling. Environmental Modelling Software, 26:1376–1388, 2011.

M. Bednarskia, W. Cholewaa, and W. Frid. Identification of sensitivities in Bayesian networks. Engineering cations of Artificial Intelligence, 17:327–335, 2004.

M. Beer, S. Ferson, and V. Kreinovich. Imprecise probabilities in engineering analyses. Mechanical Systems Signal Processing, 37: 4–29, 2013.

E. Borgonovo. A new uncertainty importance measure. Reliability Engineering and System Safety, 92(6):771–784, 2007.

Z. Botev. Kernel density estimation using matlab. 2010.

H. Chan. Sensitivity analysis of Bayesian networks and its application for service engineering. 2009 International Conference of Soft Computing and Pattern Recognition, pages 551–556, 2009.

H. Chan and A. Darwiche. Sensitivity analysis in Bayesian networks: From single to multiple parameters. Proceedings of the 20th conference on Uncertainty in artificial intelligence, pages 67–75, 2004.

VM. Coup and LCV. der Gaag. Properties of sensitivity analysis of Bayesian belief. Annals of Mathematics and Artificial Intelligence, 36(4):323–356, 2002.

FG. Cozman. Credal networks. Artificial intelligence, 120: 199–233, 2000.

K. Korb and A. Nicholson. Bayesian artificial intelligence. Faculty of Information Technology, Monash University, Clayton, Victoria 3800 Australia, 2nd, CRC Press, 2010.

C. Li and S. Mahadevan. Sensitivity analysis of a Bayesian network. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering, 4 (1):10, 2017.

X. Li and L. Zhang. Efficient Bayesian networks for slope safety evaluation with large quantity monitoring information. Geoscience Frontiers, 9(6):1679–1687, 2018.

W. Liu, S. Li, and R. Tang. Slope stability evaluation in open pit based on Bayesian networks. In G. Yang, editor, Advances in Intelligent Systems and Computing, pages 1227– 1231. Springer-verlag Berlin, Heidelberger Platz 3, D-14197 Berlin, Germany, 2013.

M. Peng, XY. Li, and DQ. Li. Slope safety evaluation by integrating multi-source monitoring information. Structural Safety, 49(SI):65–74, 2014.

KK. Phoon. Reliability-based design in geotechnical engineering: Computations and applications. Taylor Francis, London, page Chapter 1, 2008.

S. Sankararaman and S. Mahadevan. Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data. Reliability Engineering System Safety, 96(7):814–824, 2011.

B. Schweizer and E. Wolff. On nonparametric measures of dependence for random variables. Annals of Statistics, 9: 879–885, 1981.

I. Sobol, S. Tarantolab, D. Gatellib, S. Kucherenkoc, and W. Mauntz. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. Reliability Engineering and System Safety, 92:957–960, 2007.

D. Straub and A. Der Kiureghian. Combining Bayesian networks with structural reliability methods: Methodology. Journal of Engineering Mechanics, 136(10):1248–1258, 2010a.

D. Straub and A. Der Kiureghian. Combining Bayesian networks with structural reliability methods: Application. Journal of Engineering Mechanics, 136(10):1259–1270, 2010b.

P. Wei, Z. Zheng, W. Hao, J. Feng, and B. Wang. Efficient sampling methods for global reliability sensitivity analysis. Computer Physics Communications, 183(8):1728–1743, 2012.

P. Wei, Z. Zheng, and X. Yuan. Monte carlo simulation for moment-independent sensitivity analysis. Reliability Engineering and System Safety, 110:60–67, 2013.

P.Wei, Z. Zheng, and J. Song. Moment-independent sensitivity analysis using copula. Risk Analysis, 34(2):210–222, 2014.

P. Wei, J. Song, S. Bi, M. Broggi, and M. Beer. Non-intrusive stochastic analysis with parameterized imprecise probability models: I. performance estimation. Mechanical Systems and Signal Processing, 124:1728–1743, 2019.

# 6 Conclusions



In this study, we make an attempt to deal with the stability and deformation problems of soil using the advanced BN approach. The research in the article mainly focuses on the development of novel methods for enhancing the BN model in order to estimate the risk in geotechnical engineering. The improved BN model is applied in analyzing the risk of slope stability and braced excavation in the case studies, respectively, which is presented in the three research articles. Moreover, the results indicate the feasibility and efficiency of the proposed methods.

Initially, the BN models are built taking into account the case of continuous variables and incomplete information. To prove the potential of the advanced BNs in the application of the slope problem, two BNs models of soil slopes are constructed, which the structure of the model is described in Chapter II. In the network, the probability of slope failure is regarded as the actual risk level of the slope due to the variability of input parameters and the states of drainage. The research illustrates the BN model can provide the decision-makers real-time results about the safety state of slopes, and the discretization approach presented is proved to be efficient for updating the BN model when the observation given on continuous variables. On another hand, on the basis of the available information, the proper method for the definition of input variables in the network affects the uncertainty of the target variable of interest. Through the comparison of BN models of a slope under different information situation, the implementation of Credal networks shows the uncertainty of slope failure can be reduced according to the definition of causal factors related to the selection of proper methods.

In the next step of research, we enhanced the updating capability of advanced BNs (Chapter III), and the method is used in the case study of pit excavation. Based on the advanced BNs approach, soil parameters, geotechnical factors and field observation can be combined with a model. However, uncertain soil materials can mislead the estimation of the risk level during the excavation. Bayesian networks possess the powerful ability of real-time information updating. In this context, the advanced BNs model is implemented to analyze the risk level in the process of braced excavation, and material parameters are considered to be updated with monitor information. In geotechnical engineering, it is often difficult to build and compute the likelihood functions of multivariable soil parameters. In light of this, the advanced BNs are improved with the distance-based ABC updating methods to reduce the computation cost for the complex computation. After computing and analyzing with the enhanced model, the results demonstrate distribution parameters of uncertain soil factors can be efficiently updating associated with the field data.

Sensitivity analysis is suitable for the structure of the BNs model, but the computation in the large network and/or Credal networks is complex and time-consuming. To overcome this shortcoming, global sensitivity analysis is implemented in the advanced BN model. The proposed method is described in Chapter IV. An infinite slope is investigated using the methods, the results show the importance measure can identify the uncertainty contribution of causal factors to the target parameters of interest in the network. The application of the imprecise stochastic model approach improves the BN model to cope with high dimension computation in the case of imperfect information, and besides, from the importance ranking of input variables in the network, key factors can be identified. With the enhanced sensitivity of advanced BN model, decision-makers can lower the risk level of the occurrence of the failure event trough the reduction of uncertainties in inputs.

# Bibliography



Acharya, G., F. De Smedt, and NT. Long. 2006. "Assessing landslide hazard in GIS: a case study from Rasuwa, Nepal." Bulletin of Engineering Geology and the Environment 65 (1):99–107.

Aguilera, PA., A. Fernndez, R. Fernndez, R. Rum, and A. Salmern. 2011. "Bayesian networksin environmental modelling." Environmental Modelling Software 26: 1376–1388.

Alessandro, A., and M. Zaffalon. 2008. "Decision-theoretic specification of credal networks: A unified language for uncertain modeling with sets of Bayesian networks." International Journal of Approximate Reasoning 49: 345–361.

Alonso, EE. 1976. "Risk analysis of slopes and its application to slopes in Canadian sensitive clays." Gotechnique 26: 453-472.

Alvarez, DA., F. Uribe, and JE. Hurtado. 2018. "Estimation of the lower and upper bounds on the probability of failure using subset simulation and random set theory." Mechanical Systems and Signal Processing 100: 782–801.

Angelis, M. De, E. Patelli, and M. Beer. 2014. "Advanced line sampling for efficient robust reliability analysis." Structural Safety 52: 170–182.

Anjum, MF, I. Tasadduq and K. AlSultan. 1997. "Response surface methodology: A neural network approach." European Journal of Operational Research 101: 65–73.

Antonucci, A., CP. de Campos, D. Huber, and M. Zaffalon. 2015. "Approximate Credal network updating by linear programming with applications to decision making." International Journal of Approximate Reasoning 58 (IS): 25–38.

Antonucci, A., A. Salvetti, and M. Zaffalon. 2004. "Assessing debris flow hazard by Credal nets." Soft Methodology and Random Information Systems 26: 125–132.

Beaumont, MA. 2010. "Approximate Bayesian computation in evolution and ecology." Annual Review of Ecology, Evolution, and Systematics 41: 79–406.

Beaumont, MA., WY. Zhang, and DJ. Balding. 2002. "Approximate Bayesian computation in population genetics." Genetics 162: 2025–2035.

Beck, JL. 2010. "Bayesian system identification based on probability logic." Structural Control and health Monitoring 17: 825–847.

Beck, JL., and SK. Au. 2002. "Bayesian updating of structural models and reliability using Markov Chain Monte Carlo simulation." Journal of Engineering Mechanics 128: 380–391.

Beck, JL., and LS. Katafygiotis. 1998. "Updating models and their uncertainties. I: Bayesian statistical framework." Journal of Engineering Mechanics 24: 455–461.

Bednarskia, M., W. Cholewaa, and W. Frid. 2004. "Identification of sensitivities in Bayesian networks." Engineering Applications of Artificial Intelligence 17: 327–335.

Beer, M., S. Ferson, and V. Kreinovich. 2013a. "Imprecise probabilities in engineering analyses." Mechanical Systems Signal Processing 37: 4–29.

Bi, SF., M. Broggi, and M. Beer. 2018. "The role of the Bhattacharyya distance in stochastic model updating." Mechanical Systems and Signal Processing 117: 437–452.

Bi, SF., M. Broggi, PF. Wei, M. Beer. 2019. "The Bhattacharyya distance: enriching the P-box in stochastic sensitivity analysis." Mechanical Systems and Signal Processing 129: 265-281

Borgonovo, E. 2007. "A new uncertainty importance measure." Reliability Engineering and System Safety 92 (6): 771–784.

Borgonovo, E., W. Castaings, and S. Tarantola. 2011. "Moment independent importance measures: new results and analytical test cases." Risk and Analysis 31: 404–428.

Botev, ZI. 2010. "Kernel density estimation using Matlab.".

Botev, ZI., JF. Grotowski, and DP. Kroese. 2010. "Kernel density estimation via diffusion." Annals of Statistics 38: 2916–2957.

Branco, L. Pinheiro, A. Topa Gomes, A. Silva Cardoso, and C. Santos Pereira. 2014. "Natural variability of shear strength in a granite residual soil from Porto." Geotechnical and Geological Engineering 32 (4): 911–922.

Calvello, M, and RJ Finno. 2004. "Selecting parameters to optimize in model calibration by inverse analysis." Computers and Geotechnics 31: 411–425.

Cao, ZJ., Y. Wang, and DQ. Li. 2016. "Site-specific characterization of soil properties using multiple measurements from different test procedures at different locations - A Bayesian sequential updating approach." Engineering Geology 211: 150-161.

Cardenas, IC., SSH. Al-Jibouri, JIM. Halman, W. Van de Linde, and F. Kaalberg. 2014. "Using prior risk-related knowledge to support risk management decisions: lessons learnt from a tunneling project." Risk Analysis 34: 1923–1943.

Cardenas, IC. 2019. "On the use of Bayesian networks as a meta-modelling approach to analyse uncertainties in slope stability analysis." Georisk 13: 53-65.

Cassidy, MJ., M. Uzielli, and S. Lacassec. 2008. "Probabilistic risk assessment of landslides: A case study at Finneidfjord." Canadian Geotechnical Journal 45: 1250-1267.

Chakraborty, A., and D. Goswami. 2017. "Slope stability prediction of slope stability using multiple linear regression (MLR) and artificial neural network." Arabian Journal of Geosciences 10 (17): 385–395.

Chan, H., and A. Darwiche. 2004. "Sensitivity analysis in Bayesian networks: from single to multiple parameters." Artificial Intelligence 25: 67–75.

Chen, G., ZH. Jia, and JC. Ke. 1997. "Probabilistic analysis of underground excavation stability." International Journal of Rock Mechanics and Mining Sciences 34: 51.e1–51.e16.

Chen, YC., TA. Wheeler, and MJ. Kochenderfer. 2017. "Learning discrete Bayesian networks from continuous data." Journal of Artificial Intelligence Research 59: 103–132.

Ching, JY., and YC. Chen. 2007. "Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging." Journal of Engineering Mechanics 133: 816–832.

Ching J., SS. Wu, and KK. Phoon. 2016. "Statistical characterization of random field parameters using frequentist and Bayesian approaches." Canadian Geotechnical Journal 53 (2): 285-298.

Ching, J., KK. Phoon, JL. Beck, and Y. Huang. 2017. "Identifiability of geotechnical site-specific trend functions." ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering 3 (4): 04017021.

Ching, J., and KK. Phoon. 2017. "Characterizing uncertain site-specific trend function by sparse Bayesian learning." Journal of Engineering Mechanics 143 (7): 04017028.

Ching, J., and KK. Phoon. 2019a. "Constructing site-specific multivariate probability distribution model using Bayesian machine learning." Journal of Engineering Mechanics 145 (1): 04018126.

Ching, J., and KK. Phoon. 2019b. "Measuring similarity between site-specific data and records from other sites." ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, in press.

Ching, J., and KK. Phoon. 2019c. "Constructing a site-specific multivariate probability distribution using sparse, incomplete, and spatially variable (MUSIC-X) data." Journal of Engineering Mechanics, under review.

Choobbasti, AJ., F. Farrokhzad, and A. Barari. 2009. "Prediction of slope stability using artificial neural network (case study: Noabad, Mazandaran, Iran)." Arabian Journal of Geo-sciences 2: 311–319.

Cho, SE. 2013. "First-order reliability analysis of slope considering multiple failure modes." Engineering Geology 154: 98-105.

Christian, JT., CC. Ladd, and GB. Baecher. 1994. “Reliability applied to slope stability analysis.” Journal of Geotechnical Engineering 120: 2180–2207.

Chua, CG., and ATC. Goh. 2005. “Estimating wall deflections in deep excavations using Bayesian neural networks.” Tunnelling and Underground Space Technology 20: 400-409.

Chun, MH., SJ. Han, and NI. Tak. 2000. “An uncertainty importance measure using a distance metric for the change in a cumulative distribution function.” Reliability Engineering and System Safety 70: 313–321.

Cobba, BR., and PP. Shenoyb. 2006. “Inference in hybrid Bayesian networks with mixtures of truncated exponentials.” International Journal of Approximate Reasoning 41: 257–286.

Coup, VMH., and LC. Van der Gaag. 2002. “Properties of sensitivity analysis of Bayesian belief.” Annals of Mathematics and Artificial Intelligence 36 (4): 323–356.

Cozman, FG. 2000. “Credal Networks.” Artificial Intelligence 120: 199–233.

Cozman, FG. 2005. “Graphical models for imprecise probabilities.”International Journal of Approximate Reasoning 39: 167–184.

Cukier, RI., CM. Fortuin, and KE. Shuler. 1973. “Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory.” The Journal of Chemical Physics 59(8): 771–784.

Da Fonseca, A. Viana, and R. Coutinho. 2008. “Characterization of residual soils.” In Proceedings of the 3rd International Conference on Site Characterization, 195–248.

Daly, R., Q. Shen, and QS. Aitken. 2011. “Learning Bayesian networks: approaches and issues.” The Knowledge Engineering Review 26 (2): 99–157.

Dai, FC., CF. Lee, and YY. Ngai. 2002. “Landslide risk assessment and management: an overview.” Engineering Geology 64: 65-87.

Dawson, EM., WH. Roth, and A. Drescher. 1999. “Slope stability analysis by strength reduc-tion.” Geotechnique 49: 835–840.

De Angelis, M., E. Patelli, and M. Beer. 2014. “Advanced Line Sampling for efficient robust reliability analysis.” Structural Safety 52 (SI): 170–182.

Dilley, M., RS. Chen, U. Deichmann, AL. Lerner-Lam, M. Arnold, J. Agwe, P. Buys, O. Kjekstad, B. Lyon, and G. Yetman. 2005. “Natural disaster hotspots: a global risk analysis.” The World Bank Hazard Management Unit, Washington.

Do, TN., CY. Oua, and RP. Chen. 2016. “A study of failure mechanisms of deep excavations in soft clay using the finite element method.” Computers and Geotechnics 73: 153-163.

Dougherty, J., R. Kohavi, and M. Sahami. 1995. “Supervised and unsupervised discretization of continuous features.” In: Proceedings of the 12th International Conference on Machine Learning (ICML).

Dyson, AP., and A. Tolooiyan. 2019. “Prediction and classification for finite element slope stability analysis by random field comparison.” Computers and Geotechnics 109: 117–129.

Dzgn, HSB., and S. Lacasse. 2005. “Vulnerability and acceptable risk in integrated risk assessment framework.” International Conference on Landslide Risk Management, Vancouver, Canada 505-516.

Edoardo, P., M. Broggi, S. Tolo, and J. Sadeghi. 2017. “COSSAN software a multidisciplinary and collaborative software for uncertainty quantification.” 212–224.

El-Ramly, H., NR. Morgenstern, and DM. Cruden. 2002. “Probabilistic slope stability analysis for practice.” Canadian Geotechnical Journal 39 (3): 665–683.

EL-Ramly, H., NR. Morgenstern, and DM. Cruden. 2005. “Probabilistic assessment of stability of a cut slope in residual soil.” Geotechnique 55: 77-84.

Elbaz, K., JS. Shen, A. Arulrajah, and S. Horpibulsuk. 2016. “Geohazards induced by anthropic activities of geoconstruction: a review of recent failure cases.” Arabian Journal of Geosciences 9: 708.

Elkateb, T., R. Chalaturnyk, and PK. Robertson. 2003. “An overview of soil heterogeneity: quantification and implications on geotechnical field problems.” Canadian Geotechnical Journal 40: 1–15.

Fagiuoli, E., and M. Zaffalon. 1988. “2U: An exact interval propagation algorithm for poly-trees with binary variables.” Artificial Intelligence 106 (1): 77–107.

Ferreira da Rocha, JC., and FG. Cozman. 2002. “Inference with separately specified sets of probabilities in Credal networks.” In Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence, San Francisco, 430–437.

Fetz, T., and M. Oberguggenberger. 2016. “Imprecise random variables, random sets, and Monte Carlo simulation.” International Journal of Approximate Reasoning 78: 252–264.

Finlay, PJ., GR. Mostyn, and R. Fell. 1999. “Landslide risk assessment: prediction of travel distance.” Canadian Geotechnical Journal 36: 556-562.

Finno, RJ., and M. Calvello. 2005. “Supported excavations: observational method and inverse modeling.” Journal of Geotechnical and Geoenvironmental Engineering 131: 826–836.

Finno, RJ., J. Richard, JT. Blackburn, and JF. Roboski. 2007. "Three-dimensional effects for supported excavations in clay." Journal Geotechnical and Geoenvironmental Engineering 133:30–36.

Goh, ATC., KS. Wong, and BB. Broms. 1995. "Estimation of lateral wall movements in braced excavations using neural networks." Canadian Geotechnical Journal 32: 1059–1064.

Goldsworthy, JS., MB. Jaksa, GA., Fenton, WS. Kaggwa, V. Griffiths and HG., Poulos. 2007. "Effect of sample location on the reliability based design of pad foundations." Georisk 1 (3): 155-166.

Gomes, A. Topa. 2009. "Elliptical shafts by the sequential excavation method. The case of Metro do Porto, (in Portuguese)." PhD Thesis, University of Porto 195–248.

Gong, WP., Z. Luo, CH. Juang, HW. Huang, J. Zhang, and L. Wang. 2014. "Optimization of site exploration program for improved prediction of tunneling-induced ground settlement in clays." Computer and Geotechnics 56: 69-79.

Griffiths, DV., and PA. Lane. 1999. "Slope stability analysis by finite elements." Geotechnique 49: 387–403.

Griffths, DV., and GA. Fenton. 2004. "Probabilistic slope stability analysis by

finite elements." Journal of Geotechnical and Geoenvironmental Engineering 130: 507-518.

Griffths, D.V., JS. Huang, and GA. Fenton. 2011. "Probabilistic in

nite slope analysis." Computers and Geotechnics 38: 577-584.

Haque, U., P. Blum, and PF. Da Silva. 2016. "Fatal landslides in Europe." Landslides 13:1545–1554.

Hashash, YMA., S. Jung, and J. Ghaboussi. 2004. "Numerical implementation of a neuralnetwork based material model in finite element analysis." International Journal for Numerical Methods in Engineering 59: 989–1005.

Hashash, YMA., S. Levasseur, A. Osouli, R. Finno, and Y. Malecot. 2010. "Comparison of two inverse analysis techniques for learning deep excavation response." Computers and Geotechnics 37: 323–333.

Helton, JC. 2000. "Sampling-based methods." Sensitivity Analysis 101–53.

Hight, DW. and VN. Georgiannou. 1995. "Effects of sampling on the undrained behaviour of clayey sands." Geotechnique 45(2): 237-247.

Hsieh, PG., CY. Ou, and HT. Liu. 2008. "Basal heave analysis of excavations with consideration of anisotropic undrained strength of clay." Canadian Geotechnical Journal 45: 788–799.

Huang, J., AV. Lyamin, DV. Griffths, K. Krabbenhoft, and SW. Sloan. 2013. “Quantitative risk assessment of landslide by limit analysis and random

elds.” Computers and Geotechnics 53: 60-67.

Iman, R., and S. Hora. 1990. “A robust measure of uncertainty importance for use in fault tree system analysis.” Risk Analysis 10: 401-403.

ISO-2394. 2015. “General principles on reliability for structures.” ISO/TC 98/SC 2 Reliability of structures 111.

Jan, JC., SL. Hung, SY. Chi, and JC. Chen. 2002. “Neural network forecast model in deep excavation.” Journal of Computing in Civil Engineering 16: 59–65.

Janin, JP, D. Dias, F. Emeriault, R. Kastner, H. Le Bissonnais, and A. Guilloux. 2015. “Numerical back-analysis of the southern Toulon tunnel measurements: A comparison of 3D and 2D approaches.” Engineering Geology 195: 42–52.

Jensen, F. 2001. “Bayesian networks and decision graphs.” Springer-Verlag, Heidelberg Press.

Jiang, SH., DQ. Li, LM. Zhang, and CB. Zhou. 2014. “Slope reliability analysis considering spatially variable shear strength parameters using a non-intrusive stochastic finite element method.” Engineering Geology 168: 120–128.

Jiang, SH., I. Papaioannou and D. Straub. 2017. “Optimizing borehole locations for slope reliability assessment.” Georisk 284: 420-430.

Juang, CH., Z. Luo, S. Atamturktur, and HW. Huang. 2013. “Bayesian updating of soil parameters for braced excavations using field observations.” Journal of Geotechnical and Geoenvironmental Engineering 139: 395–406.

Korb, KB., and AE. Nicholson. 2010. “Bayesian artificial intelligence.” Faculty of Information Technology, Monash University, Clayton, Victoria 3800 Australia 2nd, CRC Press.

Koutsourelakis, PS. 2004. “Reliability of structures in high dimensions. Part ii. theoretical validation.” Probability Engineering Mechnique 19 (4): 19–23.

Kristo, C., H. Rahardjo, and A. Satyanaga. 2017. “Effect of Variations in Rainfall Intensity on Slope Stability in Singapore.” International Soil and Water Conservation Research4 (1):258–264.

Kung, GTC., CH. Juang, ECL. Hsiao, and YMA. Hashash. 2007. “Simplified model for wall deflection and ground-surface settlement caused by braced excavation in clays.” Journal of Geotechnical and Geoenvironmental Engineering 133: 731–747.

Kurgan, L., and K.J. Cios. 2001. “Discretization algorithm that uses class-attribute interdependence maximization.” In Proceedings of the International Conference on Artificial Intelligence (IC-AI), Las Vegas, Nevada, 980–987.

Lacasse, S., and F. Nadim. 1998. "Risk and reliability in geotechnical engineering." Proceedings: Fourth International Conference on Case Histories in Geotechnical Engineering, St. Louis, Missouri March 9-12.

Lacasse, S., ZQ. Liu, and F. Nadim. 2017. "Probabilistic characterization of soil properties recognition of Wilson Tang's contribution to geotechnical practice." Geotechnical Safety and Reliability GSP 286 40: 2–26.

Langseth, H., T. D. Nielsen, R. Rumi, and A. Salmern. 2009. "Inference in hybrid Bayesian networks."Reliability Engineering System Safety 94: 1499–1509.

Laskey, K.B. 1995. "Sensitivity analysis for probability assessments in Bayesian networks."IEEE Transactions on System 25: 901–909.

Lauritzen, SL. 1992. "Propagation of probabilities, means and variances in mixed graphical association models." Journal of the American Statistical Association 87: 1098–1108.

Ledesma, A., and A. Gens; EE. Alonso. 1996. "Three-dimensional finite element analysis ofdeep excavations."International Journal for Numerical and Analytical Methods in Geomechanics 20: 119–141.

Lee, FH., SH. Hong, Q. Gu, and PJ. Zhao. 2011. "Application of large three-dimensional finite element analyses to practical problems." International Journal of Geomechanics 11: 529–539.

Li, C., and S. Mahadevan. 2017. "Sensitivity analysis of a Bayesian network." ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering 4: 1–10.

Li, DQ., ZY. Yang, ZJ. Cao, SK. Au, and KK. Phoon. 2017. "System reliability analysis of slope stability using generalized Subset Simulation." Applied Mathematical Modelling 46: 650–664.

Li, X., and L. Zhang. 2018. "Efficient Bayesian networks for slope safety evaluation with large quantity monitoring information." Geoscience Frontiers 9 (6): 1679–1687.

Liu, WS., SB. Li, and R. Tang. 2013. "Slope Stability Evaluation in Open Pit Based onBayesian Networks." Proceedings of the ICCEAE 2012 181: 1227–1231.

Li, S., ZQ. Yue, LG. Tham, CF. Lee, and SW. Yan. 2005. "Slope failure in under consolidated soft soils during the development of a port in Tianjin, China. Part 2: Analytical study." Canadian Geotechnical Journal 42: 166-183.

Liang, SH., JS. Huang, C. Yao, and JH. Yang. 2019. "Quantitative risk assessment of slope failure in 2-D spatially variable soils by limit equilibrium method." Applied Mathematical Modelling 47: 710-725.

Li, YJ., C. Qian and K. Liu. 2019. "Sampling efficiency in spatially varying soils for slope stability assessment." Advances in Civil Engineering.

Liu, Y., YN. Ding, and G. Chen. 2019. "Quantitative risk assessment on landslides considering spatial variability of soils." Proceedings of the 29th European Safety and Reliability Conference Edited by Michael Beer and Enrico Zio: doi:10.3850/978-981-11-2724-3 1065-cd.

Liu, Y., FH. Lee, ST. Quek, JE. Chen, and JT. Yi. 2015. "Effect of spatial variation of strength and modulus on the lateral compression response of cement-admixed clay slab." Geotechnique 65: 851–865.

Liu, Y., W. Zhang, L. Zhang, Z. Zhu, J. Hu, and H. Wei. 2018. "Probabilistic stability analyses of undrained slopes by 3D random fields and finite element methods." Geoscience Frontiers 9 (6): 1657-1664.

Ma, J., H. Tang, X. Hu, A. Bobet, M. Zhang, T. Zhu, Y. Song, and A. M. Mutasim. 2017."Identification of causal factors for the Majiagou landslide using modern data mining methods." Landslides 14 (1): 311–322.

MacKay, DJC. 1992. "Bayesian interpolation."Neural Computation 4: 415–447.

Maua, DD., CP. de Campos, A. Benavoli, and A. Antonucci. 2014. "Probabilistic inference in Credal networks: new complexity results." Journal of Artificial Intelligence Research 50:603–637.

Metya, S., T. Mukhopadhyay, S. Adhikari, and G. Bhattacharya. 2009. "Efficient evaluation of reliability for slopes with circular slip surfaces using importance sampling." Geotechnical and Geoenvironmental Engineering 135 (6): 768–777.

Metya, S., T. Mukhopadhyay, S. Adhikari, and G. Bhattacharya. 2017. "System reliability analysis of soil slopes with general slip surfaces using multivariate adaptive regression splines." Computers and Geotechnics 87 (7): 212–228.

Mia, M., and NR. Dhar. 2016. "Response surface and neural network based predictive models of cutting temperature in hard turning." Journal of Advanced Research 7: 1035–1044.

Miro, S., M. Knig, Hartmann, and T. Schanz. 2015. "A probabilistic analysis of subsoil parameters uncertainty impacts on tunnel-induced ground movements with a back-analysis study." Computers and Geotechnics 68: 38–53.

Mitra, D., C. Bhandery, A. Mukhopadhyay, A. Chanda, and S. Hazra. 2018. "Landslide risk assessment in Darjeeling hills using multi-criteria decision support system: a Bayesian network approach." Disaster Risk Governance in India and Cross Cutting 361-386.

Moral, S., R. Rum, and A. Salmern. 2001. "Mixtures of truncated exponentials in hybrid Bayesian networks." European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty 156–167.

Morales-Napoles, O., DJ. Delgado-Hernandez, D. De-Leon-Escobedo, and JC. Arteaga-Arcos 2014. "A continuous Bayesian network for earth dams' risk assessment: methodology and quantification." Structure and Infrastructure Engineering 10: 589-603.

Morgenstern, H., NR. Morgenstern, and DM. Cruden. 2002. "Probabilistic slope stability anal-ysis for practice." Canadian Geotechnical Journal 39: 665–683.

Morgenstern, NR., and V. Price. 1965. "The analysis of the stability of general slip surface." Geotechnique 15 (1): 79–93.

Niu, WJ. 2014. "Determination of slope safety factor with analytical solution and searching critical slip surface with genetic-traversal random method." Scientific World Journal.

Oka, Y., and TH. Wu. 1990. "System reliability of slope stability." Geotechnical Engineering 116 (8): 1185–1189.

Oliveira, PJV., JLP. Pinheiro, and AAS. Correia. 2011. "Numerical analysis of an embankment built on soft soil reinforced with deep mixing columns: Parametric study." Computers and Geotechnics 38: 566-576.

Orense, R., K. Farooq, and I. Towhata. 2014. "Deformation behavior of sandy slopes during rain water infiltration." Soil and Foundation 44: 15–30.

Ou, C.Y., DC. Chiou, and TS. Wu. 1996. "Three-dimensional finite element analysis of deepexcavations." Journal of Geotechnical Engineering-ASCE 122: 337–345.

Panek, T., K. Silhan, P. Taborik, and J. Hradecky. 2011. "Catastrophic slope failure and its origins: Case of the May 2010 Girov Mountain long-runout rockslide (Czech Republic)." Geomorphology 130: 352–364.

Papadimitriou, C., JL. Beck, and LS. Katafygiotis. 2001. "Updating robust reliability usingstructural test data." Probabilistic Engineering Mechanics 16: 103–113.

Park, CK., and KI. Ahn. 1994. "A new approach for measuring uncertainty importance and distributional sensitivity in probabilistic safety assessment." Reliability Engineering System Safety 46: 253–261.

Park, JM., YS. Jo, and YS. Jang. 2019. "Reliability Assessment of Geotechnical Structures on Soils improved by Deep Mixing Method I: Data Collection and Problem Setting." Journal of Civil Engineering 23 (1): 63-73.

Patelli, E. 2016. "A Multidisciplinary Software Suite for Uncertainty Quantification and Risk Management." PIn: Ghanem R., Higdon D., Owhadi H. (eds.) Handbook of Uncertainty Quantification, Springer, Cham, Switzerland.

Patelli, E., M. Broggi, S. Tolo, and J. Sadeghi. 2017. "COSSAN software: A multidisciplinary and collaborative software for uncertainty quantification." Proceedings of

2nd ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering, Rhodes, Greece212224.

Pearl, J. 1988. Probabilistic reasoning in intelligent systems. San Mateo, California: Morgan Kaufmann Publishers.

Pearl, J., and S. Russell. 2000. Causality: models, reasoning, and inference. Cambridge University Press.

Peng, M., XY. Li, and D. Q. Li. 2013. “Slope safety evaluation by integrating Multi-source monitoring information.” Structural Safety 49 (SI): 65–74.

Phoon, KK. 2017. “Role of reliability calculations in geotechnical design.” Georisk 11 (1): 4-21.

Phoon, KK. 2008. “Reliability-based design in geotechnical engineering: computations and applications.” Taylor Francis, London Chapter 1.

Phoon, KK. 2018. “Editorial for Special Collection on Probabilistic Site Characterization.” ASCEASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering 4 (4): 02018002.

Phoon, KK. 2019. “Managing uncertain ground truth using Bayesian machine learning.” Proceedings of the 29th European Safety and Reliability Conference Edited by Michael Beer and Enrico Zio: doi:10.3850/978-981-11-2724-3 0002-cd.

Phoon, KK. 2020. The story of statistics in geotechnical engineering. Georisk, under review.

Phoon, KK., and F. H. Kulhawy. 1999. “Characterization of geotechnical variability.” Canadian Geotechnical Journal 35: 612–624.

Prangle, D. 2017. “Approximate Bayesian computation in evolution and ecology.” International Society for Bayesian Analysis 12: 289-309.

Rahardjo, H., SN. Alfrendo, C. L. Eng, and Y. S. Ng. 2010. “Effects of Groundwater Table Position and Soil Properties on Stability of Slope During Rainfall.” Geotechnical andGeoenvironmental Engineering 136 (10): 1555–1564.

Remondo, J., J. Bonachea, and A. Cendrero. 2008. “Quantitative landslide risk assessment and mapping on the basis of recent occurrences." Geomorphology 94: 496-507.

Rocscience. 2011. “RS2 Software.” In reocscience, version 7.0 [computer program].

Sakellariou, MG., and MD. Ferentinoue. 2005. “A study of slope stability prediction using neural networks.” Geotechnical and Geological Engineering 23: 419–445.

Saltelli, A. 2002. “Sensitivity analysis for importance assessment.” Risk analysis 22: 579–590.

Saltelli, A., P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola. 2010. "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index." Computer Physics Communications 181: 259–270.

Sankararaman, S., and S. Mahadevan. 2011. "Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data." Reliability Engineering and System Safety 96 (7): 814–824.

Schubert, M., NP. Hoj, A. Ragnoy, and H. Buvik. 2012. "Risk assessment of road tunnels using Bayesian networks." Transport Research Arena 48: 2697-2706.

Schwarza, M., F. Pretic, F. Giadrossich, P. Lehmann, and D. Or. 2010. "Quantifying the role of vegetation in slope stability: A case study in Tuscany (Italy)." Ecological Engineering36:285–291.

Schweizer, B., and EF. Wolff. 1981. "On nonparametric measures of dependence for random variables." Annals of Statistics 9: 879–885.

Scott, F., V. Kreinovich, L. Ginzburg, DS. Myers, and K. Sentz. 2003. "Constructing Probability Boxes and Dempster-Shafer Structures." Sandia National Laboratories SAND 2002–4015.

Shafer, G. 1976. "A mathematical theory of evidence." Reliability Engineering and System Safety.

Shenoy, PP., and JC. West. 2011. "Inference in hybrid Bayesian networks using mixtures of polynomials." International Journal of Approximate Reasoning 52: 641–657.

Silva, F., TW. Lambe, and WA. Marr. 2008. "Probability and risk of slope failure." Journal of Geotechnical and Geoenvironmental Engineering 134: 1691-1699.

Simoen, E., G. De Roeck, and G. Lombaert. 2015. "Dealing with uncertainty in model updating for damage assessment: A review." Mechanical Systems and Signal Processing 56-57: 123–149.

Smith, M. 2006. "Dam Risk Analysis Using Bayesian Networks." Proceedings of Geohazards .

Sobol, IM. 1993. "Sensitivity estimates for nonlinear Mathematical models." Mathematical Modelling and Computational Experiment 1: 407–414.

Sobol, I.M., S. Tarantolab, D. Gatellib, SS. Kucherenkoc, and W. Mauntz. 2007. "Estimating the approximation error when fixing unessential factors in global sensitivity analysis." Reliability Engineering and System Safety 92: 957–960.

Song, Y., J. Gong, S. Gao, D. Wang, T. Cui, Y. Li, and B. Wei. 2012. "Susceptibility assessment of earthquake-induced landslides using Bayesian network: a case study in Beichuan, China." Computer Geoscience 42: 189–199.

Song, Y. Q., J. H. Gong, and S. Gao. 2012. “Susceptibility assessment of earthquake-induced landslides using Bayesian network: a case study in Beichuan, China.” Computers and Geosciences 42: 189–199.

Sousa, RL. “Risk analysis for tunneling projects.” PhD Thesis, Massachusetts Institute of Technology.

Sousa, RL., and HH. Einstein. 2012a. “Risk analysis during tunnel construction using Bayesian Networks: Porto Metro case study.” Tunnelling and Underground Space Technology 27: 86-100.

Sousa, RL., and HH. Einstein. 2015. “A dynamic Bayesian network based approach to safety decision support in tunnel construction.” Reliability Engineering and System Safety 134: 157-168.

Spackova, O., J. Sejnoha, and D. Straub. 2013. “Probabilistic assessment of tunnel construction performance based on data.” Tunnelling and Underground Space Technology 37: 62–78.

Sperotto, A., JL. Molina, S. Torresan, A. Critto, and A. Marcomini. 2017. “Reviewing Bayesian networks potentials for climate change impacts assessment and management: A multi-risk perspective.” Journal of Environmental Management 202: 320-331.

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

Straub, D., and A. Der Kiureghian. 2010. “Combining Bayesian networks with Structural Reliability Methods: Methodology.” Journal of Engineering Mechanics 136 (10): 1248–1258.

Straub, D., and A. Der Kiureghian. 2010. “Bayesian network enhanced with structural reliability methods: methodology.” Journal of Engineering Mechanics 136 (10): 1248-1258.

Straub, D., and I. Papaioannou. 2015. “Bayesian updating with structural reliability methods.” Journal of Engineering Mechanics 141.

Terzaghi, K. 1943. Theoretical soil mechanics. New York: John Wiley and Sons.

Tessem, B. 1992. “Interval probability propagation.” International Journal of Approximate Reasoning 7: 95–120.

Tolo, S., E. Patelli, and M. Beer. 2018a. “An inference method for Bayesian networks with probability intervals.” In Proceedings of the joint ICVRAM-ISUMA UNCERTAINTIES Conference, Florianpolis, SC, Brazil.

Tolo, S., E. Patelli, and M. Beer. 2018b. “An open toolbox for the reduction, inference computation and sensitivity analysis of Credal networks.” Advances in Engineering Software 115: 126–148.

Toni, T., D. Welch, N. Strelkowa, A. Ipsen, and M. Stumpf. 2009. “Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems.” Journal of The Royal Society Interface 6: 187–202.

Tsai, F., JS. Lai, and TH. Lin. 2013. “Analysis of topographic and vegetative factors with data mining for landslide verification.” Ecological Engineering 61: 669–677.

Turner, BM, and T. Van Zandt. 2012. “A tutorial on approximate Bayesian computation.” Journal of Mathematical Psychology 56: 69–85.

Vanmarcke, EH. 1977. “Probabilistic modeling of soil profiles." Journal of the Geotechnical Engineering Division, ASCE 103: 1227-1246.

Vucetic, M., and R. Dobry. 1991. “Effect of soil plasticity on cyclic Response.” Journal of Geotechnical Engineering 117: 89–107.

Wang, Y., ZJ. Cao, and SK. Au. 2011. “Practical Reliability Analysis of Slope Stability by Advanced Monte Carlo Simulations in a Spreadsheet.” Geotechnical and Geoenvironmental Engineering 48 (1): 162–172.

Wang, Y., Z. Guan, TY. Zhao. 2019. “Sample size determination in geotechnical site investigation considering spatial variation and correlation.” Canadian Geotechnical Journal 56 (7): 992-1002.

Weber, P., G. Medina-Oliva, C. Simon, and B. Iung. 2012. “Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas.” Engineering Appli-cations of Artificial Intelligence 25 (4): 671–682.

Wei, PF., JW. Song, SF. Bi, M. Broggi, and M. Beer. 2019. “Non-intrusive stochastic analysis with parameterized imprecise probability models: I. Performance estimation.”Mechanical Systems and Signal Processing 124: 1728–1743.

Wei, PF., ZZ. Zheng, W. Hao, J. Feng, and B. Wang. 2012. “Efficient sampling methods for global reliability sensitivity analysis.” Computer Physics Communications183 (8): 1728–1743.

Wei, PF., ZZ. Zheng, and JW. Song. 2014. “Extended Monte Carlo simulation for parametric global sensitivity analysis and optimization.” AIAA Journal 52: 867–878.

Wei, PF., ZZ. Zheng, and XK. Yuan. 2013. “Monte Carlo simulation for moment-independent sensitivity analysis.” Reliability Engineering and System Safety 110: 60–67.

Weichselberger, K. 2000. “International journal of approximate reasoning.” Risk and Analysis 24: 149–170.

Whitman, Robert V. 1984. “Evaluating calculated risk in Geotechnical Engineering." Journal of Geotechnical Engineering 110: 143-188.

Whitman, Robert V. 2000. “Organizing and Evaluating Uncertainty in Geotechnical Engineering." Journal of Geotechnical and Geoenvironmental Engineering 126: 583-593.

Wu, FC., and CC. Chen. 2009. “Bayesian updating of parameters for a sediment entrainment model via Markov Chain Monte Carlo.” Journal of Hydraulic Engineering 135: 22–37.

Xia, YY., M. Mahmoodian, CQ. Li, and A. Zhou. 2017. “Stochastic method for predicting risk of slope failure subjected to unsaturated infiltration flow." International Journal of Geomechanics 17 (04017037).

Xu, Y., LM. Zhang, and JS. Jia. 2011. “Diagnosis of embankment dam distresses using Bayesian networks. Part II. Diagnosis of a specific distressed dam." Canadian Geotechnical Journal 48: 1645-1657.

Yang, R., JS. Huang, DV. Griffiths, JH. Li and DC. Sheng. 2019. “Importance of soil property sampling location in slope stability assessment.” Canadian Geotechnical Journal 56 (3): 335-346.

Yuan, R., and G. Bai. 2011. “New Neural network response surface methods for reliability analysis.” Chinese Journal of Aeronautics 24: 25–31.

Zdravkovic, L, DM. Potts, and HD. St John. 2005. “Modelling of a 3D excavation in finite element analysis.” Geotechnique 55: 497–513.

Zhang, LM, XG. Wu, LY. Ding, MJ. Skibniewski, and .Y Yan. 2013. “Decision support analysis for safety control in complex project environments based on Bayesian Networks.” EXPERTSYSTEMS WITH APPLICATIONS 40: 4273–4282.

Zhang, LM., XG. Wu, YW. Qin, MJ. Skibniewski, and WL. Liu. 2016. “Towards a fuzzy Bayesian network based approach for safety risk analysis of tunnel-induced pipeline dam-age.” Risk Analysis 36: 278–301.

Zhang, J., and HW. Huang. 2016. “Risk assessment of slope failure considering multiple slip surfaces." Computers and Geotechnics 74: 188-195.

Zhang, LM., Y. Xu, JS. Jia, and C. Zhao. 2011. “Diagnosis of embankment dam distresses using Bayesian networks. Part I. Global-level characteristics based on a dam distress database.” Canadian Geotechnical Journal 48: 1630-1644.

Zhou, A., JS. Huang, and CQ. Li. 2016. “Failure analysis of an infinite unsaturated soil slope.” Geotechnical Engineering 169: 410-420.

Zhou, Y., C. Li, C. Zhou, and H. Luo. 2018. "Using Bayesian network for safety risk analysis of diaphragm wall deflection based on field data." Reliability Engineering System Safety 180: 152–167.

# Acknowledgements

First and foremost I would like to thank my supervisor Prof. Dr. Michael Beer for giving me the opportunity of studying in Germany. I am also very appreciate it that the funding from China Scholarship Council supported me to live and study in Germany for four years. I specially thank to the help from international office in Leibniz Universität Hannover, supporting me finish the last half year of PhD study.

I sincerely thank Prof. Dr. António Topa Gomes, Prof. Dr. Kok Kwang Phoon, Prof. Dr. Yong Liu and associate Prof. Dr. Li Wang, providing me with guidance and advice about the geotechnical knowledge. Without them, the thesis would not have been produced. Additionally, Prof. Dr. Kok Kwang Phoon and Prof. Dr. António Topa Gomes are particularly gratefully acknowledged for reviewing parts of my thesis from a geotechnical professional point of views, and gave me some precious suggestions for improving my thesis.

Also, i am particularly grateful for the tolerance and moral support of my parents. They always have my back and respect all my decisions. I am the only child of them and I know it is not easy for them to meet me only once a year, so I very thank their understanding and support.

Last but not least, a big thank goes to the help I ever got from my friends, colleagues and some warm-hearted strangers I met, making my phD life in Germany warm and colorful. The one of a kind experience living in Germany I wish I could record and treasure it in my mind and heart forever.

# Curriculum Vitae

## Personal Details

Surname: He

First Name: Longxue

Date of birth: 09 November 1988

Place of birth: Jilin, China

## University Education

Since 10/2015: Studies in Institute for risk and reliability, Faculty of Civil Engineering and Geodetic Science, Leibniz Universität Hannover, Germany.

09/2012 - 06/2015: Studies in Environmental Engineering, University of Chinese Academy of Sciences(UCAS), Beijing, China. (Institute: Northeast Institute of Geography and Agroecology, Changchun, China)

09/2008 - 06/2012: Studies in Environmental Engineering, Nanchang University, China.

## Publications

**Longxue He, António Topa Gomes, Matteo Broggi, and Michael Beer. 2019:** Failure Analysis of Soil Slopes with Advanced Bayesian Networks. Periodica Polytechnica Civil Engineering. https://doi.org/10.3311/PPci.14092

**Longxue He, Yong Liu, Sifeng Bi, Li Wang, Matteo Broggi, Michael Beer. 2019:** Estimation of failure probability in a braced excavation by Bayesian networks integrating with model updating approaches. Underground Space.
https://doi.org/10.1016/j.undsp.2019.07.001

**Longxue He, Michael Beer, Matteo Broggi, Pengfei Wei and António Topa Gomes. 2019:** Sensitivity analysis of prior beliefs in advanced Bayesian networks. 2019 IEEE Symposium Series on Computational Intelligence, December 6-9, Xiamen, China (published in the proceeding of the 2019 IEEE Symposium Series on Computational Intelligence, December 6-9, Xiamen, China, and will be published in IEEE Xplore)

## Conference Contributions

**Longxue He, António Topa Gomes, Matteo Broggi, and Michael Beer. 2018:** Risk Assessment for Slope Stability with Enhanced Bayesian Networks Methods. Proceeding of the 6th ISRERM conference, Singapore.

**Longxue He, António Topa Gomes, Matteo Broggi, and Michael Beer. 2018:** Risk Analysis of Infinite Slope Failure using Advanced Bayesian Networks. Proceeding of the 8th International Workshop on REC2018, Liverpool, UK.

9