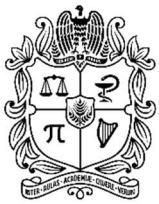


UNIVERSIDAD
NACIONAL
DE COLOMBIA

Bayesian Network Methodology for Decision Support in Forensic Geotechnical Engineering

William Mauricio García Feria

Universidad Nacional de Colombia

Facultad de Ingeniería, Área Curricular de Ingeniería Civil y Agrícola

Bogotá D.C., Colombia

2023

Bayesian Network Methodology for Decision Support in Forensic Geotechnical Engineering

William Mauricio García Feria

Tesis presentada como requisito parcial para optar al título de:

Doctor en Ingeniería Civil

Director:

Julio Esteban Colmenares Montañez. Ph.D

Profesor Titular

Línea de Investigación:

Geotecnia y Riesgo Geoambientales

Grupo de Investigación:

GENKI

(Geotechnical Engineering Knowledge and Innovation)

Universidad Nacional de Colombia

Facultad de Ingeniería, Área Curricular de Ingeniería Civil y Agrícola

Bogotá D.C., Colombia

2023

A Judy y Julieta

Declaración de obra original

Yo declaro lo siguiente:

He leído el Acuerdo 035 de 2003 del Consejo Académico de la Universidad Nacional. «Reglamento sobre propiedad intelectual» y la Normatividad Nacional relacionada al respeto de los derechos de autor. Esta disertación representa mi trabajo original, excepto donde he reconocido las ideas, las palabras, o materiales de otros autores.

Cuando se han presentado ideas o palabras de otros autores en esta disertación, he realizado su respectivo reconocimiento aplicando correctamente los esquemas de citas y referencias bibliográficas en el estilo requerido.

He obtenido el permiso del autor o editor para incluir cualquier material con derechos de autor (por ejemplo, tablas, figuras, instrumentos de encuesta o grandes porciones de texto).

Por último, he sometido esta disertación a la herramienta de integridad académica, definida por la universidad.

_____

William Mauricio García Feria

Fecha 31/01/2023

Agradecimientos

Este trabajo doctoral no hubiera sido posible sin el apoyo de muchas personas que estuvieron presentes a lo largo estos años. A continuación, mis más sinceros agradecimientos a :

El profesor Julio Esteban Colmenares por la confianza depositada, sus palabras de ánimo y su acompañamiento en todas las fases de investigación. Su entusiasmo por la academia, el amor por la Universidad Nacional y su vocación por la investigación es un modelo para esta generación y las futuras.

El profesor Laureano Hoyos por aceptarme como investigador visitante en University of Texas at Arlington. Desde mi llegada a los Estados Unidos me sentí acogido por su generosidad y su pasión por la geotecnia. Su familia fue un gran apoyo en los momentos necesarios.

Mi familia, que estuvo siempre a mi lado y soportó la alegría, la tristeza, la soledad, y el estrés que implica el desarrollo de un doctorado.

A mi gran amigo y hermano de vida Diego Munar, que en paz descanse.

Abstract

Bayesian Network Methodology for Decision Support in Forensic Geotechnical Engineering

Recent advances in engineering have increased the community's expectation for civil engineering works to operate safely. Occasionally some of these works fail because of human errors or the unpredictable behavior of materials. Forensic engineering is the branch of forensic science in charge of investigating those engineering failures. Scientific methods used in forensic engineering guarantee that conclusions regarding the causes of an engineering failure come from reliable investigation processes. However, in the case of geotechnical failures, the inherent uncertainty of soil/rock materials, difficulties in evidence collection, and multiplicity of failure scenarios (hypotheses) pose a challenge in identifying the actual causes of failure. Therefore, conclusions about the causes of geotechnical failures sometimes seem arbitrary and biased because they are mainly based on expert judgment. Bayesian probabilistic tools can support decision-making about the causes of geotechnical failures. This thesis presents a Bayesian methodology for decision support in forensic geotechnical engineering based on two probabilistic techniques: Bayesian inference via posterior odds ratio and Bayesian Networks. The methodology compares probabilistically the hypotheses formulated as causes of failure and evaluates the influence of the amount of information (evidence) included in the analysis. Two benchmark problems and a case study were used to validate the applicability of the methodology. The results show that the Bayesian methodology identifies the most likely cause of a geotechnical failure, even when the amount of evidence is sparse. The use of the proposed methodology improves decision-making processes related to the causes of geotechnical failures.

Keywords: Forensic geotechnical engineering, Bayesian Networks, Bayesian inference.

Resumen

Metodología de Redes Bayesianas para el Apoyo en las Decisiones de Ingeniería Geotécnica Forense

Los avances recientes de la ingeniería han aumentado la expectativa de la comunidad de que las obras civiles funcionen con seguridad. Ocasionalmente, algunas de estas obras fallan debido a errores humanos o al comportamiento imprevisible de los materiales. La ingeniería forense es la rama de la ciencia forense encargada de investigar las fallas en ingeniería. Los métodos científicos utilizados por la ingeniería forense garantizan que las conclusiones sobre las causas de una falla provengan de procesos de investigación confiables. Sin embargo, en el caso de fallas geotécnicas, la incertidumbre inherente a los materiales de suelo y roca, las dificultades en la recolección de evidencia y la multiplicidad de escenarios de falla (hipótesis) suponen un reto para identificar las verdaderas causas de falla. En consecuencia, las conclusiones relacionadas con las causas de fallas geotécnicas algunas veces lucen arbitrarias y sesgadas porque se basan principalmente en el juicio de los expertos. Las herramientas probabilísticas bayesianas pueden apoyar la toma de decisiones sobre las causas de fallas geotécnicas. Esta tesis presenta una metodología bayesiana de apoyo a la toma de decisiones en ingeniería geotécnica forense utilizando dos técnicas probabilísticas: Inferencia bayesiana empleando las técnicas *posterior odds ratio* y *Redes Bayesianas*. La metodología compara probabilísticamente las hipótesis formuladas como causas de una falla y evalúa la influencia de la cantidad de información (evidencia) incluida en el análisis. Se presentan dos problemas de referencia y un caso de estudio para su validación. La metodología bayesiana identifica la causa más probable de la falla, incluso cuando la cantidad de evidencia es escasa. Además, su aplicación mejora la toma de decisiones relacionadas con las causas de fallas geotécnicas.

Palabras clave: Ingeniería geotécnica forense, redes bayesianas, inferencia bayesiana

Content

List of Figures

Pag.

List of Tables

Pag.

1.Introduction

1.1 Background

Geotechnical engineering has seen remarkable advancements in recent times, driven by the dedication, scientific expertise, and innovative materials developed by enthusiastic engineers. This progress has increased the community's expectation that civil works such as buildings, dams, levees, or excavations operate safely. Occasionally some civil engineering works fail because of unpredictable behavior or human errors. Such failures remind engineers that there is still a long way to go in building safer structures and reducing risk.

Some engineering failures are catastrophic and can considerably impact society, the environment, or the economy. However, most of them merely impact the structure's serviceability or have little effect on its functionality. Regardless of the failure's magnitude or the structure's importance, each failure becomes an opportunity to enhance knowledge and test the validity of engineering models. Forensic engineering is the right tool for investigating such failures because it provides a reliable framework. In addition, it is based on scientific methods, which ensure that the results come from a sequence of phases that lead to repeatable, measurable, and testable conclusions.

Uncertainty plays a significant role in forensic engineering because the evidence is frequently incomplete, fragmented, or blurred. Uncertainty has traditionally been addressed from a qualitative perspective through techniques such as event and causal factor analysis (ECFA), human performance evaluation process (HPEP), and fault tree analysis, among others (Noon, 2009). However, a more rigorous approach to uncertainty management by combining abductive reasoning and Bayesian statistics has been proposed by researchers such as Biedermann et al. (2005), Taroni et al. (2014), and Bensi (2010).

Sometimes, the causes of failures are attributable to geotechnical/geological origins. Forensic geotechnical engineering (hereafter FGE) investigates these causes by using particular geotechnical methods but preserving the forensic engineering approach and the principles of the scientific method. FGE has been associated with legal disputes (Lacasse, 2016; Rao & Babu, 2016; Day, 2010). However, during this century, the main focus has shifted to improving geotechnical works, determining the causes of failures, and learning from mistakes (Terwel et al., 2018).

1.2 Problem Statement

Geotechnical engineers recognize that determining the cause (or causes) of failure is challenging, owing to factors such as incomplete evidence and the inherent uncertainty associated with failure investigations. On the one hand, failures encompass multiple variables, scenarios, and pieces of evidence that must be organized and classified. On the other hand, once the evidence has been evaluated, several hypotheses about causes of failure must be tested against the available evidence. The scientific method supports the testing process, but the decision regarding the most probable cause of failure usually relies on qualitative assessments or expert opinion. Moreover, the hypotheses are rarely compared using probabilistic techniques despite their recognized benefits in FGE (Bea, 2006).

As a result, conclusions about the causes of geotechnical failures sometimes seem arbitrary or biased (Kool et al., 2019). Biased conclusions may be due to the misapplication of engineering design methods in FGE assessments and the lack of a probabilistic framework for testing and comparing failure hypotheses. For example, Brady (2012) recognizes that design methods are unsuitable for FGE because they align available evidence with preconceived notions of the causes of failure. In addition, Phoon et al. (2016) identify the lack of a rigorous probability framework and the potential applicability of reliability concepts in FGE. In summary, the followings knowledge gaps have been identified:

- Conclusions regarding the causes of failure sometimes seem arbitrary and biased due to the misapplication of design procedures in FGE. For example, collected evidence is forced to conform to biased hypotheses rather than the evidence leading to the formulation of hypotheses.

- Complex geotechnical models and sophisticated constitutive soil models usually support forensic investigations. However, final decisions regarding the cause of failure are mainly based on expert opinion. Therefore, the decision-making process is frequently untraceable and irreproducible.

- Probabilistic methods, especially Bayesian tools, have the potential to be applied in FGE, yet they are seldom used.

1.3 Aim of the Thesis

The aim of this thesis is to develop a Bayesian methodology for supporting decisions about the causes of failures with geotechnical origins. In particular, the proposed methodology uses Bayesian inference and Bayesian network tools. The following three specific objectives were formulated in order to achieve the aim of the thesis:

1. To define a procedure for testing hypotheses about the causes of geotechnical failures. The procedure relies on Bayesian inference (*posterior odds ratio* technique) and *Bayesian networks.*

2. To assess the influence of multiple pieces of evidence on the conclusions about the causes of geotechnical failures.

3. To validate the proposed methodology by applying it to an actual failure case.

1.4 Outline of structure

The thesis is structured as follows: Chapter 2 briefly introduces Bayesian statistics and Bayesian networks, gives key definitions regarding probability, Bayes' theorem, hypotheses comparison, and direct acyclic graphs, and describes the steps to turn a causal graph into a Bayesian network. Chapter 3 provides an overview of forensic engineering and FGE and describes inductive, deductive, and abductive reasoning processes. It also depicts and discusses the stages of the forensic geotechnical process and some additional aspects to consider, such as back analysis and technical shortcomings. Chapter 4 presents a Bayesian-based methodology for supporting decisions about the causes of failures. The methodology

combines the abductive reasoning process with the inference properties of Bayesian Networks and Bayesian statistics. Chapter 5 analyzes an actual geotechnical failure using the proposed methodology. Chapter 6 discusses the major findings of the thesis. Finally, conclusions and suggestions for future work are presented in Chapter 7.

2. Introduction to Bayesian Inference and Bayesian Networks

This Chapter presents a general overview of the fundamentals of Bayesian statistics and Bayesian Networks (BNs). Bayesian statistics is a theory based on Bayes' theorem that assumes a different interpretation of probability and uncertainty. Within Bayesian statistics, BNs are one of the probabilistic graphic models used to support decision-making. The following paragraphs provide key concepts about statistics, probability, and Bayesian analysis. Standard textbooks such as Koller & Friedman (2009) can be consulted for a more comprehensive review.

2.1 Bayesian Statistics

2.1.1 Uncertainty and probability

Engineering systems, natural environments, and even daily life events are all subjected to *uncertainty*. Uncertainty may arise from a lack of information, unobservable or partly observable outcomes, and physical or economic constraints. Even if these problems can be overcome, systems are inherently stochastic. From a practical geotechnical engineering approach, uncertainty can be categorized as natural and epistemic.

Figure **2-1** shows the uncertainty categories proposed by Baecher & Christian (2003). Natural uncertainty refers to the intrinsic randomness of nature, which can be observed over time and space. For example, several locations within the same homogeneous soil will show different values for a physical-mechanical characteristic (e.g., deformability, shear strength, hydraulic conductivity) measured at a specific date. Also, the exact locations will show different physical-mechanical values if measurements are performed at a later point in time. On the other hand, epistemic uncertainty involves the lack of knowledge and information about an engineering

system. It may arise from measurement inaccuracy, hidden information, unsuitable models, or knowledge limitations.

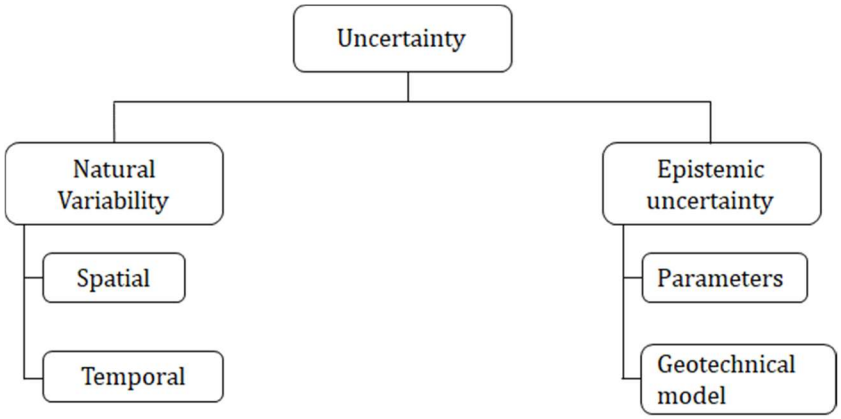


Figure 2-1. A practical approach to uncertainty categories in geotechnical engineering. Modified from Baecher & Christian (2003).

From a mathematical perspective, the measurement of uncertainty is carried out through *probability* theory. Probability assigns a number to each possible outcome (or event E_i) of the sample space (W). The sample space refers to all possible outcomes related to an experiment, parameter, or system involving chance. Outcomes and probability assignments follow specific rules (Jensen & Nielsen, 2007):

1. Outcomes are mutually exclusive. That is, two or more outcomes cannot co-occur.

2. The probability assigned to an outcome must be a non-negative real number:

$$P(E_i) \geq 0, P(E_i) \in \mathbb{R} \tag{2-1}$$

The probabilities of all potential outcomes within the sample space must add up to 1.0

$$\sum_{i=1}^{n} P(E_i) = P(\Omega) = 1.0 \tag{2-2}$$

3. The probability that two mutually exclusive outcomes E_i or E_j occur is the sum of each individual probability:

$$P(E_i + E_j) = P(E_i) + P(E_j) \tag{2-3}$$

2.1.2 Probability: Frequentists vs. Bayesians

There are two definitions of probability depending on the philosophical perspective of the term. On the one hand, the frequentist approach assumes that a population's statistical parameters are fixed but unknown (Bolstad, 2010). In this case, the probability measures the uncertainty of those unknown fixed parameters. Frequentists estimate probabilities in two different ways (Kruschke, 2015): (1) from relative frequencies calculated from long-run experiments and (2) by deriving it mathematically when an experiment and its outcomes are simple to interpret. Frequentists choose the best statistical parameters by rejecting hypotheses that include values with a low chance of being observed (Baecher & Christian, 2003).

On the other hand, the Bayesian approach assumes probability as a rational degree of belief (Baecher, 2017). This approach holds the idea that the natural world has no "fixed" statistical parameters but several probable states. Consequently, the probability measures the natural world's possible states that can be estimated through an updating process. This process, known as "*explaining away*," simulates the bidirectional pattern (i.e., backward and forward processes) of human reasoning for making decisions under uncertainty. Unlike frequentist schools, hypothesis rejection is unnecessary because the best statistical parameters are chosen based on prior assumptions and actual data.

The philosophical confrontation between these two probability schools has a long history. The main aspects of this conflict are: (i) how to analyze data, (ii) how to update beliefs in the light of new information, and (iii) how to make rational decisions under uncertainty (McGrayne, 2011). In recent years, there has been an increasing debate about the validity of p-values used in the frequentist approach. Some researchers, such as Wagenmakers et al. (2008), have suggested that frequentist inference is not an appropriate statistical interference method. However, the Bayesian school is strongly criticized for using subjective beliefs that may lead to different conclusions depending on prior beliefs.

2.1.3 Bayes' Theorem

The theorem was proposed by Reverend Thomas Bayes in 1763 and then rediscovered and widely used by Pierre Simon Laplace in 1774. For almost 150 years, the theorem was practically forgotten until the Second World War when Alan Turing decoded a German secret code using a Bayesian approach. Since then, the Bayesian theorem has been used in fields as diverse as medical diagnosis, cybersecurity, insurance, business decisions, spam filtering, and image processing. Despite some critics, the Bayesian approach to probability is gaining interest due to its advantages in decision-making problems (McGrayne, 2011).

Bayes' theorem is a precise mathematical technique for analyzing information based on experience and knowledge (Stone, 2013). The theorem can be deduced from the concept of conditional probability. Let m and d be two different events. The conditional probability states the following: the probability of m given d is the probability that m and d occur together, divided by the probability that d occurs at all (Equation 2-4.)

$$P(m|d) = \frac{P(m \cap d)}{P(d)} \tag{2-4}$$

Similarly, the conditional probability of d given m is presented in Equation 2-5.

$$P(d|m) = \frac{P(d \cap m)}{P(m)} \tag{2-5}$$

Equating 2-4 and 2-5, and admitting that $p(m \cap d) = p(d \cap m)$, algebraic manipulations lead to Equation 2-6.

$$P(m|d) = \frac{P(d|m)\, P(m)}{P(d)} \tag{2-6}$$

Equation 2-6 is known as Bayes' theorem or Bayes' rule. Bayes' theorem states that the probability of m given d is the probability of d given m times the probability of m, relative to

the probability of d. Each term in Equation 2-6 has a name describing a specific information type. $P(m)$ is the prior probability of m, i.e., the probability of m before observing any data. $P(d|m)$ also known as the likelihood term, express the probability of observing d given that m is true. $P(d)$ is the marginal probability of d or the normalization term. Finally, $P(m|d)$ is the posterior probability that describes the probability of m being true, given the observed data d. Equation 2-6 can be rewritten as Equation 2-7 or 2-8, depending on whether discrete or continuous variables are used. Equation 2-7 is suitable for discrete variables, whereas Equation 2-8 applies for continuous variables.

$$P(m \mid d) = \frac{P(d \mid m)\, P(m)}{\sum_m P(d \mid m)\, P(m)} \tag{2-7}$$

$$P(m \mid d) = \frac{P(d \mid m)\, P(m)}{\int_m P(d \mid m)\, P(m)\, dm} \tag{2-8}$$

Information for Bayesian inference may come from two sources: the prior probability $P(m)$ that express the belief of m and the likelihood $P(d|m)$ that includes actual data. Therefore, according to Equations 2-6 to 2-8, the posterior probability $P(m \mid d)$ can be described as a combination of prior beliefs and actual data. When actual information is scarce $P(m)$ controls the posterior probability. On the contrary, if actual information is available, $P(m \mid d)$ controls the posterior probability.

Although deriving Bayes' theorem is relatively easy, its application to actual inference problems may be challenging. Before the advent of computers, Equations 2-6 to 2-8 could only be solved using closed-form expressions known as conjugate distributions. These expressions include posterior, likelihood, and prior distributions belonging to the same probability distribution family. For example, a Gaussian likelihood and a Gaussian prior will lead to a Gaussian posterior distribution. Similarly, if the likelihood is defined by a Bernoulli distribution and the prior is a Beta distribution, the posterior will be a Beta distribution. Nowadays, Bayesian inference is not restricted to conjugate distributions due to the

development of efficient computer algorithms such as Monte Carlo simulations, Metropolis–Hasting sampling, and Gibbs sampling.

In the case of geotechnical engineering, some researchers have pointed out the importance of the Bayesian probability approach. Baecher (2017) argues that most geotechnical engineers intuitively practice Bayesian inference. For example, the emblematic observational method attributed to Terzaghi and then described by Peck (1969) uses a practical and qualitative Bayesian approach. Several researchers such as Wang et al. (2016); Sakurai et al. (2003); Baecher & Christian (2003); Feng (2015); Ering & Sivakumar Babu (2017); Calvello et al. (2017); Berti et al. (2012); Hasan & Najjar (2013); and Wu(2011), recognize the advantages of Bayesian approach to probability over traditional frequentist analysis in geotechnical engineering.

2.1.4 Hypotheses Comparison: Bayes Factor and Posterior Odds

As mentioned in Section 2.1.3, Bayesian statistics analyze the information by combining experience and data. It aims to update prior beliefs through the acquisition of new information. The updating process is achieved via the Bayes theorem in which prior beliefs (i.e., the prior probability term $P(m)$ in Equation 2-6) are combined with the collected information (i.e., the likelihood term $P(d|m)$) to update prior believes (i.e., the posterior probability term $P(m|d)$ in Equation 2-6). In other words, Bayes theorem estimates the probability of an event m given that data d are observed.

In addition to updating event probabilities, Bayes' theorem compares competing hypotheses. If the event m is replaced by a hypothesis H in Equation 2-6 and the marginal probability $P(d)$ is dropped, then Equation 2-9 is obtained. The term $P(d)$ can be discarded because it only normalizes the posterior probability (i.e., gives values between 0 and 1 to the posterior probability).

$$P(H|d) \propto P(d|H)\, P(H) \tag{2-9}$$

Equation 2-9 states that posterior probability is proportional to likelihood multiplied by the prior probability. If H_1 and H_2 are two competing hypotheses, then the ratio of posterior

probabilities can be used to compare these hypotheses (see Equation 2-10). The first term in Equation 2-10 is the posterior odds ratio and represents how better H_1 explains the data than H_2. The second term is known as Bayes Factor (BF). It denotes the ratio between likelihoods. The third term is the prior odds which express the ratio between prior probabilities of H_1 and H_2 before observing some data. Equation 2-10 plays a significant role in model comparison and null hypothesis testing and supposes an improvement to the classical Null Hypothesis Significance Testing (Gelman & Tuerlinckx, 2000).

$$\frac{P(H_1|d)}{P(H_2|d)} = \frac{P(d|H_1)}{P(d|H_2)} \frac{P(H_1)}{P(H_2)} \tag{2-10}$$

Jeffreys (1961) proposed an evidence scale for interpreting BF and posterior odds values (**Table 2-1**). Similarly, Kass & Raftery (1995) suggested a BF scale based on a minor modification of the Jeffreys scale. The BF or posterior odds values can be interpreted as follows: For example, a $BF_{12} = 7$ (or $posterior\ odds_{12} = 7$) means that H_1 explains the data seven times better than H_2. According to **Table 2-1**, BF=7 is interpreted as "Positive" evidence favoring H_1 over H_2. In the case of a BF or a posterior odds values less than 1, **Figure 2-2** can be used to interpret the results. For example, $BF_{12} = 0.14$ indicates that there is "positive" evidence favoring H_2 over H_1.

Table 2-1.: Jeffreys (1961) scale and Kass & Raftery (1995) modified scale.

Jeffreys scale
(1 – 3.2) Not worth more than a bare mention
(3.2 – 10) Substantial
(10 – 100) Strong
(>100) Decisive

Kass and Raftery scale

(1 – 3) Not worth more than a bare mention

(3 – 20) Positive

(20 – 150) Strong

(>150) Very Strong

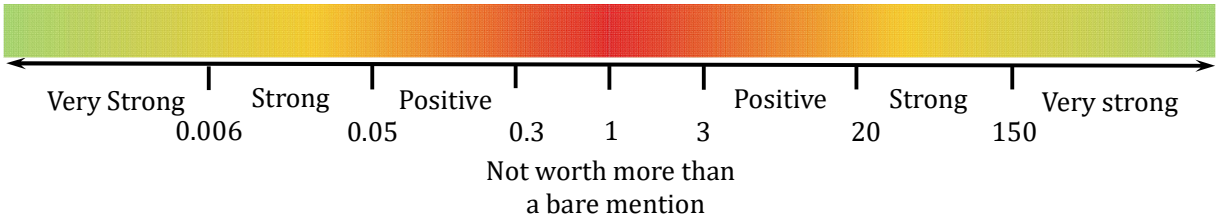


Figure 2-2.: Graphical representation for interpreting the BF and posterior odds values.

In the case of multiple comparisons hypotheses, Bayesian analysis completely overcomes the problems when several hypotheses are compared simultaneously via classical *p-values*. The classical frequentist approach resorts to correction factors to avoid encountering false-positive errors (Type I errors) during multiple comparisons (refer to Gelman et al., 2012 for further details). The correction factors are grouped in a concept known as Family Wise Error Rate and include the Bonferroni, Holm, and Hochberg correction factors. In contrast, Bayesian multiple comparisons do not require correction factors since their comparative process incorporates more information than *p-values*. Therefore, it is more adept at mitigating size effects and preventing Type I errors (Gelman & Tuerlinckx, 2000).

The Bayes factor, posterior odds, and model comparison process intend to understand how well a hypothesis explains the evidence. Their purpose is not to gather evidence to support a particular hypothesis as some statistical malpractices do (e.g., hypothesis fishing, refer to Dahl et al., 2008 for further explanation). Therefore, the Bayesian methodology proposed in this thesis is based on the ability of Bayesian analysis to compare hypotheses and determine which of them better explains a geotechnical failure. Chapter 4 expands the description of the proposed methodology.

2.2 Bayesian Networks

The primary goal of Bayesian Networks (BN) is to estimate variable values that cannot be directly observed or that can be measured but at an unacceptable cost (Jensen & Nielsen, 2007). To achieve this goal, causal graphs $\mathcal{G}$ and probability relationships $\wp$ are used to construct a Bayesian network. Once $\mathcal{G}$ and $\wp$ are defined, they may reveal valuable attributes for variables and their causality relationships. This section presents a BN overview, some basic definitions, and the process for eliciting conditional probability distributions.

2.2.1 Bayes' Theorem and Bayesian Networks

As mentioned before, Bayes' theorem is a powerful tool for updating beliefs and analyzing data based on prior information (i.e., prior beliefs) and evidence (i.e., observations or actual data). The direct use of Equations 2-6 to 2-8 may be helpful in simple problems involving a few variables and a few pieces of evidence. However, real-world problems, particularly geotechnical systems, involve several unknown events, many variables, and limited observations. Bayes' theorem can be successfully applied in those complex problems if combined with graphical models. The result is a probabilistic graphic model known as a Bayesian Network.

BNs have some advantages over the plain use of Bayes' theorem. These advantages include updating information in multivariable models, a transparent inclusion of evidence, and the capacity to track the updating process (Correa et al., 2009). Additionally, graphical representation facilitates understanding causal relationships between variables and provides a clear language of communication. Unlike most machine learning tools (e.g., neural networks and support vector machines), BNs' topology encodes dependencies between variables. Therefore, findings and outcomes can be easily explained when new information is included in BNs (Friedman et al., 1997).

2.2.2 Basic definitions

A decision-making problem under uncertainty can be structured through a graphical model known as a causal graph $\mathcal{G}$ or causal network (see Figure **2-3**). Mathematically, $\mathcal{G}$ consists of a set of nodes $\mathcal{X} = \{X_1, \dots, X_n\}$ and a set of directed arrows $\mathcal{E}$ connecting nodes (e.g., $X_i \rightarrow X_j$). Nodes may encode random variables, events, sample spaces, a set of states, hypotheses, or

propositions. Arrows are a set of information channels that represent causal relationships between nodes. When $\mathcal{E}$ does not form a loop between $\mathcal{X}$, $\mathcal{G}$ is called a direct acyclic graph (DAG).

The basic notation of a DAG consists of denoting nodes as parents and children. In the expression $X_i \rightarrow X_j$, X_i is the parent of X_j and X_j is the child of X_i. In general, $Pa(X_i)$ denotes the parents of node X_i and $Ch(X_i)$ represents the children of X_i. Figure **2-3** presents an example of a causal graph $\mathcal{G}$ with nodes $\mathcal{X} = \{A, B, C, D, E\}$. C and D are children of B, A is a parent of C and E is a child of C. Note that A and B have no parents. Consequently, A and B are denoted as root nodes.

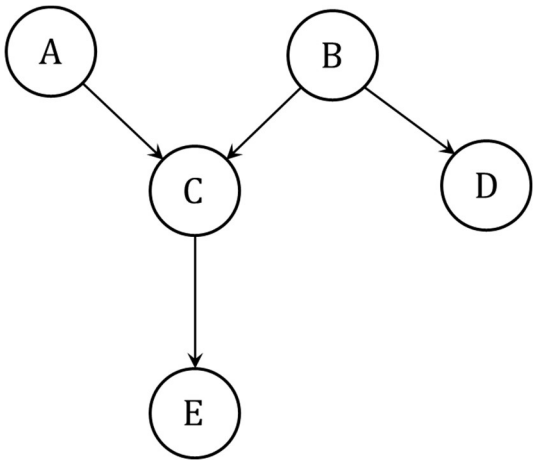


Figure 2-3.: Example of a simple causal graph (DAG)

2.2.3 Connection types

Nodes in a DAG can be connected in three simple configurations: serial, diverging, and converging connections. Each connection type is related to causal relationships between nodes and how one node may change another node's state. Also, connection types define how information is propagated through a DAG when some evidence is included in one or more nodes. As shown below, arrows represent causality direction between nodes but do not necessarily indicate the direction of information flow.

Figure **2-4** presents a serial connection. In this simple DAG, A affects B, which in turn influences C. In other words, any evidence entered in A will propagate to B and then to C (see

Figure **2-4**a). Reasoning from A to C is known as a *causal trail*. Similarly, any evidence entered in C will propagate to A through B (see Figure **2-4**b). This latter type of reasoning is known as an *evidential trail*. In the case of instantiating B (i.e., entering evidence in B), any evidence entered in A will be irrelevant for C. Therefore, when B is instantiated, the flow channel between A and C is blocked, and it is said that A and C become independent given B.

A diverging connection, also known as a common cause, is shown in Figure **2-5**. A and C are children of B or mathematically $Ch(B) = \{A, C\}$. Any information included in A will influence C (see Figure **2-5**a). Similarly, any evidence entered in C will influence A through B (see Figure **2-5**b). However, if B is instantiated, any evidence entered in A will be irrelevant to C (see Figure **2-5**c). In this case, it is said that A and C become independent given B.

A converging connection, also known as a common effect configuration, is shown in Figure **2-6**. Information propagation in converging connections differs from serial and diverging connections. The converging connection in Figure **2-6** can be expressed mathematically as $Pa(B) = \{A, C\}$. Clearly, any evidence entered in A or C will propagate to B (see Figure **2-6**a, b). By extension, any knowledge of A does not influence C. Similarly, any knowledge of C does not influence A. In other words, knowledge of one potential cause does not influence other potential causes. However, if any information is entered in B (the common effect), the information within B will influence both A and C (see Figure **2-6**c). This type of information flow, known as the *explaining away effect,* is a powerful characteristic used in abductive reasoning.

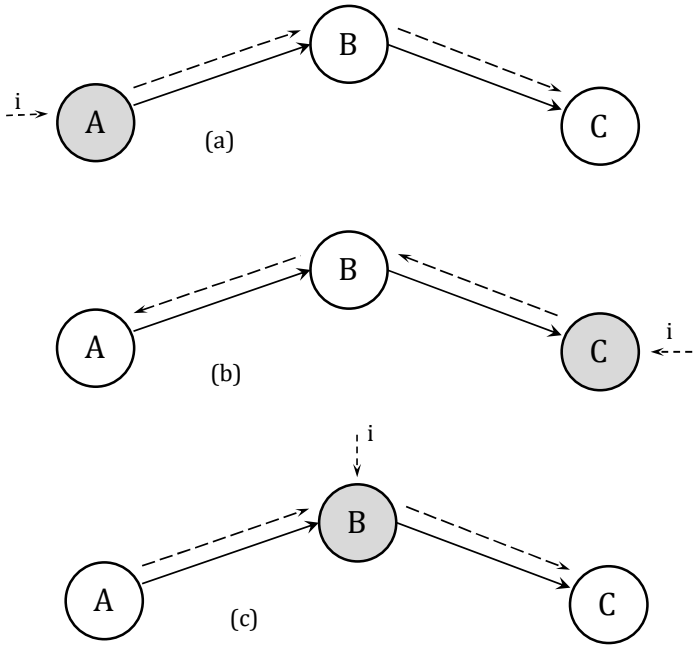


Figure 2-4.: Serial connection and its flow of information (grey nodes indicate instantiation i). (a) Node A is instantiated, (b) Node C is instantiated, and (c) Node B is instantiated.

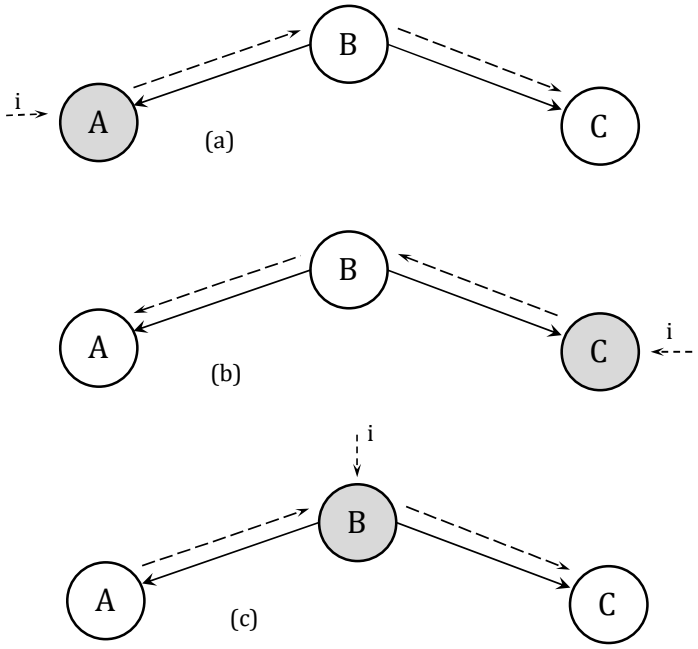


Figure 2-5.: Diverging connection and its flow of information (grey nodes indicate instantiation i). (a) Node A is instantiated, (b) Node C is instantiated, and (c) Node B is instantiated.

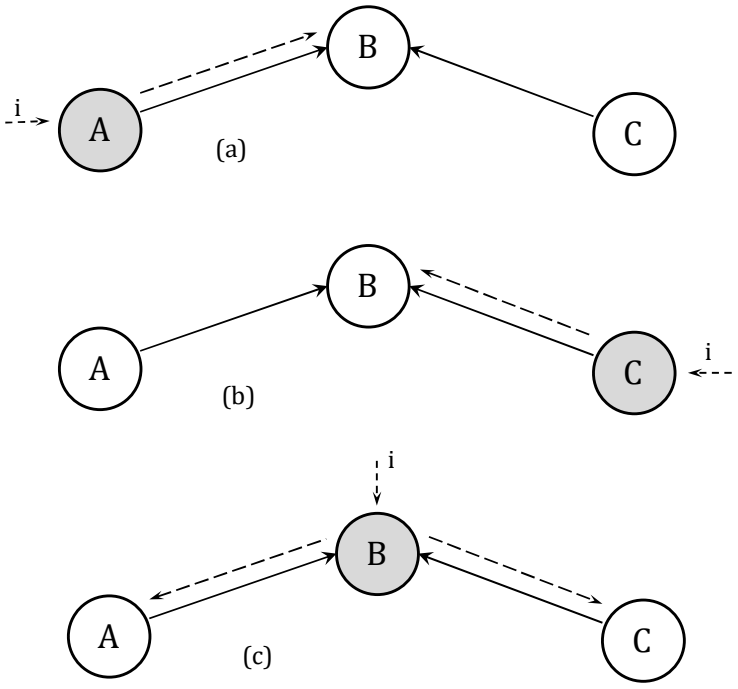


Figure 2-6.: Converging connection and its flow of information (grey nodes indicate instantiation i). (a) Node A is instantiated, (b) Node C is instantiated, and (c) Node B is instantiated.

2.2.4 D-separation and Markov Blankets

Section 2.2.3 has intuitively introduced the d-separation concept. The three basic connection types presented above encompass how information can be transmitted through a causal graph. Also, they cover the rules to decide if a pair of nodes are independent (i.e., separated) given any evidence included in a causal graph. Accordingly, the d-separation concept can be defined as follows (Jensen 2007): Two nodes A and C in a DAG are d-separated if there is an intermediate node B connecting those nodes such that:

- The connection type between A and C is serial or diverging, and B is instantiated, or
- The connection type between A and C is converging, and neither B nor any child of B is instantiated.

In any other case, A and C are d-connected.

The d-separation (or d-connection) concept is related to the idea of a DAG path that can transmit information. For example, two nodes A and C in a DAG might be connected by several paths. However, if all the paths that connect A and C are inactive, A and C are d-separated. Conversely, if there is at least one active path between A and C, it is said that A and C are d-connected.

A Markov blanket extends the idea of d-separation to causal graphs containing many nodes (Jensen & Nielsen, 2007). A node value can be inferred using only the information from a subset of nearby nodes. Information from distant nodes is frequently useless due to d-separation rules within the causal graph. For example, Figure **2-7** presents the Markov Blanket for node I. The Markov Blanket has the particular property that when it is instantiated, the node I remains d-separated from the rest of the nodes in the network (outside the red dotted line). Mathematically, the Markov Blanket of node X_i can be defined as the set that contains the children of X_i, the parents of X_i and all nodes that share a child with X_i.

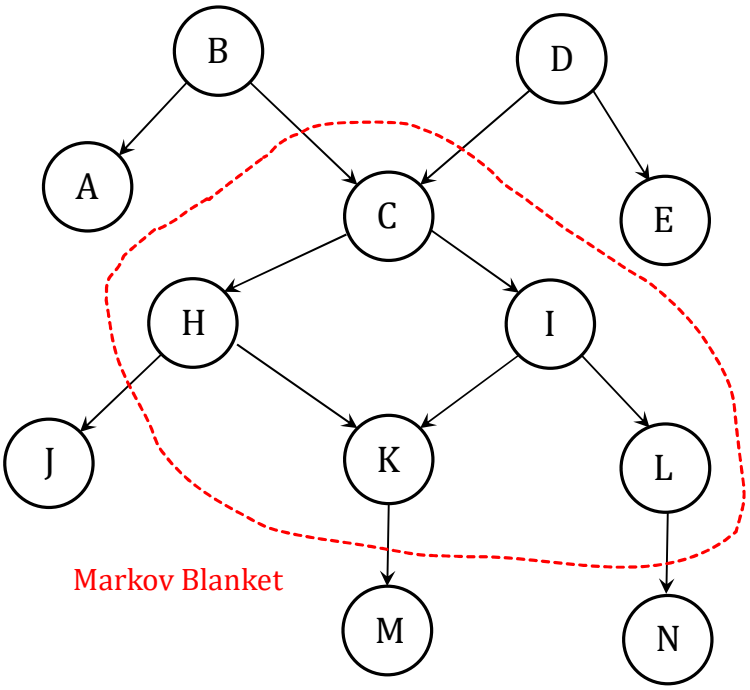


Figure 2-7.: Markov blanket for node I

2.2.5 From Causal Graphs to Bayesian Networks

The set of directed arrows $\mathcal{E}$ in $\mathcal{G}$ usually represents the strength of cause-effect relationships between nodes. This relationship can be expressed quantitatively using probability calculus (Jensen & Nielsen, 2007). For example, the arrow in Figure **2-8**a represents the relationship between A and B, which can be described probabilistically as $P(A \mid B)$. However, if node A has several parents (e.g., nodes B, C, D, E in Figure **2-8**b) then arrows represent the probabilistic interaction $P(A \mid B, C, D, E)$ between A and their parent nodes.

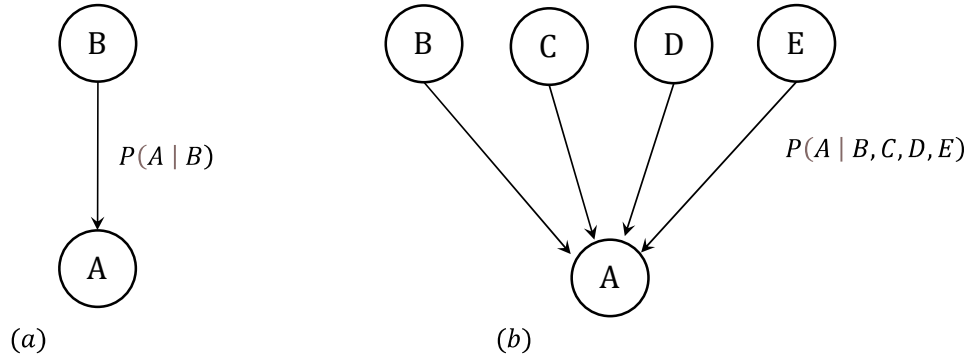


Figure 2-8.: Probabilistic relationships between nodes. (a) Simple connection. (b) multiple connections.

When arrows in a *DAG* represent probabilistic relationships and nodes have finite states, the resulting probability graph model is a Bayesian Network. The following is a formal definition based on Jensen & Nielsen (2007) and Koller & Friedman (2009): A Bayesian network is a model consisting of a set of probability relationships $\wp$ that factorize over a graph $\mathcal{G}$. $\wp$ is expressed as conditional probability distributions (*CPD*) for nodes in $\mathcal{G}$. In summary, a Bayesian Network must meet the following conditions:

- Include a set of nodes and a set of arrows connecting the nodes.
- The set of nodes and arrows must form a *DAG.*
- Nodes must have a finite set of mutually exclusive states.
- A *CPD* must be attached to each node.

2.2.6 The Chain Rule

BN's ability for decision-making problems is due to the combining properties of Bayes' theorem and the Chain Rule. On the one hand, Bayes' theorem, described in section 2.1.3, updates prior beliefs as new information is obtained. On the other hand, the chain rule reproduces the causal relationships between nodes using conditional probabilities. Let $\mathcal{X} = \{X_1, \dots, X_n\}$ be the set of variables that represents a problem. The joint probability distribution of $\mathcal{X}$ is $P(\mathcal{X}) = P(X_1, \dots, X_n)$. For example, if $n = 10$ and each variable is binary-valued, the joint probability distributions would require $2^n - 1 = 1023$ numbers for specifying $P(\mathcal{X})$. Even for a few variables with binary states, the joint distribution is inaccessible to a human expert. Furthermore, it is computationally expensive.

The required number for specifying a joint distribution grows exponentially as the number of variables and states increases. However, a more compact version of $P(\mathcal{X})$ can be achieved by using the properties of BNs related to conditional probabilities and d-separation rules between variables. The compact version of $P(\mathcal{X})$ corresponds to the chain rule, which can be expressed as shown in Equation 2-11. Using the notation $Pa(X_i)$ for denoting the parents of X_i, the chain rule for Bayesian Networks can be rewritten as Equation 2-12.

$$P(\mathcal{X}) = P(X_1, \dots, X_n) = P(X_n \mid X_1, \dots, X_{n-1})\, P(X_{n-1} \mid X_1, \dots, X_{n-2}) \dots P(X_2 \mid X_1)\, P(X_1) \tag{2-11}$$

$$P(\mathcal{X}) = \prod_{i=1}^{n} P\big(X_i \mid Pa(X_i)\big) \tag{2-12}$$

2.2.7 Causal Graph Construction

Scutari & Denis (2015) recognize three approaches to creating a causal graph $\mathcal{G}$: (i) data-driven, (ii) expert-driven, and (iii) hybrid approach. Data-driven modeling derives the underlying causal relationships using the observed data, in other words, learning the BN's structure from data and uncovering the conditional probability distribution that controls the problem. The expert-driven approach involves defining causality between variables by a human expert or deriving it using well-known physical or mathematical relationships. The

third approach combines the previous two to generate a more robust model. In this study, only the expert-driven approach is employed due to the characteristics of geotechnical failure problems.

In general, the construction of $\mathcal{G}$ using the expert-driven approach consists of the following steps:

1. *Identification of variables and node creation*: The aim is to recognize all relevant variables in the problem. Variables can be divided into two general groups. The first group includes those variables that cannot be directly observable or whose observation implies an unacceptable cost. These variables are known as hypothesis variables. The second group of variables includes those that provide some information and modify hypothesis variables. This second group is called information or evidence variables. Once relevant variables have been identified, each variable is placed at a unique node.

2. *Node type definition and number of states:* The values of the variables can be discrete or continuous. Defining variable types depends on their characteristics and complexity. Discrete variables can assume values among a countable set of values. The simplest discrete case is the Boolean type, in which the discrete space contains only two states (e.g., 0/1, yes/no, failure/no-failure). Rank discretization extends the discrete space to more than two states (e.g., low/medium/high, 0/1/3/4, excellent/good/regular/bad). Probability mass functions can be used to represent probability distributions of discrete variables

   Continuous variables assume a range of values among an infinite set of states. They are represented through probability density functions (PFD) such as Gaussian, Poisson, or beta distributions. However, using continuous variables in BNS requires a discretization process that leads to a loss of accuracy. Some efficient computational algorithms, such as dynamic discretization (Neil et al., 2007), combine discrete, continuous, and non-numeric nodes to overcome the loss of accuracy difficulties.

3. *Information channels and arrows:* The flow of information through a BN requires connecting the nodes. The connecting process is carried out in such a way that it reflects the causal relationships between nodes. In other words, the BN's topology defines connection types, d-separation properties, and Markov blankets. The BN's topology must mimic the causal dependencies or independence between variables.

In the last step, Kjærulff & Madsen (2013) distinguish two ways of eliciting the causal graph structure. The first is a basic approach that uses the natural cause-effect relationship between hypothesis and information variables. In the basic approach, Kjærulff & Madsen expand the number of variable types to include background, problem, mediating, and symptom variables. For example, **Figure 2-9**a presents a simple causal graph including hypothesis and information (evidence) variables of a geotechnical problem. **Figure 2-9**b shows an expanded model of the same problem inferred from common knowledge, empirical relationships, or theoretical models such as the geotechnical triangle (Burland, 2012). Although a causal graph can be elicited using the basic approach, domain experts must verify the correct flow of information and d-separation requirements between variables.

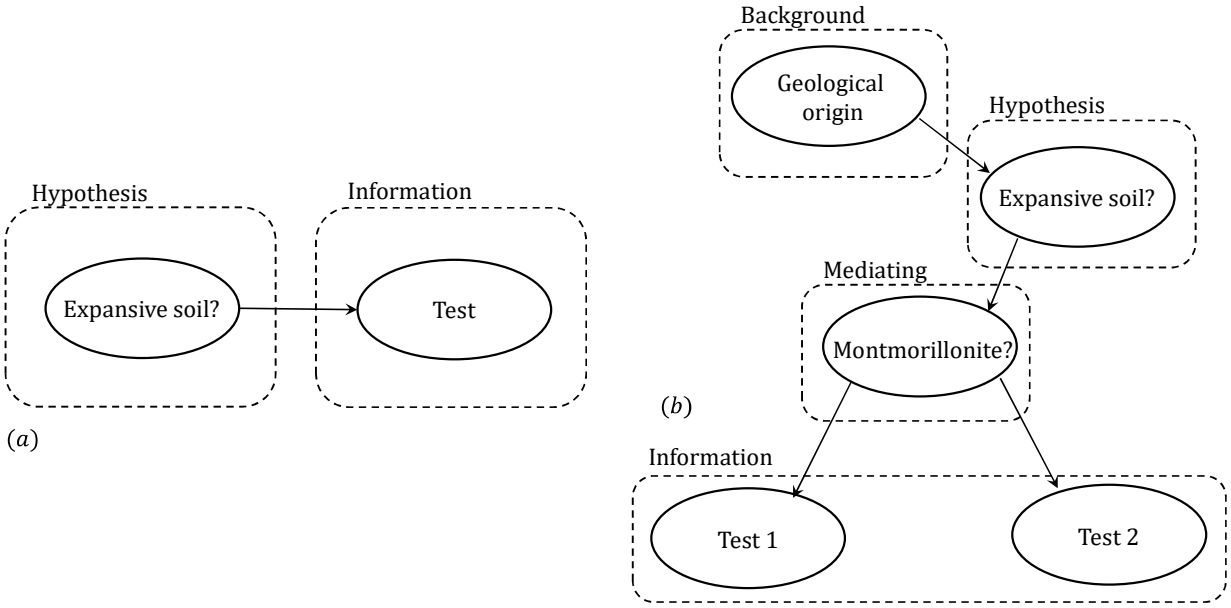


Figure 2-9.: (a) Example of a causal graph using a basic approach. (b) expanded model example.

The second approach enhances the basic approach by using idioms for constructing fragments of large BNs. Idioms denote the specialized terminology and jargon associated with a specific field of knowledge (in this case, the geotechnical engineering jargon).

Idioms operate as reasoning guidelines for building fragments of large BNs. They can represent reasoning under uncertainty through semantic relationships between variables (Jordan et al., 2013). Idioms represent an advance over the basic approach because they preserve the d-separation and Markov blanket properties required in BNs. Neil et al. (2000) argue that the following five idioms cover almost all reasoning substructures employed in a BNs:

1. Cause-consequence idioms model a cause-effect process between inputs (causes) and outputs (effects or consequences). Figure **2-10** presents cause-consequence idioms expressed as causal graphs. This idiom can be easily identified when a process follows a well-defined chronological sequence. Commonly they include words and expressions such as "*attributable to," "due to," "as a consequence of," "impact," "cause,"* and *"resulted in,"* among others. The following sentences from geotechnical literature are examples of cause/consequence idioms in geotechnical engineering jargon.

    a. *"One of the major factors contributing to the settlements in Mexico City is the regional subsidence as a consequence of the consolidation of the soft clay layers."* (Puzrin, Alonso, & Pinyol, 2010, page 6). See Figure **2-10**a
    b. *"Surface fault rupture caused by an earthquake is important because it severely damages buildings, bridges, dams, tunnels, canals, and underground utilities."* (Day, 2010, page 217). See Figure **2-10**b
    c. *"The decrease in the pore water pressure might have been caused by inward movement of the diaphragm wall, which resulted in volume expansion of the soil behind the wall."* (Iwasaki, 2016, page 535). See Figure **2-10**c
    d. *"Most of the past embankment dam failures were caused by either overtopping or seepage erosion/piping."* (Xu & Zhang, 2016, page 105). See Figure **2-10**d

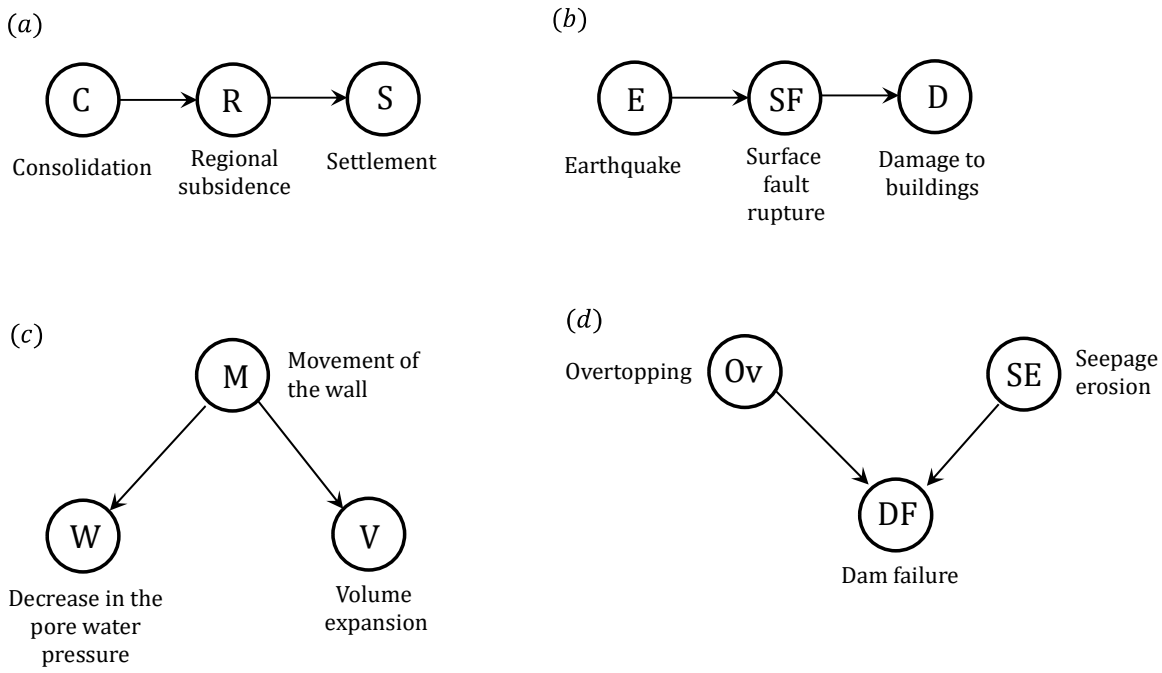


Figure 2-10.: Examples of BNs constructed from cause-consequence idioms in geotechnical engineering. (a) Settlements in Mexico City, (b) Surface fault rupture caused by an earthquake, (c) Inward movement of a diaphragm wall, and (d) dam failure.

2. Measurement idioms: Some nodes (variables) within a BN can be difficult to observe or measure directly. However, some related nodes can indirectly measure the actual value of the unobserved variable. These nodes are known as *indicators.* Measurement idioms include words and expressions such as "monitoring," "measurement," "estimation," "test," "indication," "symptom," and "features," among others. The following sentences from geotechnical literature are examples of measurement idioms used in geotechnical engineering jargon.

    a. *"The deep failure of caissons was also an indication that the foundation soil had experienced an additional reduction in strength most likely associated with soil liquefaction."* (Puzrin, Alonso, & Pinyol, 2010, page 128). See Figure **2-11**a
    b. *"One common laboratory test used to determine the expansion potential of the soil is the expansion index test. […] Other laboratory tests such as hydrometer analysis and Atterberg limits, can be used to classify the soil and estimate its expansiveness."* (Day, 2010, page 106). See Figure **2-11**b

c. *"The most significant advantages of the CPT are its simplicity, repeatability, accuracy, and continuous record. […] This paper outlines alternate ways in which the CPT can and has been correlated to liquefaction resistance…"* (Robertson & Campanella, 1985, page 384–403). See Figure **2-11**c

3. Definitional idioms model two or more nodes combined into a single node. Usually, these idioms represent a deterministic relationship between a node and its parents. For example, in Figure **2-12**, the parent node μ (pore water pressure) and σ (total stress) define the synthesis node σ' (effective stress).

4. Induction idioms are used to estimate parameter values from available data. They can be used to make predictions or model the process of statistical inference. Figure **2-13** shows a causal graph to estimate the mean value of water content from several measurements. Note that induction idioms do not reflect causality.

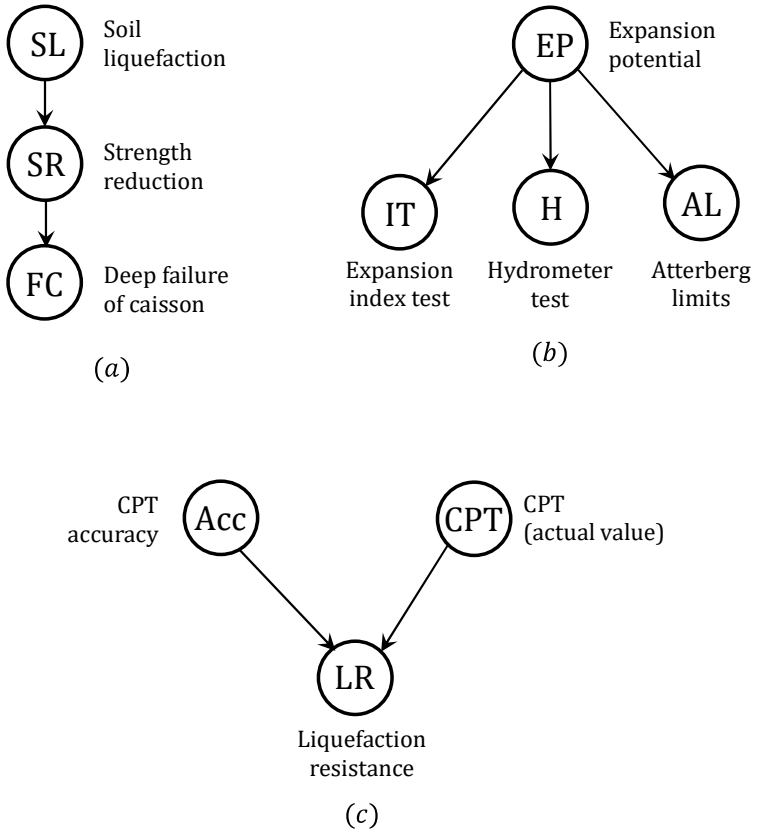


Figure 2-11.: Examples of BNs constructed from measurement idioms. (a) Deep failure of caissons, (b) Laboratory test for expansion potential, and (c) CPT and liquefaction resistance.

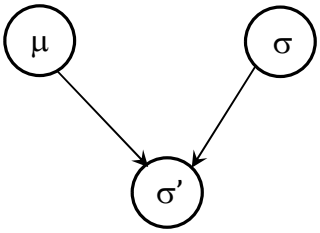


Figure 2-12.: Example of a BN constructed from a definitional idiom.

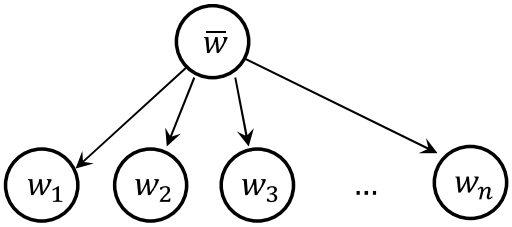


Figure 2-13.: Example of a BN constructed from an induction idiom.

5. Reconciliation idioms aim to reconcile diverse and independent sources of information into a single variable (Neil et al., 2000). They can be used when different models (or BN) are used to estimate the same variable. The reconciliation idiom may merge information from several sources or choose the most appropriate source. Figure **2-14** presents an example of a reconciliation idiom applied to estimate the shearing resistance angle from both triaxial and CPT measurements.

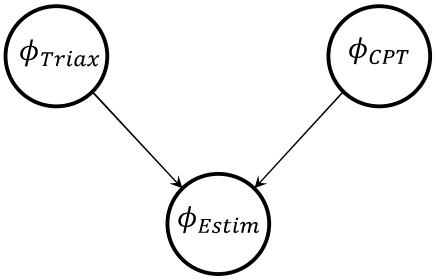


Figure 2-14.: Example of a BN constructed from a reconciliation idiom.

2.2.8 Eliciting Conditional Probability Distributions – CPD

The probability relationships $\wp$ that factorized over $\mathcal{G}$ are expressed as conditional probability distributions (CPD). A CPD encodes the probability of X_i given its parents $P(X_i \mid Pa(X_i))$. That

is, $P(X_i \mid Pa(X_i))$ represents the probability of every single value of X_i given each possible combination of $Pa(X_i)$ values. When $\mathcal{G}$ contains few nodes, and each node has few states, it is possible to express $P(X_i \mid Pa(X_i))$ using conditional probability tables (CPT) or tabular representations. However, tabular representation becomes intractable when X_i has several states and numerous parent variables. Moreover, CPT cannot be used when variables have infinite domains, as in the case of continuous variables.

Several authors have recognized that eliciting CPDs is one of the most challenging tasks in constructing a BN (Mkrtchyan et al., 2016; Druzdzel & Gaag, 2000, Marcot et al., 2006; Chen & Pollino, 2012, Fenton & Neil, 2019, Kjærulff & Madsen, 2013). That is true not only for the reasons described in the preceding paragraph but also because of the difficulties in assigning probability values when information is limited or comes from the domain of human experts. Therefore, populating CPD should be carried out right after the careful construction of $\mathcal{G}$. Otherwise, if $\mathcal{G}$ does not describe the causal relationships correctly, the elicited numbers may be useless.

The numbers composing the CPD can be estimated from expert knowledge, mathematical models, and databases. Due to the unavailability of significant and reliable databases in geotechnical engineering, especially those containing actual geotechnical failures, the examples presented in this research are focused on the former two sources. However, when databases are required, these can be retrieved from the outcomes of physical-mathematical models representing the geotechnical problem. A further description of the three sources of information for eliciting CPD is presented below

Expert knowledge

Large databases about rare and uncommon situations are scarce or even non-existent. In geotechnical engineering, this lack of data becomes critical in problems involving geotechnical failures. As a result, the primary source of information for populating CPD in reliability analysis comes from the domain of expert knowledge (Rohmer, 2020). Expert knowledge is rarely expressed by employing numerical probabilities. Therefore, qualitative assessments should be translated into quantitative measurements using questionaries, interviews, or technical literature.

Some indirect methods are available to transform qualitative knowledge into quantitative probability values. Rohmer (2020) describes three general approaches: (i) direct elicitation, (ii) assumptions on the causal structure, and (iii) filling-up techniques. Direct elicitation inquiries about the strength of causal relationships between variables using statements, analogies, or scales. The probability wheel is one of the direct technics in which a circle is divided into n segments corresponding to the possible states of a variable. Then, the human expert is queried about the probability of each state given some background information.

The probability scale is another example of direct elicitation. In this method, a human expert is queried about probabilities using predefined verbal statements. A numerical value according to a probability scale is assigned to each verbal statement. Figure **2-15** presents an example of a probability scale. In this case, emphasis has been placed on low and high values.

Additional direct techniques, such as gambling analogy and fuzzy set analysis, can assign values to expert knowledge (Rohmer, 2020). Whichever technique is used, all of them focus on populating CPD using tractable pieces of information elicited from human experts. However, direct elicitation rapidly becomes inefficient when states and parents grow. For example, a binary variable X_i with three binary parents require eight values to specify its CPD. Maybe eight values can be elicited by a human expert, but if X_i has ten binary parents, 1024 numbers must be elicited for a full CPD specification. Undoubtedly, this amount is unmanageable for a human expert.

Mathematical models

Instead of defining a conditional probability $P(X_i \mid Pa(X_i))$ for every state of X_i given every possible combination of $Pa(X_i)$, it is possible to use mathematical functions to approximate this probability. Mathematical functions include deterministic CPD, standard statistical distributions, comparative expressions, and logical operators. These functions facilitate the CPD construction and overcome the obstacles encountered in real BN applications related to computational efficiency and storage.

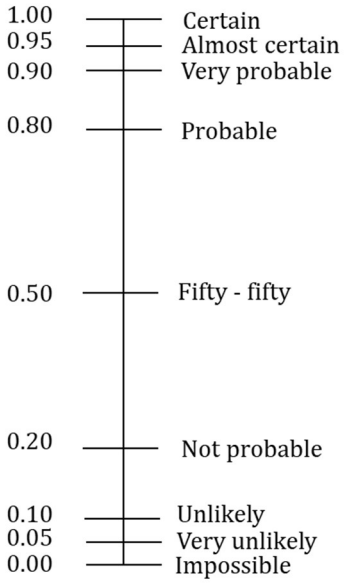


Figure 2-15.: A probability scale example for translating verbal statements into probability values (Kjærulff & Madsen, 2013).

Databases

Defining a CPD from a database assumes that some underlying process has generated the data (Kjærulff & Madsen, 2013). Consider two DAG structures, DAG1 and DAG2, defined by human experts HE1 and HE2, respectively. If each DAG represents two different underlying processes that generate the data, then the CPDs estimated by each DAG will be different. Therefore, each BN constructed from a DAG and a CPD will produce different results, but only one of them may be useful. This condition reveals the importance of constructing functional BNs that reflect the variables' causal relationships.

Constructing CPD from databases requires data to cover all relationships between variables defined in a DAG. Otherwise, the estimated parameters will not be statistically significant or may not be estimated at all. These conditions imply that databases should be long enough to cover those relationships. In this case, CPD values can be estimated from probability frequencies using the maximum likelihood estimation (MLE) or Bayesian estimation (Jensen & Nielsen, 2007).

2.2.9 Inference and Belief Updating

The primary goal of any BN is to answer questions such as (i) *What is the probability of X_1 given that some evidence in X_2 is known?, (ii)* W*hat are the most probable states for the set of variables $X' = \{X_1, X_2, \dots, X_n\}$ given that the state of X_i is known?* or (iii) *Is X_1 independent from X_2?* The first question is know as a conditional probabilistic query in probabilistic and computer science. The second corresponds to a *most probable explanation*, and the third is a *conditional independence query*. Therefore, a BN can be viewed as an expert system capable of answering questions related to some specific problem.

The process of reaching conclusions using evidence is known as inference. BNs are used as powerful probabilistic tools because they can emulate inference processes. For example, evidence can be included directly in a BN because reasoning processes are encoded in its topology. BNs allow experts to reach conclusions about the most probable values given some evidence. Besides, BNs can perform inference tasks from imperfect knowledge and incomplete information, just as human experts do (Fenton & Neil, 2019).

In the context of BNs, inference analysis has two opposite directions: forward and backward. On the one hand, forward analysis refers to BN's predictive abilities. In other words, they estimate the probability distribution of an evidence variable based on the states of hypothesis variables (refer to Figure **2-9**). Forward tasks are comparable to learning machine techniques such as neural networks, support vector machines, and regression analysis. On the other hand, the backward direction is related to diagnosis and hypothesis testing. Unlike most machine learning techniques, BNs can perform backward analysis. The backward analysis involves estimating the probability distribution of any node given some evidence entered in one or more nodes. Evidence may refer to actual observations on a particular state of a variable (i.e., hard evidence) or uncertain and incomplete observations (i.e., soft evidence). For the purposes of this thesis, only hard evidence is considered.

As described above, forward and backward inference requires entering information in some nodes to update the probability distributions in the remaining nodes. Specifically, the inference task is about computing the probability distributions of the unobserved nodes, given evidence entered in the observed nodes. Inference tasks are performed using conditional probability queries such as $P(Y \mid X = e)$, where Y denotes the unobserved nodes, and $X = e$

denotes nodes in which evidence is included. In Bayesian Network terminology, the expression $P(Y \mid X = e)$ corresponds to a belief-updating process that can be solved using exact and approximate algorithms.

Exact algorithms include variable elimination algorithm methods and junction tree methods. Conceptually, the variable elimination algorithm estimates the marginal probability distribution over a subset of variables using a sequential elimination process of those variables not involved in the subset of interest (Bensi, 2010). The order in which variables are eliminated impacts computational demand. Moreover, every conditional query made to the network requires the algorithm to be re-run. Therefore, although the elimination algorithm is relatively simple, it is not helpful in practice because computational requirements increase exponentially for BNs with multiple nodes. Many standard BN textbooks, such as Jensen & Nielsen (2007) and Koller & Friedman (2009), describe this algorithm.

The junction tree (also called clique tree) is an algorithm equivalent to the elimination algorithm because it is based on the same basic operations: factor multiplication and variables summing out. However, it relies on predefined data structures, known as cliques, which convert a BN into a tree. Cliques operate as super-nodes that include multiple variables of the original BN. These super-nodes recycle some intermediate operations of the elimination algorithm and provide answers to multiple queries using a unique data structure. Although the clique algorithm is computationally more efficient than the elimination algorithm, it demands more storage capacity since cliques may significantly increase the size of the original BN. Reference textbooks such as Koller (2009) comprehensively describe the clique algorithm.

Approximate algorithms are based on a set of techniques known as stochastic simulation. They include, among others, the rejection sampling algorithm and Gibbs sampling. The basic idea behind these techniques is to draw thousands of random observations using the BN's properties. Observations are then used for estimating conditional probabilities of interest. The most straightforward approximate algorithm is the rejections sampling technique. It generates random observations using Monte Carlo simulation and counts how many matches the evidence. The probability is estimated by the ratio presented in Equation 2-13, in which $N_{(Y=y)}$ represents the number of observations that match the evidence, and N is the number of

observations generated. A basic rejections sampling algorithm, such as the Metropolis-Hasting, is presented in **Algorithm 2-1**.

$$P(Y) = \frac{N_{(Y=y)}}{N} \tag{2-13}$$

Algorithm 2-1.: The Metropolis – Hasting algorithm.

Procedure: The Metropolis–Hasting Algorithm (MHA) is a particular implementation of the Markov Chain Monte Carlo Methods (MCMC). The MHA uses a proposed distribution (usually a $normal(0,\sigma)$) to draw samples from the posterior distribution $P(m|d)$. The algorithm starts at any arbitrary point of $P(m)$ with a non-zero value m^0. Then, it proceeds to a new position by proposing a movement based on the proposed distribution. The movement is accepted if a decision procedure is fulfilled.

0. Initialize the $\boldsymbol{m}$ values $(\boldsymbol{m^0})$. $\boldsymbol{m^0}$ values are sampled from the prior distribution $p(\boldsymbol{m})$.
   For iteration $i = 1, 2, \dots, n$
1. Generate a proposed jump sample value from the proposed distribution:
   $m^{prop} \sim r\left(m^i \middle| m^{i-1}\right)$
   If Gaussian distribution is used as a proposed distribution use
   $m^{prop} = m^{i-1} + \Delta m.$
   Where $\Delta m \sim normal(0,\sigma)$
2. Calculate the acceptance probability using the acceptance function (see equation 9):
   $\alpha = min\left\{1, \frac{q(m^{prop}|\, d)}{q(m^{i-1}|\, d)}\right\}$
   Draw a value (u) from an independently uniform (0,1) distribution.
3. Accept the proposal: if $u < \alpha$, then $m^{prop} = m^i$
   Reject the proposal: if $u > \alpha$, then $m^{prop} = m^{i-1}$
   Return m^{prop}

The Gibbs sampling is another approximate algorithm based on the rejection sampling technique. The basic idea is to explore the parameter space of the conditional probability of interest by changing one input parameter at a time. With this change, greater efficiency is achieved. **Algorithm 2-2** describes the general Gibbs sampling algorithm proposed by Yildirim (2012).

Algorithm 2-2.: The Gibbs sampling algorithm.

Procedure: Initialize all the parameters with initial values $m_1^0, m_2^0, m_3^0 \dots m_k^0$. At iteration $i = 1$, a value for m_1^1 is randomly drawn and conditioned on $m_2^0, m_3^0 \dots m_k^0$ values. Then, a value for m_2^1 at iteration $i = 1$, is randomly drawn conditioned on $m_1^1, m_3^0, m_4^0 \dots m_k^0$ values. In general, for m_k^1 at iteration $i = 1$, is drawn conditioned on $m_1^i, m_2^i, m_3^i, \dots, m_{k-1}^i$.

0. Initialize the $\boldsymbol{m}$ values $(\boldsymbol{m^0})$.

For iteration $i = 1, 2, \dots, n$

1. Sample $m_1^i \sim P\left(m_1^i | m_2^{(i-1)}, m_3^{(i-1)}, \dots, m_k^{(i-1)}\right)$
2. Sample $m_2^i \sim P\left(m_2^i | m_1^i, m_3^{(i-1)}, \dots, m_k^{(i-1)}\right)$
3. Sample $m_3^i \sim P\left(m_3^i | m_1^i, m_2^i, \dots, m_k^{(i-1)}\right)$
4. Sample $m_k^i \sim P\left(m_k^i | m_1^i, m_2^i, m_3^i, \dots, m_{k-1}^i\right)$

2.2.10 Most Probable Explanation

As mentioned in Section 2.2.9, a BN can answer questions about the most probable states for unobserved variables, given some observed variables. These types of questions are known as *Most Probable Explanations* (MPEs). Suppose that $\mathcal{X}_O = \{X_{O1}, \dots, X_{Oi}\}$ and $\mathcal{X}_U = \{X_{U1}, \dots, X_{Uj}\}$ represent the subset of observed and unobserved nodes, respectively, included in a BN. An *explanation* for a particular state of observed nodes $\mathcal{X}_O = x_O$ is defined as a configuration of states for unobserved nodes $\mathcal{X}_U = x_U$ such that x_U is consistent with x_O (Campos et al., 2001). A BN can provide several explanations to $\mathcal{X}_O = x_O$. However, as described by Pearl (1988), the abductive process is interested in finding the best explanation, or in BN terms, obtaining the *maximum a posteriori probability* (MAP). In other words, given x_O and x_U, the MAP aims to find the configuration of x_U that maximizes $P(x_O|x_U)$. Equation 2-14 represents the MAP, where

$\dot{x}_U$ is the *Most Probable Explanation* (MPE) for x_O. In general, a BN can provide the K-most probable explanations (K MPE) to an observed subset of variables x_O.

$$\dot{x}_U = argmax\ P(x_O|x_U) \tag{2-14}$$

2.3 Bayesian Networks in Civil Engineering Applications

Bayesian Networks in civil engineering applications are a new field of research. Although technical literature about this topic is limited, it has grown during the last few years. A search for publications in scientific databases yields the results presented in **Figure 2-16**. Significant growth has occurred since 2005, with 232 publications by 2021. The most relevant work is presented and discussed below.

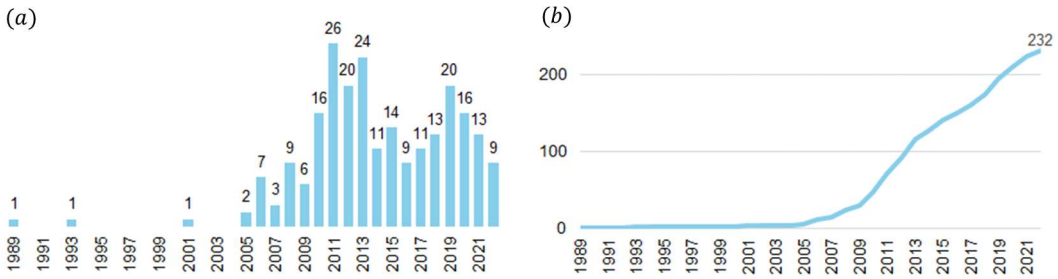


Figure 2-16.: Publications about Bayesian Networks in Civil Engineering (a) Number of publications by year. (b) Cumulative publications by year

Straub (2005) presented one of the earliest applications of BNs to civil engineering. The author proposed a general BN framework for analyzing natural hazards and demonstrating BN's viability for rockfall hazard analysis. Furthermore, he demonstrated that the flexibility of BNs improves risk assessment outcomes when interdisciplinary natural hazards share several parameters. In the same way, Smith (2006) used BNs to estimate the risk of multiple failure mechanisms in dams and the interrelation between them. For example, a simple dam failure analysis showed that although internal erosion and overtopping failure mechanisms follow different failure paths, they share some critical variables. Identifying these variables and their relative influence on failure mechanisms for programming maintenance and interventions is a necessary task.

Straub & Grêt-Regamey (2006) and Kiureghian et al. (2009) presented two applications of BNs in natural hazard analysis. The first study proposes a probabilistic framework for avalanche analysis. The framework uses BNs to combine information from diverse sources, such as dynamic and empirical models based on experience and observation. The results allowed researchers to create maps containing avalanches' annual probability and travel distances. They conclude that BNs improve avalanche prediction due to the integration of diverse information. The second study deals with infrastructure risk management in post-earthquake events. The study proposes a BN-based methodology for uncertainty management of post-earthquake events based on the ability of BNs to update their status as new information is obtained. This ability is used in making critical decisions in post-earthquake events such as road closures and the evacuation of people. Finally, the authors highlight the importance of real-time decision-making based on the information provided by BNs.

Nadim & Liu (2013) analyzed the quantitative risk of earthquake-triggered landslides and their effects on exposed buildings. Traditionally, earthquakes and landslides are studied separately and then combined using deterministic methodologies that usually do not include the assessment of the joint probability of the events. The authors propose using a BN to evaluate the combined effect of landslide-earthquake interaction from a multilevel perspective. Although the study demonstrates the ability of BNs to assess the joint probability of natural hazards, the need to include expert opinion and actual observations of earthquake-triggered landslides is recognized.

Probabilistic Bayesian network models have also been used to study dam safety. Morales-Nápoles et al. (2014) propose the use of nonparametric continuous variables to analyze dam safety, including three failure modes (global stability, overturning, and internal erosion) and three triggering factors (rainfall, earthquake, and maintenance type). Additionally, the BN model includes nodes for modeling consequences such as floods, loss of life, and economic costs. The use of expert opinion for eliciting conditional probability tables (CPT) is the most remarkable aspect of this study.

The phenomenon of soil liquefaction and its prediction during an earthquake has also been studied using BNs. This phenomenon holds significance in geotechnics, given the substantial uncertainty associated with soil and earthquake variables. Hu et al. (2016) constructed a BN

with twelve of the most critical variables in liquefaction analysis and found that BNs better predict the liquefaction behavior of soils than other artificial intelligence tools. The authors conclude that BNs predict better because their structure can represent the interdependencies between variables. Moreover, BNs combine information from multiple sources using a probabilistic framework.

Uncertainty monitoring and geotechnical reliability assessment in slope stability have also been studied using BNs. In this regard, Mohan et al. (2019) performed a study in which the influence of additional exploration on the reliability of slope stability was analyzed using geo-mechanical and spatial correlation variables. The BNs were used as a metamodel to shorten the calculation time and replace computationally expensive numerical models. A similar study developed by Špačková & Straub (2011) analyzed the performance of tunnel stability using Dynamic Bayesian Networks. The authors recognized few advances in fields such as stability analysis, advance rates, and excavation processes for tunnel design and construction. A framework based on dynamic BN was proposed based on historical information on excavation performance. The authors concluded that conditional probabilities estimated from past information are more reliable than probability estimations suggested by experts.

2.4 Summary

The preceding chapter briefly introduces Bayesian statistics, inference, and Bayesian Networks. These probabilistic concepts are the foundation on which the proposed methodology is based. The main ideas of this chapter can be summarized as follows:

- Bayesian statistics defines probability as a rational degree of belief. In other words, the natural world has no “fixed” statistical parameters but several probable states. Bayesian statistics mimics the “*explaining away*” process that simulates the bidirectional pattern (i.e., backward and forward processes) of human reasoning.

- The Bayes’ theorem (or Bayes’ rule) is the central concept of Bayesian Statistics. The theorem is a precise mathematical technique for analyzing information based on experience and knowledge.

- Bayes theorem can be used to compare competing hypotheses that explain a phenomenon (or the causes of a geotechnical failure). The comparison is performed through the Bayes Factor or Posterior Odds Ratio (POR). The POR represents how better a hypothesis H_i explains a phenomenon than a hypothesis H_j. Jeffreys scale and Kass & Raftery modified scale are used to interpret POR results.

- A Bayesian network (BN) is a model consisting of a DAG and a set of probability relationships. The DAG shows the causal relationships between variables (nodes), and the probability relationships are expressed as conditional probability distributions (CPD) or conditional probability tables (CPT). Common geotechnical knowledge and geotechnical jargon can be used to identify causal relationships between variables. These cause-effect relationships are translated into several small DAGs that can be joined together.

- The primary goal of any BN is to answer probability questions. In particular, questions related to the condition of a variable after the inclusion of some evidence in other variables. This process is known as inference, and BNs are used as powerful probabilistic tools because they can emulate inference processes.

3.Forensic Science and Forensic Engineering

3.1 Introduction and definitions

Forensic science aims to provide a logical connection between past events in order to discover causal relationships between them. Many authors have defined forensic science as a historical science based on the analysis of information (Carper, 2000; Cleland, 2001; Houck, 2006; Noon, 2009). Information in forensic science, also known as evidence, is usually sparse, scattered, vague, and incomplete (Noon, 2009). Pieces of evidence are the components from which an event is reconstructed. The result of this reconstruction is a narrative in which: (i) the event is described, (ii) a reasonably logical sequence of events is presented, and (iii) the most reasonable causal relationships between the different pieces of evidence are analyzed.

Since forensic science has traditionally studied events related to threats to life, health, or safety, it has been associated with criminal investigations. In fact, forensic science is aided by various specialties such as ballistics, DNA analysis, anthropometry, and fingerprint analysis, among others. Surprisingly, DNA analysis is the only one of these techniques that gives a probabilistic framework to traditional forensic science (Noon, 2009). The word *forensic* has a broader connotation, despite the common association of forensic science with criminalistics, disputes, or offenses. The definition of *forensic* is related to the context in which it is used. For example, digital forensics involves recording and analyzing digital information to support criminal investigations. Similarly, it is called forensic engineering when the purpose is to investigate the origin of failures in engineering systems. In this sense, the word forensics can be associated with any field of knowledge in which it is necessary to investigate cause-effect relationships.

In the case of forensic engineering, it can be defined as the branch of forensic science in charge of studying failures in engineering systems. Forensic engineering has two main focuses. First,

it can support the resolution of legal lawsuits associated with engineering failures (Carper, 2000). Second, it can investigate the origin or causes of system failures to propose technical solutions and improve engineering practices. It is important to note that the failure of an engineered system can be understood as a sudden collapse, system degradation, accident, incident, or any deviation from the behavior foreseen in the design.

3.2 Characteristics for the development of a forensic engineering study

Carper (2000) presents an account of the main characteristics of forensic engineering and the responsibilities of the professionals involved in its practice. The first characteristic refers to the professionals' technical competence and extensive experience in designing similar engineering works. Although this background is desirable and necessary in most cases, Babu (2016) and Brady (2012) argue that experience is not enough and may not be desirable because the design and forensic processes are opposite. In other words, design requires inductive skills to predict the behavior of a system, while forensic research requires deductive and retrospective skills to explain how and why a failure occurred (Day, 2010). The main differences between design and forensic processes are shown in **Table 3-1**.

A second desirable characteristic in a forensic investigation is the competence of the professionals involved in the forensic study. It refers to the ability to collect, organize, select, and interpret evidence. Babu (2016) points out that the collection of evidence is critical due to the need to clean the failure site and reconstruct the engineering works. In most cases, the evidence is destroyed, washed, lost, or concealed. Therefore, the interpretation of evidence must almost always be made on incomplete information. In other cases, conclusions can be drawn, or conflicts can be resolved with complementary and redundant information (Carper, 2000).

Table 3-1.: Comparison of design and forensic processes. (After Brady, 2012)

Design Process	**Forensic process**
Problem-solving	Causation determination
It is based on assumptions	It is based on evidence

Design Process	**Forensic process**
Deals with unknowns to give engineering solutions	Investigates the unknowns of an engineering solution
A priori (starts with theories and ends with evidence)	A Posteriori (starts with the evidence and ends with theories)
Synthesizes (has multiple solutions, chooses one)	Analyzes (only one sequence of events)
Discards the less probable hypothesis	Discards hypotheses based on evidence
The heuristic method is used (i.e., follows standards, strategies, rules, or codes)	No heuristic method is used (Based on evidence and scientific principles)

The third and last characteristic refers to the ability to communicate the results of a forensic investigation. Communication ability includes both oral and written skills. Verbal ability is vital when the investigation results are used to support legal litigation, while written ability must reflect the competence and truthfulness of the forensic investigation process (Carper, 2000).

3.3 Use of the Scientific Method

The scientific method is a systematic approach that aims to reveal the cause-effect relationships of a phenomenon through the collection, organization, and analysis of evidence or factual data (Noon, 2009). Lord Bacon formally introduced the scientific method in the early seventeenth century, and since then, its application has been widely extended to the physical sciences. Two versions of the scientific method are recognized. The first and oldest corresponded to the application of the original version of the method. This version included the following stages:

1. Proposing a hypothesis based on data, measurements, facts, or observations.

2. Contrasting the data with the hypothesis to verify its consistency.

3. Verifying or modifying the initial hypothesis with new data or available evidence.

The second and more recent version of the scientific method includes the following stages:

1. Proposing a hypothesis based on the collection, organization, and analysis of data, evidence, measurements, and observations.

2. Contrasting the data with the proposed hypothesis. This hypothesis must also be able to predict unobserved data or anticipate consequences or effects not yet seen.

3. Confirming, modifying, or changing the hypothesis as required if new evidence or experimental data becomes available.

The critical difference between the two versions of the scientific method is the predictive ability of the hypothesis of the second version. Consequently, further experiments or additional measurements could verify the predictive ability. For example, the predictive ability of a hypothesis can forecast consequences, information, and data found later that could have been dismissed, misinterpreted or considered unimportant in the early stages of the forensic investigation (Noon, 2009).

The third phase in both versions of the scientific method confirms that hypotheses are not static premises but, on the contrary, evolve as new evidence is acquired. This evolutive process implies that the third phase is an iterative process in which the hypothesis adjusts according to the evidence. For example, after a preliminary evaluation, a forensic investigator proposes that an unforeseen surcharge caused an excavation failure (initial hypothesis). Based on this initial hypothesis, an engineering solution is proposed, a legal investigation process is initiated, and preliminary conclusions are communicated to the media. Sometime later, after a more exhaustive investigation, new evidence related to an elevation of the water table and additional lateral pressures on the excavation contradicts the initial hypothesis. The new hypothesis states that the excavation failure was caused by additional hydrostatic pressures on the excavation not included in the design. As a result, the engineering solution is substantially modified, the legal process could take a different direction, and public opinion could be negatively impacted.

3.4 Inductive, Deductive, and Abductive Processes

The reasoning process used in forensic engineering can be framed within three types: inductive, deductive, and abductive. Inductive reasoning dates back to ancient Greece and

attempts to infer a general conclusion from particular cases. In other words, a proposition or set of propositions is tested against evidence obtained through unbiased observations (Noon, 2009). Once the proposition is evaluated, a commonality is proposed. Thus, the inductive process starts from particular cases and ends up with the statement of a universal premise. Since the universal premise is formulated from particular cases, inductive reasoning admits a reasonable degree of uncertainty that can be addressed using probability theory.

Deductive reasoning can also be traced back to ancient Greece and was recognized as the preferred method of reasoning by the Greeks (Noon, 2009). The deductive process relies entirely on the intellect to explain the phenomena. That is to say, conclusions on particular aspects are deduced from a set of premises universally accepted as truths. Contrary to inductive reasoning, deductive reasoning begins from general aspects and ends with particular statements. Since the general premises are assumed to be true, deductive reasoning does not involve uncertainty and therefore does not admit probability analysis.

Abductive reasoning is the third type of logical reasoning. Contrary to inductive or deductive reasoning, which use particular cases or universal premises, abductive reasoning is based on statements known as hypotheses. Hypotheses are suppositions or possible explanations of a phenomenon generated on the basis of evidence, observation, and measurements. Campos et al. (2001) represent the abductive reasoning process by Equation 3-1. This equation explains that if some phenomenon ω is observed and some norm or rule φ leads to ω, then φ is defined as a possible explanation for ω. The nature of abductive reasoning implies a component of uncertainty associated with its process, i.e., since φ is defined as a possible explanation of the phenomenon, it is assumed that φ is not unique, and several rules could explain the phenomenon ω. Consequently, each possible explanation φ associated with the phenomenon ω has a probability of occurrence.

$$\frac{\varphi \rightarrow \omega, \omega}{\varphi} \tag{3-1}$$

Several authors, such as Campos et al. (2001), Jensen & Nielsen (2007), and Fenton & Neil (2019), argue that abductive reasoning actually mimics the human reasoning and decision-

making process. **Figure 3-1**, proposed by Campos et al. (2001), presents the process of abductive reasoning. The process unfolds as follows: (i) facts, observations, and measurements are employed to (ii) generate several hypotheses and (iii) develop the set of possible explanations of the phenomenon. After contrasting the facts and hypotheses (iv), one or a few are selected based on their capacity to explain the phenomenon. (v) This selection constitutes the best explanation of the phenomenon. Bayesian networks can systematically represent the abductive reasoning process (refer to Chapter 2).

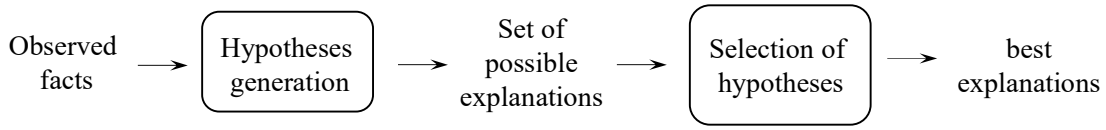


Figure 3-1.: Abductive process (modified from Campos et al., 2001)

3.5 A review of some general methods used in forensic engineering

Although standardization does not apply to forensic methods, some specific characteristics are common to all forensic engineering investigation processes. These characteristics are linked to applying the scientific method and are reproducible in structural and geotechnical engineering. The forensic investigation methods presented below are general and applicable to any engineering failure. However, since the most studied failures have structural origins, their application focuses on structural engineering.

Noon (2009) defined the forensic process using the "investigation pyramid" idea. This idea compares the forensic process to a pyramid with a large amount of information (i.e., evidence, data, and facts) at its foundation. On top of this foundation, information is analyzed using the scientific method. Finally, at the top of the pyramid, the analyzed information supports a small number of conclusions. The pyramid idea suggests that conclusions must be based on the analysis of the evidence and not on other hypotheses or conclusions. **Figure 3-2** depicts the concept of the "investigation pyramid."

Based on a set of steps common to all forensic investigations, Bell (2000) proposed the flow chart of the investigative process presented in **Figure 3-3**. One implication of Bell's diagram is that several failure hypotheses cannot be re-evaluated simultaneously using the same

evidence. In fact, new failure hypotheses are not allowed in the latter stages. Therefore, even though new evidence may disprove the initial failure hypotheses, the process prevents the formulation of new hypotheses.

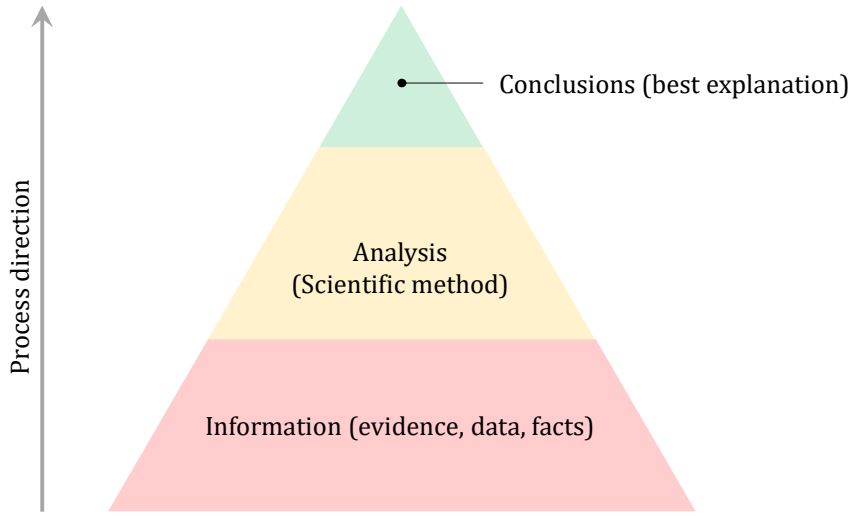


Figure 3-2.: The Investigation Pyramid (Modified from Noon, 2009).

In 2018, the American Society of Civil Engineers (ASCE) published a set of guidelines for failure investigations. These guidelines (ASCE, 2018) update a first edition published in 1989. The ASCE recognizes that each investigation is unique and that there is no single approach to studying failures. However, the ASCE outlines five steps common to all forensic engineering studies. The steps shown in **Figure 3-4** are based on forensic investigations focused on civil engineering. The Bayesian framework for forensic geotechnical assessments described in Chapter 4 is based on the recommendations suggested in the ASCE (2018) guidelines.

A forensic engineering research methodology based on civil, biomechanical, or aerospace engineering practices was proposed by Terwel et al. (2018). According to these authors, the methodology integrates three essential elements identified in every forensic investigation: (i) the life cycle of the product or structure, (ii) the categorization of the failure causes according to the life cycle stage, and (iii) the standard forensic investigation approach that includes steps such as data collection, hypothesis generation, hypothesis testing, conclusions, and report findings. In addition, the authors implemented the "ring of trustworthiness" concept to increase the validity and reliability of forensic findings.

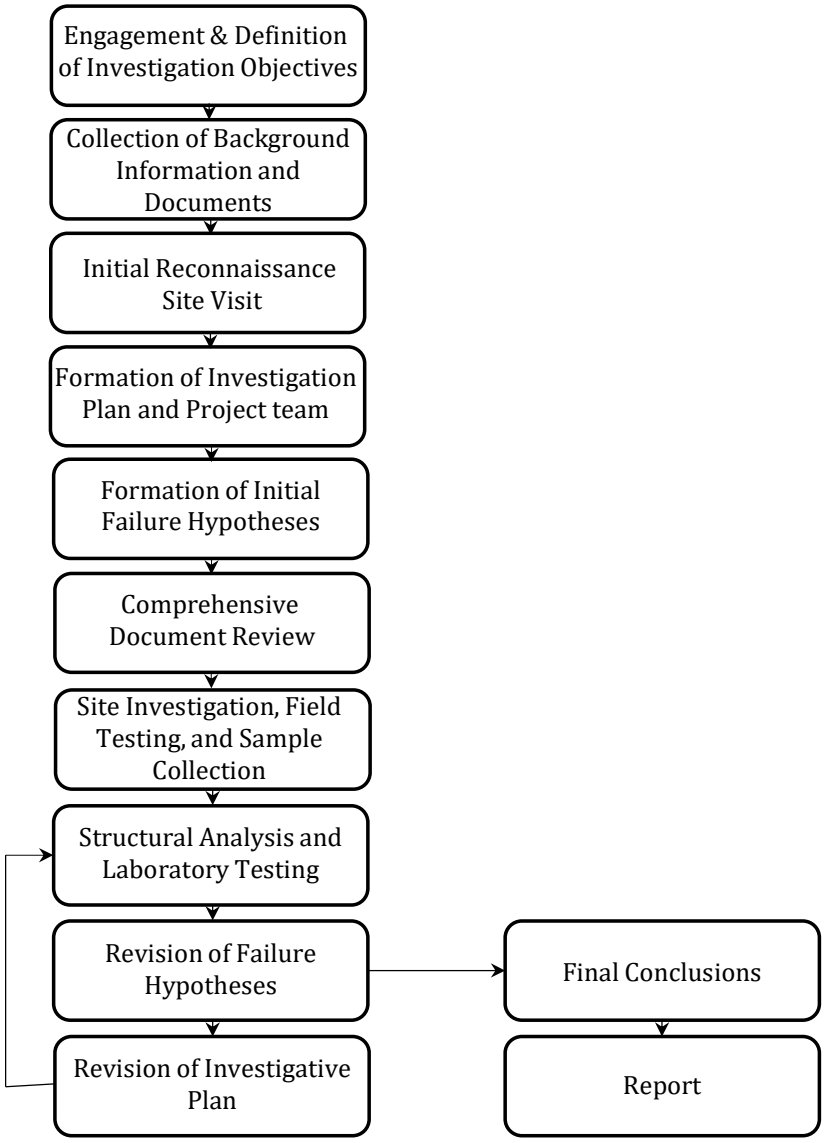


Figure 3-3.: The investigative process proposed by Bell (2000)

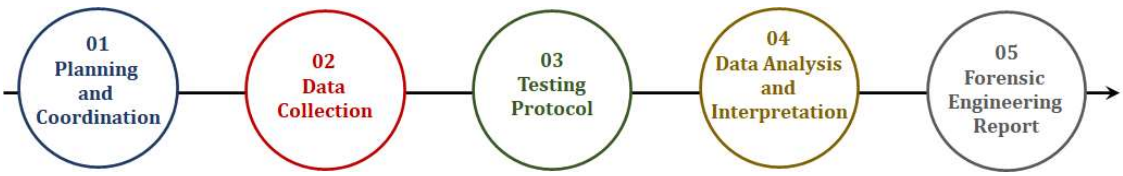


Figure 3-4.: Typical steps of a forensic investigation suggested in ASCE (2018)

3.6 Forensic Geotechnical Engineering

Forensic geotechnical engineering is a sub-branch of forensic engineering concerned with studying and investigating engineering system failures associated with geological/geotechnical origins (Rao, 2016). The traditional practice of forensic geotechnical engineering has focused primarily on resolving legal disputes, supporting legal decisions, and determining liability (Carper, 2000; Day, 2010; Lacasse, 2016). However, a more recent version of forensic geotechnical engineering has broadened its focus to investigate causal relationships leading to geotechnical failures. This latter version focuses primarily on safety and *learning from mistakes* to prevent and improve geotechnical engineering systems (Terwel et al., 2018). In addition, it allows forensic geotechnical engineering to include the tools of the Scientific Method, the deductive and abductive reasoning methods, and the application of probability and statistics techniques.

3.6.1 Stages of the forensic geotechnical process

Brady (2012) and the ASCE guidelines (ASCE, 2018) recognize that the forensic geotechnical engineering process involves at least the four stages shown below:

1. Collection of available evidence and its analysis.

2. Development of credible hypotheses about the causes of geotechnical failure.

3. Testing each of the credible hypotheses against the available evidence.

4. Selection of the hypothesis related to the most likely cause of failure.

Brady (2012) also suggests a fifth step relating to identifying and effectively communicating the cause of failure. Similarly, Terwel et al. (2018) recognize that communication of failure causes should not be undervalued.

The first step involves collecting the available information related to the geotechnical failure. This information, referred to as evidence in forensic terminology, enables the identification of failure aspects and their associated consequences. Evidence collection begins with the field

investigation, which includes site reconnaissance, field observation, photographic recording, eyewitness interviews, and the collection of design and construction documents. The use of modern tools such as drone photography, satellite imaging, and laser scanning is recommended in any forensic geotechnical investigation. Although evidence collection may come from various and dissimilar sources of information, it should be selected and organized in easily accessible databases. Databases make information available to understand the conditions, circumstances, and factors leading to geotechnical failures. In all cases, the evidence collection process should include verification, preservation, and chain of custody of the information.

Stage two is concerned with establishing hypotheses related to failure causes. Bell (2000) and Brady (2012) agree that inductive and synthesis reasoning processes employed in conventional engineering design are insufficient for hypothesis generation. These conventional methods are not applicable because they require apriorism reasoning, i.e., starting with general theories and ending with evidence. In other words, multiple solutions to an engineering problem are proposed, and one is chosen based on experience, codes, or standard design methods. Therefore, apriorism reasoning is harmful to hypothesis generation because it can force the evidence to be matched with biased hypotheses. For example, a common bias in forensic investigation occurs when a possible failure hypothesis is identified, and forensic experts attempt to find evidence confirming that hypothesis. In such cases, forward reasoning is favored over backward reasoning, contrary to the requirements of forensic methods.

Noon (2009) proposes that hypothesis generation should not be based on traditional engineering methods but on deductive and abductive reasoning. Abductive reasoning is particularly interesting in forensic engineering because it allows information and evidence to guide the generation of failure hypotheses. In such cases, abductive reasoning identifies distinct components of a geotechnical structure and combines them to generate a logical sequence of the failure and its associated causes. Unlike traditional engineering design, where there are many options for solving a problem, in forensic geotechnical engineering, the failed structure behaves and fails in only one way. For this reason, in forensic engineering, there are no standardized processes for the investigation of a failure. Only the tools provided by the scientific method and abductive reasoning are used.

The third stage of the forensic investigation process involves developing and analyzing the failure hypothesis established in the second stage. The main task of this stage is to compare each hypothesis against the evidence collected. According to Bell (2000), all credible hypotheses must be systematically analyzed before being approved or disapproved. Each hypothesis must be continually tested against the evidence to be validated or falsified during the forensic investigation. The result of the process is the elimination of all but one hypothesis that explains the evidence. However, most forensic investigations do not reach individual results due to the uncertainty involved in the process. Usually, the conclusions reached refer to several causes of failure. In other cases, the conclusion refers to some of the most probable causes of the failure.

In a conventional forensic geotechnical investigation, the hypotheses are contrasted with the evidence using back-analysis calculations. Back-analysis consists of implementing either numerical or physical geotechnical models, in which the results are compared with the evidence collected during the first stage of the forensic process. In general, if the results of the back-analysis agree with the evidence, the hypothesis is considered feasible. However, if the results of the back-analysis do not agree with the evidence, the hypothesis is rejected (falsified) or considered unlikely. Although back-analysis is commonly employed to test failure hypotheses, Hwang (2016) points out that only experts can interpret its results. This author further points out that computational capacity limit back-analysis when complex constitutive models are used.

As mentioned above, the third stage is about testing hypotheses against the evidence and re-evaluating their validity, which implies that stage three is an iterative process. Some of the hypotheses initially put forward are discarded during the iterative process, and new ones are proposed. Iteration is performed until one (or a few) hypotheses are consistent with the evidence. This iterative process is one of the characteristics of the scientific method applied in forensic investigations, in which abductive reasoning and probabilistic tools offer support in determining the causes of geotechnical failures.

The last stage focuses on selecting the most probable cause (or causes) that led to the geotechnical failure. The cause is selected among the hypotheses formulated in stage two and

from the results of the iterative process of stage three. The selected hypothesis and the conclusions of the forensic investigations may support legal disputes and assist remediation designs. Several authors, such as Bell (2000), Poulos (2016), and Terwel et al. (2018), recognize the fourth stage as a critical task in the forensic investigation process. Moreover, authors such as Kool et al. (2019) acknowledge that the conclusions drawn from forensic studies sometimes appear arbitrary and subjective due in part to a lack of rigor in the process of hypothesizing, contrasting evidence, and selecting the most probable cause. Combining the scientific method, probability theory, and statistical tools may help reduce the subjectivity and arbitrariness of the conclusions about the causes of failure of geotechnical origin. This doctoral research uses Bayesian tools to support decision-making regarding the causes of geotechnical failures in order to reduce the cognitive bias associated with the forensic process.

3.6.2 Additional aspects to consider

Collection of geotechnical evidence

Collecting geotechnical evidence is the first step after a failure. Day (2011) warns that people involved are urged to clean up the site and reconstruct the collapsed structures quickly after a failure. This cleanup may limit the time available for evidence collection and cause forensic investigators to work against the clock. The information collected within this limited timeframe must be sufficient to develop a reliable forensic investigation. In addition, the parties involved in the investigative process must agree on the plausibility of the information. For this purpose, guidelines such as ASCE (2018) for collecting and storing information should be used.

Pre-failure signs

Before a failure, geotechnical structures show signs of a possible deviation from the behavior predicted during the design stage. Geotechnical monitoring, visual inspection by human experts, and maintenance are the sources of information from which these pre-failure signs can be identified. In geotechnical structures, pre-failure signals may include cracking, settlement, deformation, stress increase, and changes in the water table and pore water pressures. Information from these signs can be helpful during the forensic investigation process. It can give clues about the origin and early development of the failure.

Back Analysis

As mentioned in Section 3.6.1, back analysis is widely used in forensic geotechnical engineering. Its purpose is to validate the hypotheses proposed as causes of failure. Back analysis can consist of several levels of analysis (Hwang, 2016). The most basic analyses employ hand calculations, stability numbers, empirical relationships, or rules of thumb. Intermediate analyses employ numerical models with simple constitutive equations. More sophisticated analyses require complex geometries and more sophisticated constitutive models. Regardless of the complexity of the analyses, the interpretation of back analysis results should be viewed only as a decision support tool in which engineering judgment should prevail. For this reason, the interpretation should be made by engineers experienced in forensic and back analysis.

Technical shortcomings

Technical shortcomings can appear at any stage of an engineering project. In the design stage, errors related to the misapplication of standards and technical specifications, inadequate subsurface exploration, and poor assignment of geotechnical models and parameters can be expected. There may be substandard construction practices, inadequate monitoring and quality control, and an inadequate inspection and maintenance plan during the construction and maintenance stages (Babu, 2016).

In any stage of a geotechnical project, failures can occur due to human factors. Melchers & Beck (2018) classify human factors into (i) human error and (ii) human intervention. On the one hand, failures due to human errors are related to ignorance, carelessness, negligence, or insufficient knowledge. On the other hand, human intervention can act in both directions: sometimes leading to failure or sometimes minimizing human errors by applying positive actions. In this regard, Sowers (1993) evaluated 500 cases and found that 88% of civil engineering failures originated from human shortcomings. However, he also emphasized that continuous education and retraining could reduce the number of engineering failures and negative consequences. Similarly, Jessep et al. (2016) discuss the shortcomings that led to 100 geotechnical failures. They found that 50% of the failures were caused by inadequate design. The remaining 50% were caused by shortcomings in site investigation, unforeseen phenomena, and construction malpractice.

Observational Method

The Observational Method is a geotechnical engineering tool developed by K. Terzaghi and later established by Peck (1969). It involves designing a geotechnical structure based on the best knowledge of materials and considering a wide range of possible behaviors. The overall structural performance is monitored during construction based on measuring specific parameters. For each possible behavior, corrective actions are stipulated. A comprehensive review of the Observational Method combined with Bayesian statistics is described by Spross (2016), Spross et al. (2016), and Spross & Johansson (2017). Extending the observational method beyond the construction process can help detect the causes of failure in geotechnical structures. For example, Zhang et al. (2010) applied the observational method and some specific measurements to evaluate the safety of civil structures using Bayesian updating techniques. However, the advantages of the observational method in forensic geotechnical engineering have not been extensively documented.

Reliability considerations

Although the variability of geotechnical materials is sufficiently accepted, probability and reliability aspects are seldom employed in forensic geotechnical engineering evaluations. Phoon et al. (2016) highlight the lack of literature on this subject but pointed out the need and potential use of probability tools in forensic geotechnical engineering. Moreover, these authors argue that if geotechnical failure is defined as an unacceptable difference between expected and observed behavior (Leonards, 1982), a failure should be quantified in a probabilistic sense. In other words, probability tools could provide additional information about the causes of geotechnical failures and probabilistically estimate the difference between expected and actual behavior.

Communicating the causes of failure

As mentioned above, communicating the causes of failure is an essential step in the forensic investigation process. The engineer must present the findings of the forensic studies in a manner that experts and non-experts understand. Effective communication about the causes of failure plays a prominent role in courts and trials where liability is sought, but it must also improve engineering practice through case studies. In this regard, Bell (2000) emphasizes that a case study should be approached by considering two types of failure processes: technical and procedural. The former refers to analyzing the physical conditions and the interactions

between components that led to the failure. The latter refers to constructive deficiencies, design deficiencies, lack of quality, and human errors. According to Bell (2000), procedural errors could represent more than 90% of the causes of the failure of engineering systems.

Common errors in a forensic investigation:

Noon (2009) identified at least four (4) common errors in forensic investigations: (i) The first error refers to the hasty search for evidence favoring a hypothesis. This error is typical in forensic practice because the investigator anticipates conclusions before completing the forensic analysis. It is of particular interest in the later stages of the study when only two or three failure hypotheses remain under study. (ii) The second common mistake is related to possible conflicts that may arise from the conclusions of a forensic study. For example, the results of the forensic study could blame an individual or a company concerned about its reputation. In this case, forensic investigators must act unbiasedly and maintain professional integrity. (iii). The third error is caused by overconfidence in determining the causes of failure. In other words, the failure cause is presumed to be so evident that collecting evidence, formulating failure hypotheses, and analyzing those hypotheses is considered unnecessary. (iv) Finally, the fourth typical shortcoming in forensic investigation is a procedural error related to people involved in corrective actions. Although this error is outside the failure investigation itself, Noon (2009) points out that implementing such corrective actions cannot be left to those involved or responsible for the failure. In this case, from an ethical perspective, the solution requires assigning different personnel.

3.6.3 Advances and research in forensic geotechnical engineering: description and discussion

Civil engineering has traditionally employed forensic engineering to determine failure causes, especially failures associated with structural causes. In the forensic study of structures, the classic texts of Ratay (2000), Bell (2000), Noon (2001), Kardon (2003), Brady (2012), and ASCE (2018) are well-known. However, forensic science in geotechnical engineering was only officially recognized in 2006 with the creation of the ISSMGE Technical Commission TC 40 “Forensic Geotechnical Engineering” (Rao & Babu, 2009). The creation of this technical committee does not imply a total absence of forensic geotechnical analyses in previous decades, rather studies have tended to focus on particular cases. For example, the Vaiont landslide in Italy (Müller L., 1968; Chowdhury, 1987; Dykes & Bromhead, 2018), the Leaning

Tower of Pisa (Burland et al., 1998) and the settlement of the Metropolitan Cathedral in Mexico City (Guerra, 1992) are well-known forensic geotechnical cases.

Geotechnical engineering has used forensic science deterministically and sometimes without the rigorous application of the scientific method. This practice has negatively impacted forensic geotechnical studies because conclusions about the causes of failure sometimes seem arbitrary or biased (Kool et al., 2019). Several authors have warned about this malpractice and have expressed the need to implement new approaches to studying forensic geotechnical engineering (Phoon et al., 2016; Gilbert, 2016; Xu & Zhang, 2016). The following paragraphs discuss some of the most relevant works in forensic geotechnical engineering developed in the last few years.

The Vaiont landslide is perhaps one of the first well-documented cases of forensic geotechnical engineering. The landslide occurred in October 1963 due to the reactivation of an old landslide and the rising water levels in the reservoir (Alonso et al., 2010). Müller (1968) reviewed the primary studies of the Vaiont catastrophe up to 1968 and developed a fault model that disproves the failure hypothesis related to the existence of clay on the slip surface. The study criticized the use of *a posteriori* analyses performed by other authors in which critical aspects of the limit equilibrium calculation were not considered. He further explains that such calculations cannot account for certain aspects of failure due to reservoir level variation. Müller's approach to the Vaiont failure has some characteristics of a forensic study. For example, he developed a model of failure and tested some hypotheses. He then compared the results of the failure model with the evidence and concluded on the hypotheses. Although Müller's work was not strictly a forensic study, it was the first attempt to test hypotheses against the evidence from a geotechnical perspective.

Day (2010) presented one of the first forensic engineering textbooks devoted exclusively to geotechnical engineering, especially to the forensic study of foundations. Day provided an overview of the practice of forensic foundation engineering in the late 20th century. Extensive descriptions of the process of evidence collection through field testing, laboratory testing, monitoring, and documentation acquisition are presented. Although the first chapters are focused on explaining the process of assignment and investigation based on the scientific method, Day presented an investigative process based on an ASCE guideline (Greenspan et al.,

1989) that does not include hypotheses testing against evidence (**Figure 3-5**). Moreover, the examples presented throughout the book do not reveal how each failure hypothesis was tested against geotechnical evidence. Although the text attempts to make the forensic geotechnical investigation more rigorous by formalizing some of the processes, it still lacks two essential steps of the scientific method: hypothesis development and verification (Falsification principle. Popper, 2002).

As was mentioned before, by the beginning of the 21st century, the international geotechnical community had identified the need for a more rigorous approach to forensic geotechnics. Thus, in 2006 the ISSMGE created the technical committee TC 40, "Forensic Geotechnical Engineering," which later became the TC 302. The purpose of this committee was to prepare guidelines for forensic studies, present to the geotechnical community learned lessons, and interact with other disciplines. Among TC 302's functions were the realization of several working sessions, dissemination events, and the publication of the book "Forensic Geotechnical Engineering" (Rao & Babu, 2016). The publication of this book formalized the practice of forensic geotechnical engineering and presented an overview of how researchers were approaching the study of geotechnical failures.

In addition to the above, Poulos (2016), Lacasse (2016), and Rao (2016) contributed a set of articles that proposed a general framework and guidelines for the practice of forensic geotechnical engineering. Poulos focuses his framework on discussing the geotechnical and structural factors affecting forensic investigations. He also suggests developing and testing credible hypotheses by comparing them with the field, laboratory, and computational model results. In addition, The author warns of the need to consider the variability of the terrain, thus opening the possibility of using probabilistic tools. Finally, he draws attention to the iterative process of hypothesis verification.

On the other hand, Lacasse emphasizes the importance of forensic engineering in resolving legal conflicts and highlights the principles of its practice. Such principles are related to the "standard of care," expert evidence, and the litigation process. In other words, it refers to scientific and conflict resolution skills required by professionals involved in forensic investigations. Rao provides an additional viewpoint on the characteristics of forensic investigations. He offers an overview of the general procedures used in forensic analysis and

points out that traditional sampling, testing, and design techniques do not fit the requirements of forensic work. He further emphasizes the importance of legal issues and the relevance of final report writing.

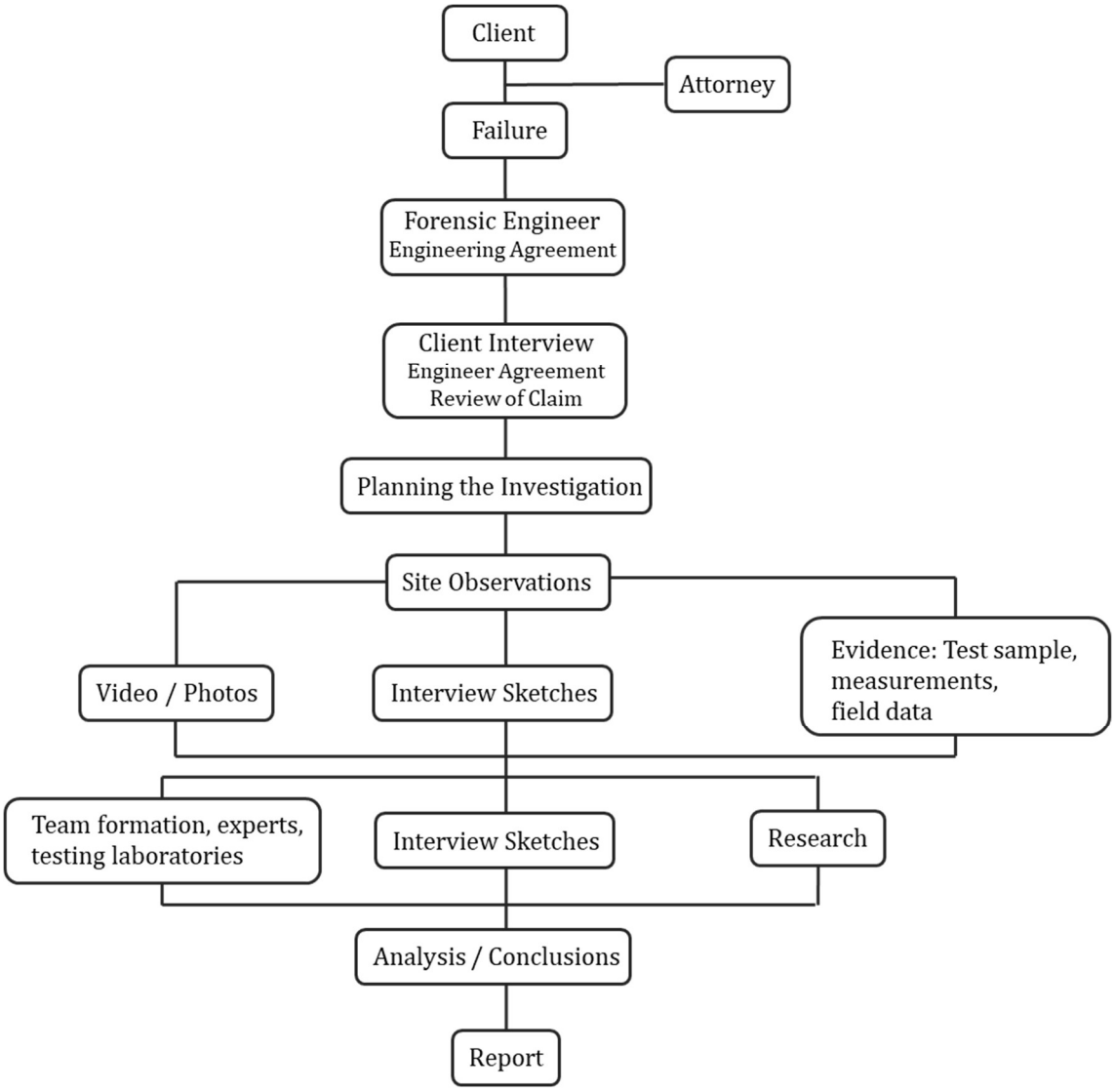


Figure 3-5.: Typical steps in forensic investigations. Modified from Day (2010) and Greenspan et al. (1989).

The use of back analysis has been and continues to be the most widely employed tool in verifying failure hypotheses. For example, Popescu & Schaefer (2016) employed a back analysis to compare strength parameters before and after the installation of piles for landslide stabilization. Hwang (2016) noted that back analysis is employed as technical evidence to

validate hypotheses. His study provides an example of back analysis, the information required, and its limitations. The limitations include the complexity of some constitutive models and the need for the results to be interpreted by expert engineers. Other notable research papers in which back analysis was used for hypothesis validation include (i) Alonso et al. (2016), who described the failure of a Caisson produced by storm waves, (ii) a study by Iai (2016) in which the failure of a Caisson quay wall due to an earthquake is described, and (iii) the forensic analysis of the failure of a retaining wall presented by Babu et al. (2016)

The most recent proposals on the use of probabilistic tools in forensic analyses are presented by Gilbert (2016) and Phoon et al. (2016). Gilbert (2016) shows that uncertainty plays a crucial role in forensic analysis. He demonstrates from a probabilistic approach that if several hypotheses can explain the evidence, it is almost impossible to determine the causes of failure. He also argues that collecting additional evidence does not always mean that the uncertainty of failure hypotheses is reduced. Furthermore, the author claims that ignoring credible hypotheses can lead to erroneous conclusions.

Similarly, Phoon et al. (2016) emphasize the need to include reliability analysis in forensic geotechnical engineering to provide objective conclusions. To this end, they propose a preliminary methodology consisting of two components: (i) a reliability index describing the behavior of the geotechnical system and (ii) a conventional statistical hypothesis test to check if the reliability index was met. Although Phoon's method provides innovation in forensic evaluation, it continues to employ the classic null hypothesis significance testing (NHST) that Bayesian methods have strongly criticized (Allenby, 1990; Johnson, 1999; Kruschke, 2010; Masson, 2011; Szucs & Ioannidis, 2017; Tendeiro & Kiers, 2019). The methodology proposed in this doctoral research is based on Phoon's proposal to include probability and reliability tools for forensic geotechnical evaluations. However, it goes further by incorporating Bayesian analysis tools to overcome the drawbacks of the classical probability approach.

The Breitenhagen levee failure is a recent and well-documented forensic geotechnical case that focuses on the breach of a section of the Saale River levee near Breitenhagen in Germany in 2013. According to Grubert (2013), an instability process apparently caused the failure due to the river's sustained rise in water levels. Kool et al. (2019) developed a forensic study of this levee failure using a systematic approach that included: (i) estimating a range of values for the

geotechnical parameters, (ii) establishing a baseline stability model using the expected values for the parameters, (iii) defining possible failure scenarios and including the uncertainty in the parameters, (iv) determining the most likely failure scenario using sensitivity calculations. Kool et al. proposed ten failure scenarios and estimated the factor of safety (FoS) and the geometry of the slip surface for each of them. The FoS and failure geometry values for each scenario were compared against the available evidence. They concluded that the combination of high river levels and locally weak soils caused the failure of the Breitenhagen levee.

In a later study, Kool et al. (2020) reassessed the Breitenhagen levee failure using a Bayesian probabilistic approach. As in the 2019 study, a base model and several failure scenarios were proposed. For each scenario, a probabilistic model was developed, including pore water pressures, method of analysis, and soil behavior model. The likelihood of each scenario was estimated by calculating its probability of failure. Subsequently, Bayes' theorem was used to calculate the probability of each scenario given the failure and slip geometry. Unlike the study by Kool et al. (2019), the probabilistic results showed that the most likely cause of failure was the combination of locally weak soils and high pore pressures inside the levee due to an aquifer.

Garcia-Feria et al. (2022) revisited the data of the Breitenhagen levee failure using a probabilistic model approach based on Bayesian Networks. Hypotheses and evidence nodes were included in the Bayesian networks in order to test each failure hypothesis against the collected evidence. Probability queries and the K-Most Probable Explanation (KMPE) algorithm were used to find the cause of failure. Unlike Kool et al. (2020), the Bayesian network approach concluded that a combination of weak soils and high phreatic levels led to the levee failure.

3.7 Summary

The main objective of this chapter has been to present a brief introduction to forensic science and forensic engineering and to illustrate how these concepts are used in forensic geotechnical engineering (FGE). Forensic science is defined as a historical science based on the analysis of information. Any forensic analysis should: (i) describe the event (phenomenon or failure), (ii) present the logical sequence of events, and (iii) analyze the causal relationships between pieces of evidence. In addition, the main ideas can be summarized as follows:

- Forensic engineering (FE) is the branch of forensic science in charge of studying failures in engineering systems. It has two main focuses: (i) support the resolution of legal lawsuits associated with engineering failures, and (ii) study the origin or causes of system failures to propose technical solutions and improve engineering practices.

- The scientific method is a systematic approach that aims to reveal the cause-effect relationships of a phenomenon through the collection, organization, and analysis of evidence or factual data. The scientific method is a central concept in forensic engineering. It includes the following stages: (i) hypotheses formulation, (ii) contrasting evidence with the proposed hypotheses, and (iii) confirming, modifying, or changing hypotheses when new evidence is available.

- Three reasoning processes support forensic engineering: inductive, deductive, and abductive. In FE, abductive reasoning is preferred because it is based on hypotheses. Hypotheses are explanations of a phenomenon generated on the basis of evidence, observation, and measurements. The nature of abductive reasoning implies a component of uncertainty because a hypothesis is a possible explanation of a phenomenon (failure).

- A forensic geotechnical investigation includes at least four steps: (i) Collection and analysis of available evidence, (ii) formulation of credible hypotheses about the causes of failure, (iii) testing hypotheses against the available evidence, and (iv) selection of one hypothesis related to the most likely cause of failure.

- Due to the unique characteristics of geotechnical engineering, some aspects of forensic geotechnical analysis need special attention. For example, although back analysis has been widely used in FGE, its results should be viewed only as a decision support tool in which engineering judgment should prevail. The methodology proposed in this thesis uses back analysis as a support tool, but in addition, it includes engineering judgment through probabilistic techniques such as POR and BN.

4. Bayesian Methodology for Decision Support in Forensic Geotechnical Engineering

This chapter describes the proposed Bayesian methodology for decision support in forensic geotechnical engineering (FGE). The methodology focuses on decision-making about the most probable causes of geotechnical failures by formulating multiple hypotheses about the conditions that may have led to the failures. Then, the hypotheses are evaluated probabilistically using two techniques: (i) Bayesian hypotheses comparison via posterior odds ratio and (ii) Bayesian networks (BN). The result is the selection of one (or several) hypotheses as the most probable causes of failure.

The first section introduces the methodology and describes the proposed steps for collecting evidence, formulating hypotheses, constructing the probability model, and comparing hypotheses. The methodology includes the elements presented in Chapters 2 and 3 and some additional elements proposed by the author of this thesis based on posterior odds ratio and Bayesian networks techniques. The second section presents a benchmark example (Schweiger, 2006) formulated from a well-known geotechnical problem to describe and validate the methodology for decision support using the posterior odds ratio technique and Bayesian networks. The third section uses a well-documented levee failure analysis (Kool et al., 2019) to apply the proposed methodology.

4.1 Proposed Bayesian Methodology for Decision Support

The proposed Bayesian methodology for decision support in FGE consists of three main stages. Each stage includes several steps focused on providing exhaustive information regarding evidence, failure models, and hypotheses comparison. The stages and their steps follow the principles of forensic engineering methodologies suggested in the past (Bell, 2000; Noon, 2009; Brady, 2012; Poulos, 2016; ASCE, 2018; Terwel et al., 2018). **Figure 4-1** depicts a

flowchart of the proposed methodology. The main stages and their steps are described in detail in the following subsections.

4.1.1 Stage 1. Preliminary steps

Preliminary steps include: (i) collecting evidence and (ii) formulating failure hypotheses about the causes of failure.

Collecting evidence

The purpose of collecting evidence is to validate failure hypotheses. The evidence and their analysis should be able to validate or disapprove hypotheses about the causes of failure with geotechnical origins. For example, if a hypothesis states that an elevation in the water table caused an excavation failure, the collected evidence should focus on finding information on water table levels before, during, and after the failure.

Since the primary purpose of evidence is to validate hypotheses, the collection process must be unbiased and objective (Rao, 2016). Objectivity ensures that no part of the evidence is altered or hidden to favor or disprove a hypothesis. Unbiasedness is especially important because models are susceptible to modifications if the input information (i.e., evidence) is disturbed. Therefore, alteration of water table records, pore water pressures, stress states, construction process, or geo-mechanic parameters can decisively influence conclusions about the causes of failure.

The evidence collected from geotechnical failures can be qualitative or quantitative. Qualitative evidence is common in geotechnical engineering. It describes stability conditions, deformation magnitudes, or the characteristics under which a failure occurred. The stability of a geotechnical system is the most common qualitative evidence used in forensic assessments. In this case, an expert defines the failure of a geotechnical system based on some functionality or deformation criteria. The expert can also recognize a failure if its consequences are observable. For example, a levee failure can be recognized by flooding in nearby areas, and an excavation failure can be identified by the damage caused to adjacent structures (Kool et al., 2019).

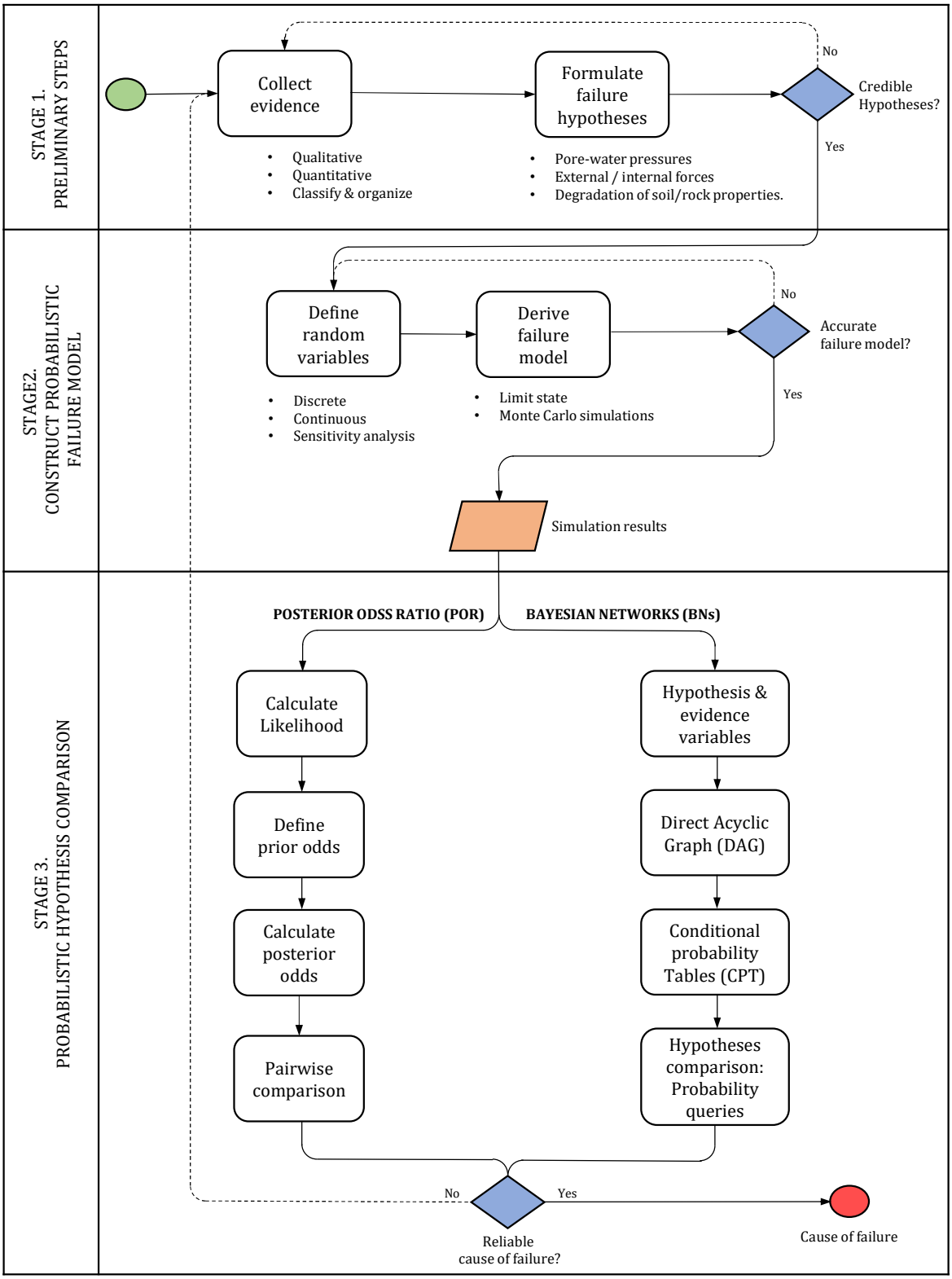


Figure 4-1. Proposed Bayesian methodology for decision support in Forensic Geotechnical Engineering.

Quantitative evidence focuses on monitoring the performance of critical variables that describe the geotechnical system. In general, quantitative evidence is more reliable and accurate than qualitative evidence. However, collecting quantitative evidence is challenging because, in most cases, it requires a monitoring system conceived in the design and implemented during the construction and operation stages. Deformations, pore water pressures, and stress measurements are standard monitoring variables used as quantitative evidence. Unfortunately, quantitative evidence is not always available in forensic assessments because most failed structures do not include an appropriate monitoring system, or the information was not acquired.

In some cases, qualitative and quantitative evidence are mutually interchangeable. In other words, qualitative evidence can be translated into quantitative and vice versa. For example, when limit equilibrium methods are used, a qualitative descriptor such as "unstable" can be translated into a quantitative factor such as the Factor of Safety. In this case, a value less or equal to 1.0 means an "unstable" stability condition. Likewise, qualitative descriptors such as *high, medium, low,* or *yes* and *no* can be used to simplify the range of quantitative variables used in complex numerical models.

The evidence must be carefully classified and organized regardless of its origin. The Bayesian methodology proposed in this thesis relies on this organization to draw conclusions about the causes of geotechnical failures. For the purposes of this thesis, evidence will be assembled into a set D, where each element d_i corresponds to a piece of evidence. As demonstrated in example of Section 4.2, the amount of evidence from the set D included in the forensic analysis is decisive for drawing conclusions. In summary, the classification and proper use of pieces of evidence are the basis for formulating failure hypotheses and identifying the most probable causes of failure.

Formulating failure hypotheses

Formulating competing hypotheses about the causes of failure is one of the most challenging tasks in forensic geotechnical analyses. Each hypothesis must be able to explain the geotechnical failure and must be tested against the evidence. As mentioned in Chapter 3, the inductive process used in the conventional design is inappropriate for formulating hypotheses because it uses apriorism reasoning (Brady, 2012). Therefore, abductive reasoning should be

used instead of apriorism reasoning, and evidence should guide the formulation of failure hypotheses. Some critical factors for formulating failure hypotheses are described below.

Formulating "credible" failure hypotheses

A critical aspect of the formulation process is that competing hypotheses must be credible. All competing hypotheses should be based on the predictable behavior of soil/rock materials and expected external/internal forces acting on the geotechnical structure. In geotechnical engineering, credible hypotheses are associated with changes in stress states. These changes result from the variation of:

- Pore-water pressures: Changes in pore-water pressures within soils/rocks may be caused by the increase (or decrease) in groundwater levels. For example, sustained heavy rainfalls may raise the water table in just a few hours and produce changes in pore-water pressures, resulting in variations of effective stress states within the soil. An excavation may lead to a two/ three-dimensional water flow through the soil/rock, causing a decrease in the water table. This decrease also results in changes in effective stress states.

- External/internal forces: Variations in external or internal forces can also change the stress state within the soil/rock mass. External forces such as point or linear forces caused by vehicular traffic, construction equipment, nearby buildings, earthquakes, or excavation activities may significantly change stress states. Internal forces are less susceptible to variation because they are related to the unit weight of soils and rocks. However, a variation in pore water pressures within the soil/rock can cause a significant change in internal forces, especially those related to effective stresses.

- Soil/rock-environment interaction and their influence on geomechanical properties: Weathering, erosion, and chemical and biological changes illustrate the soil/rock-environment interaction that may impact strength and deformation properties. Although the impact of these processes is well known in geotechnical engineering, their quantification is difficult due to the complex physical-chemical processes that occur within the soil/rock (Gens, 2010).

Use of semantic expressions (idioms)

The cause-effect relationships or semantic expressions (idioms) described in Section **2.2.7** can be used to formulate failure hypotheses. Hypotheses should be verbalized in a way that cause-effect relationships are identifiable. In other words, input (cause) variables should be clearly distinguished from output (effect) variables. Geotechnical idioms are helpful for this classification because they usually reflect how geotechnical models work. The following are some examples of semantic expressions (idioms) used as failure hypotheses:

- Example 1: The rise of the water table to level -1.0 m led to slope instability.
- Example 2: An unexpected surcharge of 60 kN/m near the excavation caused a settlement of 0.60 m.
- Example 3: Weathering of the rock mass reduced joint strength. Therefore, several wedges collapsed.
- Example 4: The earth retaining structure was underdesigned. Therefore, the structure was not able to support lateral earth pressures.
- Example 5: The failure was a random event. Therefore, although soil/rock materials were correctly characterized and geotechnical structures adequately designed, the failure occurred because of the randomness of natural materials (i.e., sometimes the unpredictable happens).

Example 1 implies a direct cause-effect relationship between the water table and slope stability. This relationship can be estimated through a physical-mathematical model such as limit equilibrium or finite elements. Example 2 is a typical cause-effect relationship between external forces and deformations. The settlement caused by the unexpected force can be estimated using finite element methods with constitutive soil models such as cam-clay, hardening, or soft soil. Example 3 relates the degradation of rock properties with wedge failure in a rock mass. In this case, a physical-mathematical function needs to define the joint strength reduction due to weathering. Example 4 involves a cause-effect relationship in which an erroneous design caused a failure. Hypotheses like the one in Example 4 require careful consideration because they involve human errors. Finally, Example 5 is considered the null hypothesis or the baseline scenario (i.e., the original design). Its purpose is to be compared with other hypotheses and, if feasible, to be discarded.

Collectively exhaustive and mutually exclusive hypotheses

In order to fulfill probability requirements, hypotheses about the causes of a geotechnical failure must be collectively exhaustive. In other words, credible hypotheses should together encompass the entire range of possible causes that could explain the geotechnical failure. This condition guarantees that at least one hypothesis can explain the causes of failure. Therefore, if there is one hypothesis that seems feasible, it should be included in the forensic analysis. If some hypotheses are not included in the analysis, they cannot be probabilistically evaluated using the proposed methodology.

Regarding the concept of mutually exclusive hypotheses, geotechnical failures can be a combination of multiple causes. Therefore, hypotheses are not mutually exclusive because multiple causes can co-occur. One typical example is about slope instabilities caused by unexpected surcharges and heavy rainfalls infiltrating soils. Another example is an excavation failure caused by an elevation of the water table and uncontrolled construction processes. Although the hypotheses do not necessarily have to be exclusive, the possible states of each hypothesis must be exclusive. For example, the failure of a geotechnical system cannot be explained by both rising and falling water tables simultaneously.

As a final remark about hypotheses formulation, it is worth mentioning that all credible hypotheses must be able to be translated into a probabilistic model. In practice, that means that cause-effect relationships between variables must be able to be represented by physical-mathematical models in which uncertainty is included. Illustrations of probabilistic models, such as the posterior odds ratio using performance indicators (e.g., the factor of safety) and Bayesian networks, are presented in Sections **4.2** and **4.3**.

4.1.2 Stage 2. Constructing the Probabilistic Failure Model

The probabilistic failure model includes two steps: (i) defining random variables and (ii) deriving the probabilistic failure model. Model accuracy is evaluated using a decision node. Each step is described below.

Defining random variables

Once the failure hypotheses have been formulated, the next step is to define the random variables included in the hypotheses. By default, all geotechnical variables are random.

However, the randomness of some of them may be minor or have limited influence on geotechnical behavior. Therefore, random variables are selected according to the following criteria:

- Variables related to one or more hypotheses should be considered random variables. For example, if a hypothesis associates the geotechnical failure to the elevation of the water table, then the variable "water table" should include all possible credible states of the water table to be evaluated.

- In the case of geotechnical failures related to the strength of soil/rock materials (i.e., Strength Limit State analysis), strength variables should be defined as random variables. The effective angle of shear resistance ϕ', effective cohesion c', pre-overburden pressure POP, shear strength ratio S, and strength increase exponent m, are examples of random strength variables in both limit equilibrium and finite element methods.

- For Service Limit State analysis, i.e., when deformations are conditioning the geotechnical analysis, variables such as Young's modulus E, shear modulus G, compression index C_c, swelling index C_s, and initial void ratio e_0, among others, should be considered as random variables.

- Performance variables (i.e., evidence variables) such as Factor of Safety FoS, stability condition SC, settlement ρ, or inclination i are random because they depend on random input variables.

Random variables can be defined as discrete or continuous variables. If a variable can only take a finite number of states, then the variable should be treated as discrete. On the other hand, if the variable can take an infinite number of states, then the variable should be treated as continuous. Most random geotechnical variables are continuous due to the characteristics of physical-mathematical models. For example, strength and deformation parameters are continuous random variables. However, for the sake of simplicity, some geotechnical models can be simplified by discretizing the range of a continuous variable. A classic example of

discretization is the Factor of Safety (FoS). The FoS is continuous in the interval [0, ∞), but it is commonly discretized into two states: *stable* for FoS>1.0 and *unstable* for FoS≤1.0.

The complexity of some geotechnical models and the number of random variables could make Bayesian probabilistic analysis intractable, mainly because of a significant increase in computational cost. In these cases, it is necessary to resort to two strategies: (i) simplify geotechnical models and (ii) reduce the number of random variables. Geotechnical model simplification is feasible when the simplified model can reliably explain the relationships between input (hypotheses) variables and output (evidence) variables. For instance, a limit equilibrium method for a slope stability model could be used instead of a computational strength reduction method (i.e., finite element method). However, model simplification is not always viable, and sometimes complex models should be used.

Regarding the number of variables, comparison techniques such as Bayesian networks and posterior odds ratio (refer to **Figure 4-1**) require a limited number of random variables. Critical random variables should be identified by exploring the sensitivity of the geotechnical model. In other words, once the model is selected, only the variables with the most significant impact on failure should be randomized. In addition, the randomization should consider the geotechnical variables involved in the hypotheses about the causes of failure.

Deriving Failure Model

A geotechnical failure model describes how performance variables behave. A reliable geotechnical failure model should be capable of including failure hypotheses defined in previous stages. For probabilistic analysis, performance variables are random because the model's input variables are also random. Depending on the performance variables, the failure model can consider failure by resistance (Ultimate Limit State -ULS-) or failure by deformation (Serviceability Limit State -SLS-). The factor of safety (FoS) and the Margin of Safety (MS) are the most common performance variables used in ULS analysis (Melchers & Beck, 2018). Both variables can be represented by probability density functions in which a limit state defines the probability of failure. In the case of the FoS, the probability of failure is estimated as $P_f = P(FoS \leq 1.0)$. For the MS, the probability of failure is calculated as $P_f = P(MS \leq 0)$.

A probabilistic failure model for SLS can be defined similarly to a ULS analysis. Deformations and displacements are performance variables frequently used in SLS analysis. However, unlike ULS, the performance variables in an SLS model do not have a unique limit state. For example, the allowable limit deformation for a road slope is larger than the allowable limit deformation for a dam slope. Similarly, the maximum allowable differential settlement for a residential building is greater than for a nuclear power plant. Therefore, each geotechnical structure should define its performance variables for SLS analyses.

In Bayesian forensic analyses, computing the failure model is the most time-consuming and computationally expensive task. Several simplified techniques for estimating failure models, such as direct integration and second-moment methods, are available in the literature. However, these methods limit the number of random variables and their type of probability distribution functions. In Sections 4.2, 4.3 and Chapter 5, only numerical Monte Carlo simulations are used. The Monte Carlo technique artificially simulates thousands of experiments based on the probability distribution of input (hypotheses) variables. Each experiment samples a value from each random variable and estimates performance variables using a deterministic geotechnical (physical-mathematical) model. Since the experiment is repeated thousands of times, enough observations of the performance variables are obtained. Consequently, the probability of failure can be estimated using Equation 4-1.

$$P_f = \frac{n_f}{N} \tag{4-1}$$

where n_f is the number of experiments for which a failure (or undesired performance) is observed, and N is the total number of experiments.

The accuracy of the failure model and hence the probability of failure is a function of the total number of experiments N. On the one hand, too few experiments will result in low model accuracy and, therefore, a poor approximation of the probability of failure. On the other hand, too many experiments will result in a longer computational time and higher computational cost. A trade-off between accuracy and computational cost can be estimated by convergence analyses (Melchers & Beck, 2018). Convergence analysis estimates the value of the probability of failure for a progressive number of experiments using a convergence plot. The value N is

defined once stability is reached in the convergence plot. Additional convergence analysis using standard deviation for performance variables can also be used to estimate N. For example, **Figure 4-2** (a) shows a convergence plot for the probability of failure $P(f)$, and **Figure 4-2** (b) a convergence plot for the standard deviation of $P(f)$. Note that in this case after 600 experiments, both values converge to unique values.

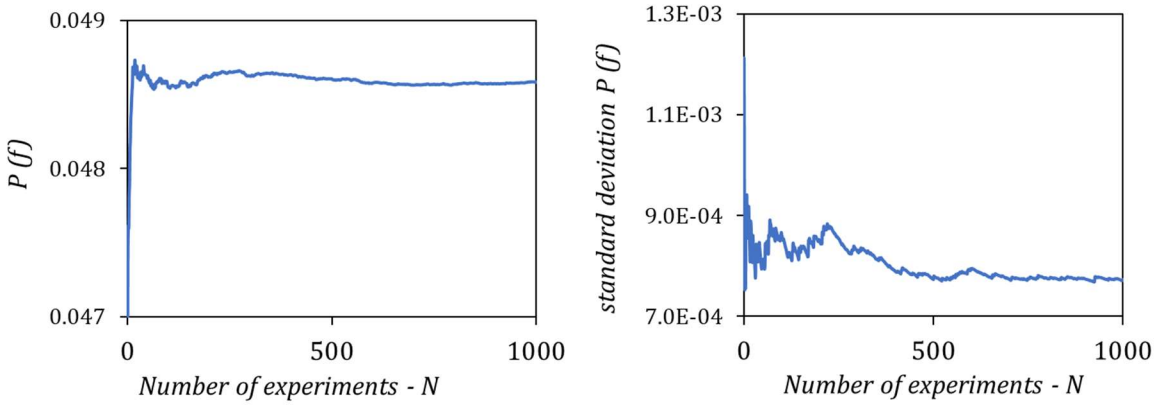


Figure 4-2. Typical convergence plot for $P(f)$ and standard deviation of $P(f)$.

4.1.3 Stage 3. Probabilistic Hypothesis Comparison

The hypothesis comparison is the core of the Bayesian methodology proposed in this thesis. Two Bayesian techniques for hypothesis comparison are chosen: posterior odds ratio and Bayesian networks. Both techniques are based on Bayesian probability but differ in how they deal with information (i.e., evidence) and how the results are reported. Each Bayesian comparison technique is described below.

Posterior Odds Ratio (POR)

The method described in this section compares the probability of multiple hypotheses using the odds ratio. The odds of hypotheses H_1 and H_2 (denoted as $O(H_{12})$) is the ratio of the chance of H_1 being true to the chance of H_2 being true. Hypotheses comparison via posterior odds ratio includes the following steps:

Calculate the Likelihood

Use the available evidence to estimate the likelihood term in the Bayes theorem (see Equation 2-7). This step consists in estimating the probability of observing the evidence d_i, given that hypothesis H_i is true (i.e., $P(d_i \mid H_1)$). The probability can be estimated from the results of Monte Carlo simulations described in Section 4.1.2 by conditioning the evidence d_i on hypothesis H_i. For example, let's assume that evidence d_1 corresponds to an observed slope instability, evidence d_2 corresponds to a circular slip surface, and hypothesis H_1 states that a slope failure was caused by a high-water table. Therefore, the probability of observing an instability and a circular slip surface, given that a high-water table caused the failure, can be estimated from Monte Carlo simulations by counting the number of failures with circular slip surfaces conditioned on a high-water table.

Pieces of evidence can be included one at a time in order to assess their impact on the hypotheses. However, the complete analysis requires all available evidence to be included and contrasted with each hypothesis.

Define prior odds for failure hypotheses

Prior odds (written as $O(H_{ij})$) compares the probability of hypotheses H_i and H_j before any evidence is observed or included in the analysis. In other words, prior odds represent how strongly a forensic investigator believes in a hypothesis H_i compared to hypothesis H_j before conducting a Bayesian analysis. Prior odds can be assigned based on well-established knowledge, experience in similar geotechnical failures, or expert opinion. Given that an incorrect assignment of prior odds can lead to erroneous results and unrealistic conclusions about causes of failure, its definition must be implemented carefully. Incorrect prior odds assignments are evident when the value of $O(H_{ij})$ is unreasonably large or small. For example, two hypotheses are formulated to explain the causes of an excavation failure. The hypothesis H_1 states that an unexpected surcharge value caused the failure, whereas the hypothesis H_2 affirms that an inappropriate construction sequence led to the failure. If the odds ratio for H_1 and H_2 is defined as $O(H_{12}) = 100$, it means that the hypothesis H_1 is one hundred times more likely to be true than the hypothesis H_2 before any evidence is observed. Clearly H_1, is favored over H_2 without a convincing technical or probabilistic argument. When there is no well-established prior knowledge about the causes of a geotechnical failure, prior odds should be

defined as $O(H_{12}) = 1.0$. That is, both hypotheses H_1 and H_2 are equally likely before including any evidence or conducting a Bayesian analysis.

Calculate the posterior odds ratio

The posterior odds ratio is calculated using Equation 2-10. Forensic investigators can use the boundaries defined in Table 2-1 or Figure 2-2 as a guide to interpret results and decide how many times the hypothesis H_{i} explains the failure better than the hypothesis H_{j}. In the case of $O(H_{ij}) = 1.0$, the posterior odds value equals the Bayes Factor (see Section 2.1.4). In order to verify the influence of prior odds on posterior odds, additional analysis can be carried out by modifying prior odds values and the amount of evidence included in the analysis.

Pairwise comparisons

Pairwise comparisons contrast all hypotheses to each other using the posterior odds ratio. The probabilistic comparison allows the forensic investigator to support or discredit some hypotheses formulated in previous stages and draw conclusions about the most probable causes of geotechnical failures.

Bayesian Networks

Bayesian networks (BNs) are the second suggested technique for comparing hypotheses about causes of failure. They encode all the information of the probabilistic failure model in its structure, allowing for more complex probabilistic queries. The construction of BNs and its hypotheses comparison technique is described below.

Identifying hypotheses and evidence variables

All random variables should be identified and classified into input (hypothesis) or output (evidence) variables, following the criteria described in Section 4.1.2. On the one hand, input (hypotheses) variables are those related to the failure hypotheses formulated in the preliminary steps described in Section 4.1.1. Soil constitutive parameters, pore-water pressures, and stress states are good examples of input (hypothesis) variables. On the other hand, output (evidence) variables are those that can be easily measured or observed. For example, slope stability conditions, deformations, or water table levels are output (evidence) variables easily measured by devices or defined through expert opinion.

Input (hypotheses) and output (evidence) variables are usually represented by discrete random variables. BNs use discrete random variables to simplify calculations through conditional probability tables (CPT). Since most geotechnical variables are continuous, it is necessary to discretize them into a finite set of states (intervals). For example, a discretization process for a continuous random variable such as the angle of shear resistance (ϕ') involves the definition of several discrete states that cover the entire range of credible ϕ' values. Three methods for discretization are suggested: (i) Define the number of states according to geotechnical criteria. Low, medium, high or stable and unstable are common descriptors used in geotechnical engineering. However, these descriptors should have a physical meaning and must be widely accepted. (ii) Divide the range of the random variable into n bins. Although this process is straightforward, loss of information is a major concern when the value of n is small (Antonucci, 2018). (iii) Use discretization algorithms such as those presented by Drezner & Zerom (2016) and Fenton & Neil (2019).

Constructing the Direct Acyclic Graph (DAG)

Before constructing the DAG for hypotheses comparison, causal relationships (i.e., causality) between hypotheses and evidence variables should be identified. Causality is the influence of a variable on another represented by deterministic or probabilistic relationships. Since the proposed methodology is based on Bayesian tools, probabilistic relationships are the focus of the analysis. Causality can be inferred from natural cause-effect relationships represented by physical models or semantic relationships known as idioms (refer to Chapter 2). Physical models use mathematical functions to describe the causality between variables. A classic example is the principle of effective stress (σ') within a soil mass. σ' is a function of the total stress (σ) and the pore water pressure (u). Although the relationship is deterministic, it can become probabilistic by including uncertainty in its parameters. In the case of idioms, geotechnical knowledge is encoded in its structure. The five basic idioms described in Section 2.2.7 can be used to infer the probabilistic relationship between variables.

Once causality relationships between variables are identified, the DAG is constructed by assigning one node to each random variable and connecting nodes through edges (arrows). Edges indicate the direction of the cause-effect relationship between the connected nodes and operate as a channel for transferring information (see Section 2.2.7). The final DAG structure must reflect the dependence (or independence) between the variables included in the forensic

geotechnical analysis. In other words, the hypotheses about the causes of failure and the pieces of evidence must be represented in the DAG structure.

The process of constructing the DAG begins by identifying cause-effect relationships among a limited number of variables. Each set leads to a small DAG, and these individual DAGs can subsequently be merged to create a more extensive DAG. An example of constructing an expanded DAG for analyzing a geotechnical failure is detailed in Chapter 5.

Eliciting the Conditional Probability Tables CPT

Databases, expert knowledge, or probabilistic failure models can be used to elicit the CPTs. If a failure model is available, probabilistic relationships can be deduced from the results of Monte Carlo simulations. Refer to Section 2.2.8 for a more detailed description.

Comparing hypotheses using probability queries

A direct acyclic graph (DAG) and its conditional probability tables (CPTs) form a BN. From a broad perspective, a BN is considered both a metamodel and an expert system. A BN operates as a metamodel because it is a simplified model of a more extensive probabilistic model obtained from Monte Carlo simulations. Additionally, a BN is considered an expert system because it can answer probability questions (i.e., conditional probability queries). In conclusion, a BN emulates the decision-making process of a human expert for making decisions under uncertainty.

In FGE, probability queries are used to determine the probability of a failure hypothesis under some evidence. The evidence is included by instantiating one or several nodes (i.e., assigning a unique value to an evidence node) and examining the configuration for the rest of the nodes. For example, a BN can be queried about the probability of observing an unstable condition *(ST=unstable)* near an excavation and a *settlement*(ρ) greater than 0.80 m, given that a high elevation of the *water table* (WT) and an unexpected *surcharge* (Sch) higher than 100 kPa were measured. In mathematical terms, the query is written as $P(ST = unstable, \rho > 0.80m \mid WT = high,\ Sch > 100\ kN)$. In this case, the WT and Sch nodes are instantiated, whereas the configuration of ρ and ST nodes are evaluated.

The above query is an example of a complex probability query that can be solved through a BN. BNs generally allow for more complex probability queries than the posterior odds ratio (POR) technique. This characteristic is advantageous because additional hypotheses can be formulated without further computational experiments or Monte Carlo simulations. For example, the same BN can be used to estimate the probability of several hypotheses, such as:

- The probability of observing an unstable condition and a settlement equal to or greater than 0.80 m near the excavation, given that the water table elevation remains constant, and the surcharge is lower than 100 kN:

$$P(ST = unstable, \rho \geq 0.80m \,|WT = const, Sch < 100\ kN) \tag{4-2}$$

- The probability of observing an unstable condition and a settlement equal to or greater than 0.80 m near the excavation, given that the water table elevation is low, and the surcharge is equal to 100 kN:

$$P(unstable, \rho \geq 0.80m \,|WT = low, Sch = 100\ kN) \tag{4-3}$$

- The probability of observing an unstable condition and a settlement between 0.5 m and 0.8 m near the excavation, given that the water table elevation remains constant, and the surcharge is higher than 100 kN:

$$P(unstable, 0.50m \geq\ \rho \geq 0.80m \,|WT = const, Sch > 100\ kN) \tag{4-4}$$

As described in Section 2.2.10, finding Most Probable Explanations (MPE) is an additional feature of BNs. The proposed Bayesian methodology uses the MPE for finding the configuration of the unobserved nodes (i.e., hypothesis nodes) consistent with the evidence included in the observed nodes (i.e., evidence nodes). Several configurations of hypotheses nodes can explain the evidence included in the observed nodes. For example, an excavation failure can be

explained by two different hypotheses: one related to an elevation of the water table and another to an unexpected surcharge. In this case, the MPE calculates which of these two hypotheses is most probable and designates it as the most probable cause of geotechnical failure. However, if the forensic investigator is interested in examining several causes of failure consistent with the evidence, the K-MPE can be implemented. In the K-MPE, the K (K is a natural number) explanations with the highest probabilities are designated as the most probable causes of failure.

4.2 The ERTC7 Benchmark Exercise

The ERTC7 benchmark exercise presented by Schweiger (2006) is used in this section to validate the posterior odds ratio (POR) and Bayesian network (BN) techniques for hypotheses comparison in FGE. Some modifications are implemented from the original benchmark exercise. For example, the embedded length of the wall is deliberately defined to be short in order to induce a failure. A variation in the water table position is also included. Then, the POR and BN techniques are used to validate if they are able to detect the short embedded length as the main cause of failure. The main stages and the specific steps of the POR and BN branches of **Figure 4-1** are applied to the ERTC7.

4.2.1 General description of the ERTC7

The original ERTC7 benchmark exercise (Schweiger, 2006) aimed to determine the required embedded length of the wall for a deep excavation supported by a strut (**Figure 4-3**). The following are some additional features considered in the original modeling of the ERTC7 benchmark exercise:

- A permanent surcharge of 10kPa and a variable surcharge of 50kPa are located at the top of the excavation (**Figure 4-3**).

- The water table (WT) is located at -4.0 m below the ground surface.

- The bedrock is more than 20m below the ground surface.

- Lowering of the WT and pore water pressures are modeled via steady-state flow calculations.

In the original benchmark exercise, thirteen researchers from different countries submitted their results regarding the required embedded length of the wall. After a comprehensive revision, Schweiger (2006) found significant differences in the results due to several analysis

methods and assumptions made during the calculations. In particular, the embedded length of the wall varied between 1.5 m and 5.5 m, with an average of 3.5 m. According to Schweiger (2006), the differences were mainly due to: (i) the assumed soil-wall friction values, (ii) the methods of analysis employed, such as finite element, finite difference, and limit equilibrium, and (iii) different design methods, strength reduction factors, load magnification factors, and safety factors used in calculations.

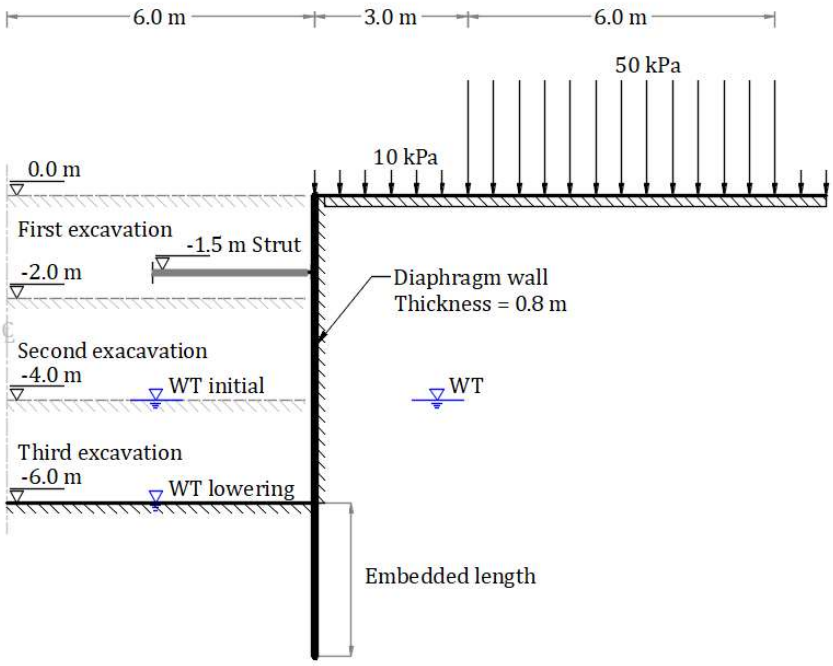


Figure 4-3. The ERTC7 benchmark exercise (Modified from Schweiger, 2006).

Table 4-1 presents the average values of soil and wall parameters and the type of limit equilibrium equations used in this POR example. Additionally, a sequence of the construction phases is shown in **Figure 4-4**.

Table 4-1. ERTC7 benchmark exercise. Adopted values for forensic assessment.

Characteristic	**Variable**	**Value**
Geometry	See Figure 4-3	-
Soil	Unit weight above WT γ_s (kN/m^3)	19
	Unit weight below WT γ_u (kN/m^3)	20
	Effective cohesion c' (kPa)	10
	Effective shear resistance angle ϕ' (º)	27.5
	Young's modulus (kPa)	3.0x10^4
	Poisson´s ratio (-)	0.3
Wall	Unit weight γ_c (kN/m^3)	24
	Young's modulus (kPa)	3.0x10^7
	Poisson´s ratio (-)	0.18
	Soil-wall friction (º)	$\frac{2}{3}\phi'$
	Thickness (m)	0.8
Construction phases	0 – Initial phase: Generation of initial effective stress based on pore water pressures and state parameters. K_0 (-)	0.5
	1 – Wall and surcharge: Installation of wall and activation of surcharge. Elastoplastic drained analysis.	-
	2 – First excavation: Excavation to level -2.0. Elastoplastic drained analysis.	-
	3 – Strut and second excavation: Strut installation at -1.5 m and second excavation to level -4.0 m. Elastoplastic drained analysis.	-
	4 – WT lowering and third excavation: Pore pressure calculation using steady-state flow conditions. Excavation to level -6.0 m using elastoplastic drained analysis.	-
Limit equilibrium analysis	Jambu's equations with circular slip surfaces	-

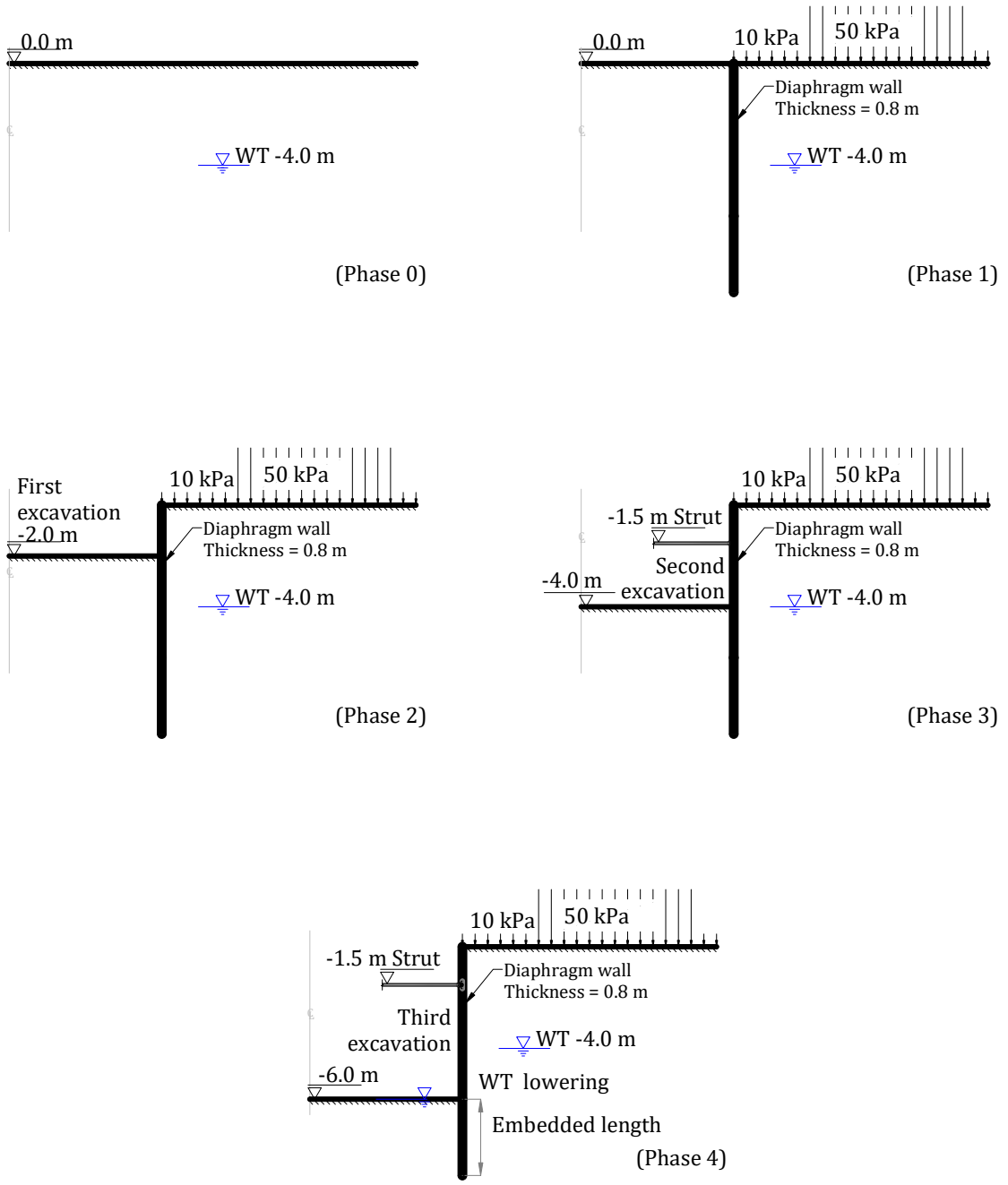


Figure 4-4. ERTC7 Benchmark exercise: Phases of the construction process for numerical modeling.

4.2.2 Stage 1. Preliminary Steps: Evidence and Failure Hypotheses (Hypothetical failure scenario)

A hypothetical failure scenario of the ERTC7 is defined to validate the POR's ability to detect the cause of failure. The hypothetical scenario consists of a deep excavation failure with the characteristics presented in **Figure 4-5** and **Table 4-2**. The excavation's stability condition and the slip surface's geometry are used as the main pieces of evidence. The variables represented by a question mark in **Figure 4-5** are considered hypothesis variables (i.e., the surcharge - *SCh*, the elevation of the water table - *WT*, and embedded length – *El*).

For the hypothetical ERTC7 failure, it is known that a short embedded length of the wall equal to 2.0 m (i.e., $El \leq 2.0\ m$) led to the failure. However, in an actual forensic assessment, the cause of failure is not known in advance. Therefore, several failure hypotheses are proposed and then evaluated to verify if the proposed Bayesian methodology can probabilistically predict the actual cause of failure.

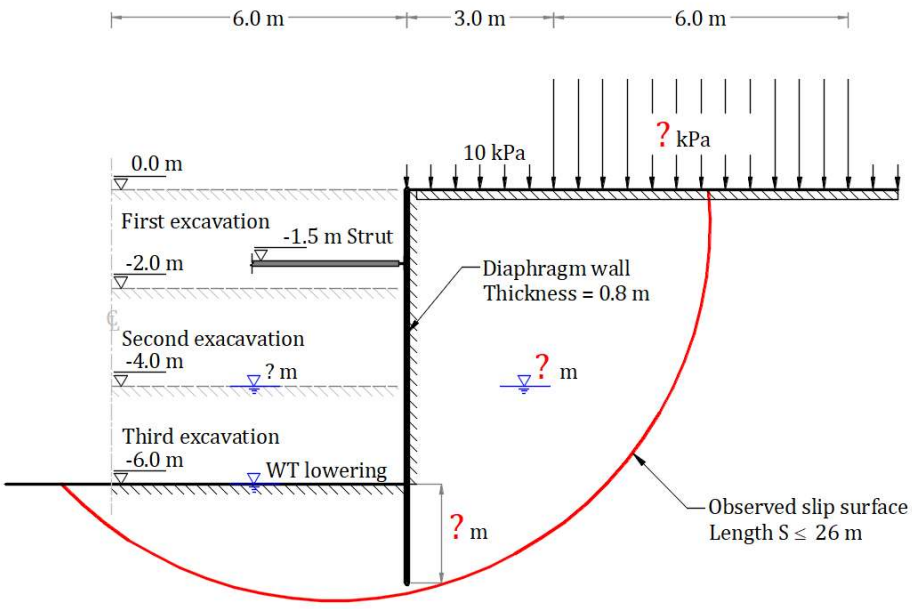


Figure 4-5. The ERTC7 hypothetical failure scenario.

Table 4-2. ERCT7. List of variables used as main pieces of evidence.

Observed variable	**Condition**	**Value**
Excavation stability	Unstable	$FoS \leq 1.0$
Slip surface	Length of the slip surface	$S < 26\ m$

Assume that for the excavation, the actual values of the embedded wall length *(El)*, the water table position *(WT)*, and the overburden *(Sch)* at the time of failure are unknown. Three failure hypotheses ($H_1$to H_3) related to these unknown variables are proposed by different parties involved in the forensic investigation. Additionally, the design values (baseline scenario) are defined as the null hypothesis H_0. The proposed hypotheses are as follows:

- The first hypothesis H_1 (proposed by the construction contractor) states that the wall was constructed according to the design (i.e., $El = 3.5\ m$). However, after construction, there was an unexpected elevation of the water table *(WT)* to the -2.0 m level. This change in the WT level caused an increase in the hydrostatic pressures and the subsequent excavation failure.

- The second hypothesis H_2 (proposed by the project owner) suggests that the wall was constructed according to the design. However, the failure was triggered by an unforeseen increase in the surcharge caused by vehicular and machinery traffic near the excavation site.

- The third hypothesis H_3 (proposed by the affected neighbors and the local authority) suggests that some errors during the construction process caused the failure. In particular, the wall has a shorter embedded length than contemplated in the design.

- The null hypothesis H_0 (baseline scenario) states that the construction process fulfilled the design requirements. Consequently, the excavation failure was caused only by chance (i.e., it was a random event). **Table 4-3** summarizes the failure hypotheses and their values.

Table 4-3. ERCT7. Summary of failure hypotheses and their variable values.

Variable	Hypothesis H_0 Simple chance or random event	Hypothesis H_1 Elevation of the water table (WT)	Hypothesis H_2 Increase in the surcharge	Hypothesis H_3 Short embedded length of the wall
Embedded length of the wall (m)	3.5	3.5	3.5	< 3.5
Water table position (m)	≤-4.0	≥-2.0	≤-4.0	≤-4.0
Surcharge magnitude (kPa)	≤ 50	≤ 50	> 50	≤ 50

4.2.3 Stage 2. Constructing the Probabilistic Failure Model

Random Variables

The original ERTC7 Benchmark exercise uses deterministic values for both material and geometry variables. However, for the purposes of this example, uncertainty is assigned to some variables via probability density functions (PDF) and probability mass functions (PMF). The PDF and PMF are assigned based on the recommendations from technical references (Phoon & Kulhawy,1999a; Phoon & Kulhawy,1999b; Griffiths & Fenton, 2007). Whereas PDFs are used to characterize the uncertainty of soil properties, PMFs are used to characterize variables associated with hypotheses H_0 to H_3 described in Section 4.2.2. **Table 4-4** and **Figure 4-6** present the PDFs and PMFs assigned to continuous and discrete random variables. Truncate normal and discrete uniform distributions are assigned to continuous and discrete variables respectively.

Table 4-4. Probability functions for the random variables involved in the ERTC7 model.

Variable	Symbol	Unit	Variable type	Probability distribution	Parameters*
Effective cohesion	c'	kPa	Continuous	Normal truncate	Mean: 10kPa COV = 20% LB =4.0 kPa UB = 16 kPa
Effective shear resistance angle	ϕ'	°	Continuous	Normal truncate	Mean: 27.5° COV = 12% LB =17° UB = 38°
Soil Young's modulus	E_s	kPa	Continuous	Normal truncate	Mean = $3.0x10^4$ kPa COV = 20% LB =$1.5x10^4$ kPa UB =$4.5x10^4$ kPa

Variable	Symbol	Unit	Variable type	Probability distribution	Parameters*
Wall embedded length	EL	M	Discrete	Uniform	LB = 1.5 m UB = 4.5 m n = 7
Water table location	PL	M	Discrete	Uniform	LB = -4.0 m UB = 0.0 m n = 5
Surcharge	Sch	kN/m^2	Discrete	Uniform	LB = 40 kN/m^2 UB =70 kN/m^2 n = 4

*COV: Coefficient of variation, LB: lower bound, UP: Upper bound, n: number of states.

Probabilistic Failure Model (Limit Equilibrium)

A probabilistic model is created to determine the causes of the hypothetical ERTC7 failure described in Section 4.2.2. The model is based on the probability functions for discrete and continuous random variables presented in **Table 4-1**. Slope stability calculations are estimated via limit equilibrium analysis using Jambu's corrected equations. In order to ensure the representativeness and reproducibility of the probabilistic failure model, a Python script was developed to implement $N = 140{,}000$ computer experiments of the ERTC7 model using Slide v 5.0 (Rocscience Inc., 2006). The Python script included a "crude" Monte Carlo simulation in which the values of the random variables are drawn from their probability distributions. **Figure 4-7** shows the histogram of the Factor of Safety (FoS) and the fitted log-normal probability distribution derived from the computer experiments.

The convergence of the probability of failure $P(f)$ (i.e., probability of $FoS \leq 1.0$) and its estimated standard deviation ($sd_{P(f)}$) derived from the N Monte Carlo simulations are shown in **Figure 4-8** and **Figure 4-9**, respectively. Note that $P(f)$converges to the value of 0.143 after approximately 50,000 runs. Similarly, $sd_{P(f)}$ converges to 0.0023 for a similar number of runs. Consequently, the representativeness of the model is guaranteed with the proposed number of simulations.

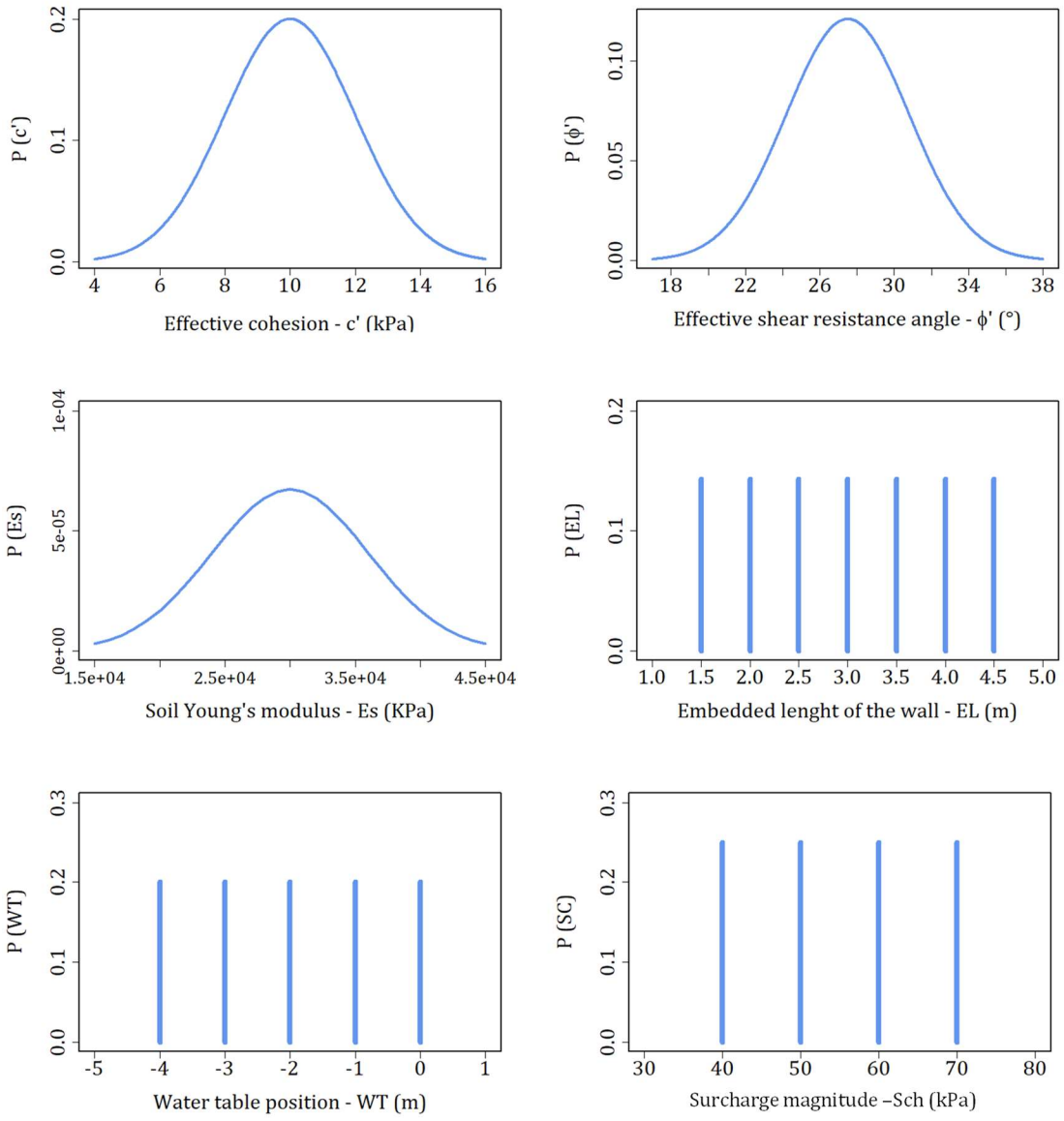


Figure 4-6. ERTC7 Benchmark exercise: Probability density function (PDF) and probability mass functions (PMF) for continuous and discrete random variables.

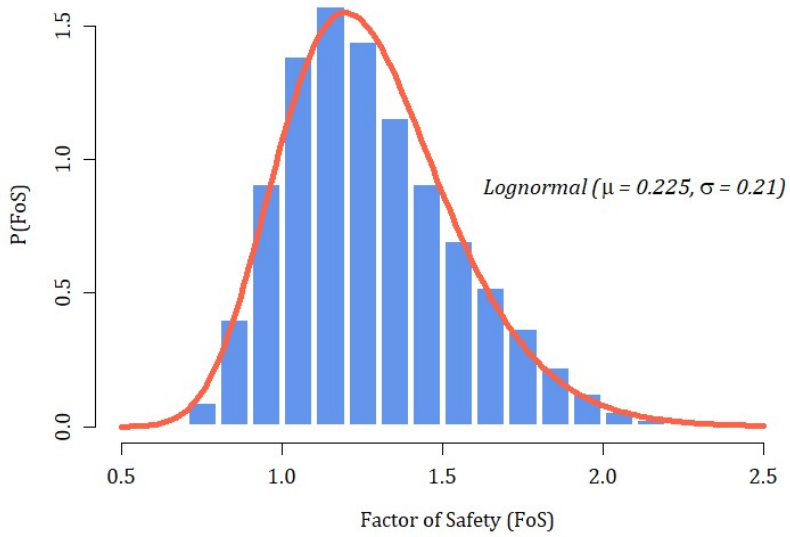


Figure 4-7. ERTC7 Benchmark exercise: Histogram and fitted Log-normal distribution for the Factor of Safety (FoS).

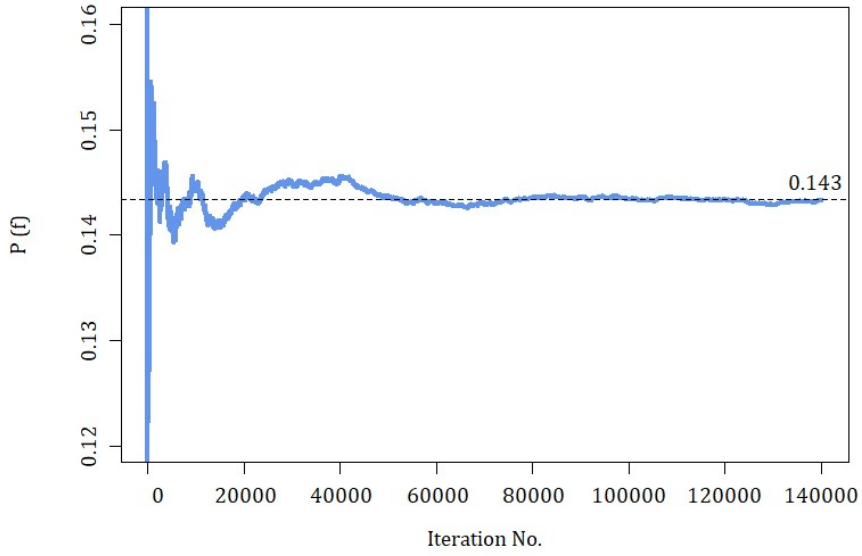


Figure 4-8. ERTC7 Benchmark exercise: Convergence plot for $P(f)$

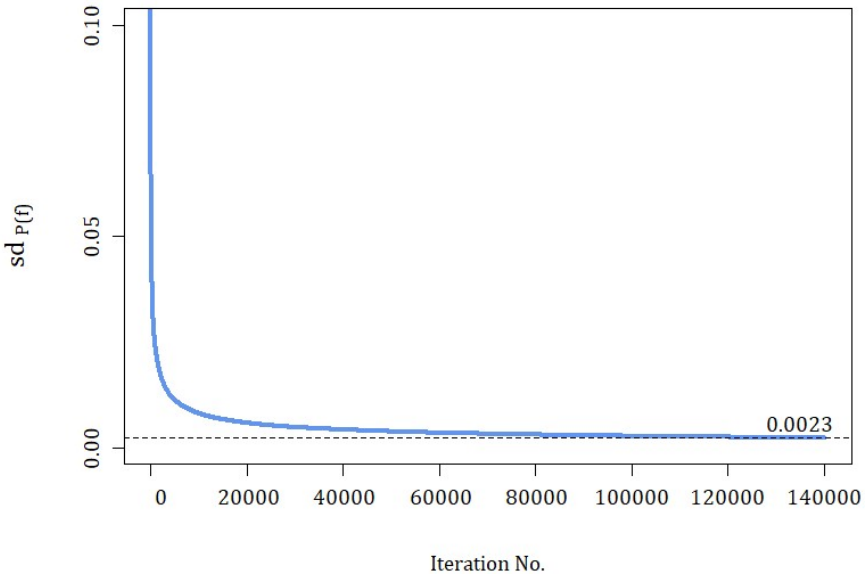


Figure 4-9. ERTC7 Benchmark exercise: Convergence plot for the standard deviation of $P(f)$.

4.2.4 Stage 3(a). Probabilistic Hypotheses Comparison via POR Technique

The proposed hypotheses regarding the causes of excavation failure are compared against the baseline scenario (i.e., the proposed design) and between them. The posterior odds ratio (POR) technique is used and includes the following steps:

- Use the available evidence to estimate the likelihood term in Bayes' theorem (see Equation 2-6). In other words, estimate the probability of observing the evidence d_i given that hypothesis H_i is true. In the case of the ERTC7 exercise, the likelihood refers to the probability of simultaneously observing an unstable condition and a slip surface length shorter than 26 m, given that the failure hypothesis H_i is true. In mathematical terms: $P(d = FoS \leq 1.0 \,\&\, S < 26\,m \mid H_i)$. The $N = 140{,}000$ results from stability calculations are used to estimate these probabilities. For example, the likelihood $P(d \mid H_1)$ is estimated by counting the number of unstable cases ($FoS \leq 1.0$) with a slip length shorter than 26 m ($S < 26\,m$) conditioned on the values of hypothesis H_1 (third column of **Table 4-3**).

- Establish the prior odds of H_i. as $P(H_i/H_j) = O(H_{ij})$ based on actual data, experience, or expert opinion. For the ERTC7 exercise, several $O(H_{ij})$ can be proposed based on prior knowledge. For example, if there are severe concerns about the construction process and the embedded length of the wall, then it can be argued that $P(H_3/H_0) = O(H_{30}) = 10$. It means that hypothesis H_3 is ten times more likely than the null hypothesis H_0. On the other hand, when $O(H_{30}) = 0.1$, it is presumed that the construction process meets the design. In this case, H_0 is ten times more likely than H_3.

- Calculate the posterior odds using Equation 2-10 and Table 2-1 to determine how many times better H_i explains the evidence than H_j. For example, in the case of H_1 vs H_0, the posterior odds assess how much the elevation of the water table better explains the excavation failure than the baseline scenario (i.e., a failure due to a random process).

- Comparison between failure hypotheses $H_1$to H_3. For the ERTC7, the comparison allows the investigator to support or discredit some hypotheses as the causes of excavation failure.

Bayesian Hypotheses Comparison for H_1 vs H_0

The probability of excavation failure due to the elevation of the WT (H_1) and the probability of failure due to random chance (H_0) are compared. Equation 4-5 shows the ratio of posterior odds for this analysis. The likelihood terms for H_1 and H_0 are estimated from the slope stability calculations. The analysis defines $k = 2$ stages for the amount of evidence included in the analysis. In the first stage ($k = 1$) only the observed stability condition $d_1 = FoS \leq 1.0$ is included as evidence. The second stage ($k = 2$) includes the stability condition and the observed slip surface length ($d_2 = FoS \leq 1.0$ & $S < 26\ m$). The likelihood values estimated from conditioning the evidence d_k on hypotheses H_i are presented in **Table 2-1**.

$$\frac{P(H_1 \mid d_k)}{P(H_0 \mid d_k)} = \frac{P(d_k \mid H_1)}{P(d_k \mid H_0)} \frac{P(H_1)}{P(H_0)} \tag{4-5}$$

Table 4-5. Likelihood values for H_1 and H_0 for stages of evidence inclusion

Amount of evidence (k)	Description	$P(d_k \mid H_0)$	$P(d_k \mid H_1)$
1	Evidence $d_1 = FoS \leq 1.0$ (Unstable)	9.90×10^{-4}	4.65×10^{-2}
2	Evidence $d_2 = FoS \leq 1.0$ & S < 26m (Unstable and slip surface length < 26 m)	1.61×10^{-3}	6.80×10^{-3}

The prior odds in Equation 4-5 for the amount of evidence $d_1 = FoS \leq 1.0$ are initially defined as $P(H_1/H_0) = O(H_{10}) = 1.0$ (or 1:1). It means, there is no prior knowledge about the causes of failure, and the investigator assumes that the probability for each hypothesis to be true is the same. The posterior odds calculated from $O(H_{10}) = 1.0$ are 46.9 (Equation 4-6). Consequently, by including only the amount of evidence d_1, a failure due to an elevation of the WT is 46.9 times more likely than a failure due to random chance, despite the initial assumption that both hypotheses (H_1 and H_0) were equally likely. In other words, there is "strong" evidence (refer to Table 2-1) in favor of an elevation of the WT as the cause of the excavation failure.

When an additional amount of evidence is included ($d_2 = FoS \leq 1.0 \,\&\, S < 26\,m$), the posterior odds reduce to 4.2 (Equation 4-7). This posterior odds value means the failure due to an elevation of the WT is only 4.2 times more likely than a failure caused by random chance. Therefore, the inclusion of additional evidence reduces the probability of hypotheses H_1 from "strong" to "positive" as the cause of failure (refer to Table 2-1). This result is not surprising, given that the actual failure cause was defined as the short embedded length of the wall (Section 4.2.2).

$$\frac{P(H_1 \mid d_1)}{P(H_0 \mid d_1)} = \frac{P(d_1 \mid H_1)}{P(d_1 \mid H_0)} \frac{P(H_1)}{P(H_0)} = \frac{0.0465}{0.00099} \frac{0.5}{0.5} = 46.9 \tag{4-6}$$

$$\frac{P(H_1 \mid d_2)}{P(H_0 \mid d_2)} = \frac{P(d_2 \mid H_1)}{P(d_2 \mid H_0)} \frac{P(H_1)}{P(H_0)} = \frac{0.0068}{0.00161} \frac{0.5}{0.5} = 4.24 \tag{4-7}$$

Table 4-6 and Figure 4-10 show the posterior odds' variation with prior odds and the regions based on the boundaries defined in Table 2-1. Note that the posterior odds are sensitive to changes in prior odds. When the probability of the hypotheses is not known in advance, the value $O(H_{10}) = 1.0$ is considered as a reference point. However, in case of having prior information such as piezometric measurements or similar past failures, the value $O(H_{10})$ can be modified. For example, prior odds defined as $O(H_{10}) = 0.1$, (or 1:10) means that before any analysis, H_0 is 10 times more likely than H_1.

Table 4-6. Prior and posterior odds values for H_1 and H_0 given evidence d_1 and d_2

Prior Odds $O(H_1)$	**Posterior odds for $d_1 = FoS \leq 1.0$**	**Posterior odds for $d_2 = FoS \leq 1.0$ & $S < 26\ m$**
0.01	0.469	0.042
0.1	4.69	0.424
1	46.9	4.24
10	469	42.4
100	4691	424

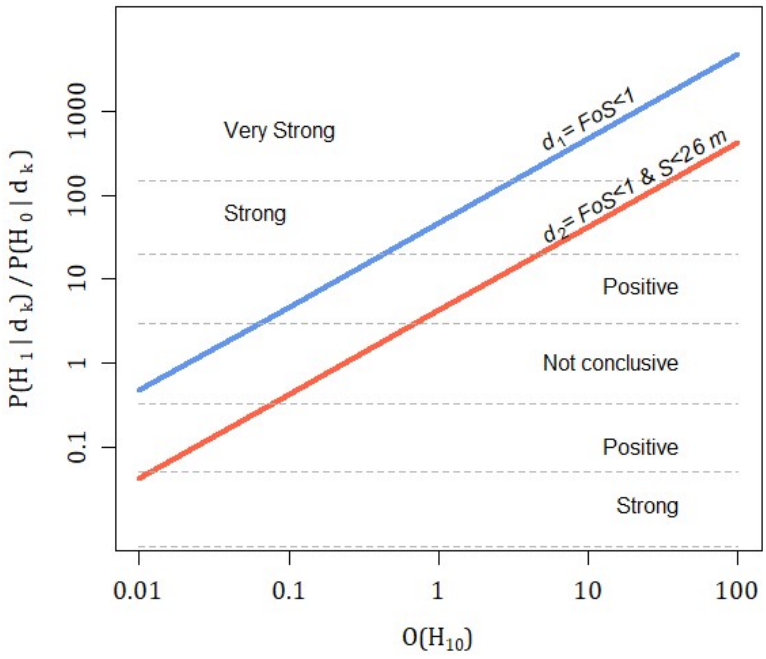


Figure 4-10. Prior and posterior odds values for H_1 vs H_0. Evidence d_1 and d_2.

Pairwise Bayesian Hypotheses Comparison

Pairwise comparison for hypotheses H_0 to H_3 is performed by including different amounts of evidence (d_1 and d_2). Table 4-7 shows the likelihood values estimated from conditioning the evidence on each hypothesis, and Table 4-8 presents the posterior odds for the amount of evidence d_1 and d_2 when prior odds are fixed to $O(H_{ij}) = 1.0$. Additionally, several plots for pairwise analysis $H_i - H_j$ considering variation of $O(H_{ij})$ values are depicted in Figure **4-11**.

Table 4-7. Likelihood values for H_2 and H_3, and evidence d_1 and d_2

Amount of evidence(k)	Stage description	Likelihood $P(d_k \mid H_2)$	Likelihood $P(d_k \mid H_3)$
1	Evidence $d_1 = FoS \leq 1.0$ (Unstable)	1.05×10^{-3}	5.62×10^{-2}
2	Evidence $d_2 = FoS \leq 1.0$ & $S < 26$m (Unstable and slip surface length < 26 m)	1.76×10^{-3}	5.63×10^{-2}

Table 4-8. Pairwise Bayesian hypothesis comparison $H_i - H_j$. (a) Posterior odds for a fixed value $O(H_{ij}) = 1.0$ and evidence $d_1 = FoS \leq 1.0$. (b) Posterior odds for a fixed value $O(H_{ij}) = 1.0$ and evidence $d_2 = FoS \leq 1.0$ & $S < 26\ m$.

(a)

$H_i - H_j$		j		
		H_1	H_2	H_3
i	H_0	46.9	1.1	56.7
	H_1	-	0.02	1.2
	H_2	-	-	53.3
	H_3	-	-	-

(b)

$H_i - H_j$		j		
		H_1	H_2	H_3
i	H_0	4.24	1.1	34.9
	H_1	-	0.25	8.2
	H_2	-	-	31.8
	H_3	-	-	-

Table 4-8 and **Figure 4-11** lead to the following conclusions:

- If only evidence d_1 is included in the analysis, the posterior odds show that a failure due to an elevation of the WT (H_1) is approximately 47 times more likely than a failure due to random chance (H_0). However, when additional evidence (d_2) is considered in the analysis; posterior odds are reduced to 4. This result confirms the importance of

including additional evidence in forensic analysis when comparing two failure hypotheses.

- In the case of H_2 vs H_0, including additional evidence does not have any effect on posterior odds. In other words, no conclusive evidence exists suggesting that an increase in the surcharge led to the failure.

- There is "strong" evidence that the hypothesis H_3 is more likely than H_0 when evidence d_2 is considered. A short embedded depth of the wall explains the excavation failure 35 times better than random chance.

- When comparing H_2 vs H_1 (Table 4-8-b), there are $1/0.25 = 4$ times more chances that an elevation of the WT led to the excavation failure than an increase in the surcharge. On the other hand, a short embedded depth of the wall explains 8.2 times better the excavation failure than an elevation of the WT.

- Finally, there is "strong" evidence in favor of failure excavation due to a short embedded depth of the wall when H_3 and H_2 are compared.

The summary plot of **Figure 4-11** (f) confirms that hypothesis H_3 is the most probable cause of excavation failure. This result demonstrates that the comparison methodology proposed in this research can probabilistically predict the most probable cause of failure. In conclusion, the POR technique indicates that the short embedded length of the wall is the most probable cause of excavation failure, as defined in the hypothetical failure scenario described in Section 4.2.2.

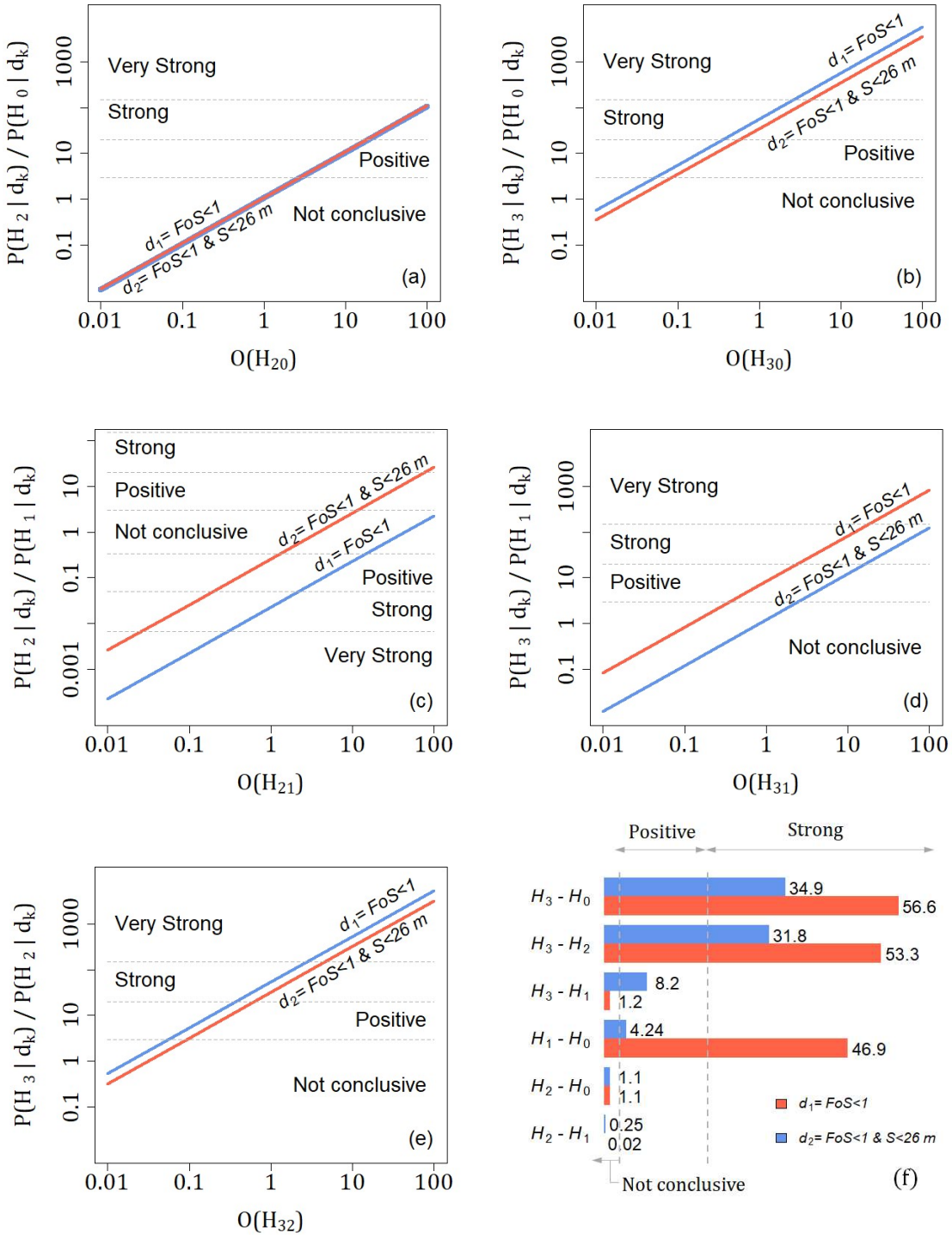


Figure 4-11. Prior and posterior odds values. (a) H_2 vs H_0, (b) H_3 vs H_0, (c) H_2 vs H_1, (d) H_3 vs H_1, (e) H_3 vs H_2, (f) Comparison summary of H_i *vs* H_j for $O(H_{\mathrm{ij}}) = 1.0$.

4.2.5 Stage 3(b). Probabilistic Hypotheses Comparison Using BN Technique

Node classification: Hypothesis and evidence nodes

The BN technique for hypothesis comparison demands assigning a node to each random variable defined in **Table 4-2** and **Table 4-3**. Nodes contain the set of mutually exclusive events or states that form the variable's domain. Also, they require classification as evidence or hypothesis nodes. In the case of the ERTC7 benchmark exercise, the embedded length of the wall *(El)*, the water table position *(WT)*, and the overburden *(SCh)* are selected as hypotheses nodes because they are related to the hypotheses of failure formulated in Section 4.2.2. The slope stability condition (*SC*) and the slip surface (*SS*) are used as evidence nodes because their information can be measured and used as evidence. For example, the *SC* node contains two mutually exclusive states *(stable and unstable)* that an experienced engineer can identify.

On the other hand, the SS node may be more challenging to measure, and only starting and ending locations of the slip surface are identifiable right after a failure. However, if circular slip surfaces are assumed, the total length of the slip surface can be estimated. **Figure 4-5** presents an example of a potential circular slip surface for excavation failure.

Direct Acyclic Graphs (DAG) or Causal Graphs

Constructing the DAG for the ERTC7 excavation failure involves defining the causal relationships between nodes. Causal relationships (i.e., causality) represent the influence of one node (variable) on another. It can be inferred from (i) physical relationships or (ii) semantic substructures that reflect standard knowledge. For example, in the case of standard knowledge, the semantic substructure of the hypothesis H_3 (a short embedded length *(El)* of the diaphragm wall caused the excavation failure) can be translated into the DAG shown in **Figure 4-12**a. However, the semantic structure does not indicate the strength of probability relationships between nodes *EL* and *SC*. Therefore, a physical or mathematical model is needed to describe causality. The influence of *WT* and *Sch* on *SC* and *SS* (hypotheses H_1 and H_2) can be inferred using similar reasoning. **Figure 4-12**b. presents the complete DAG for the ERTC7 exercise in which the blue and red circles represent hypothesis and evidence nodes, respectively.

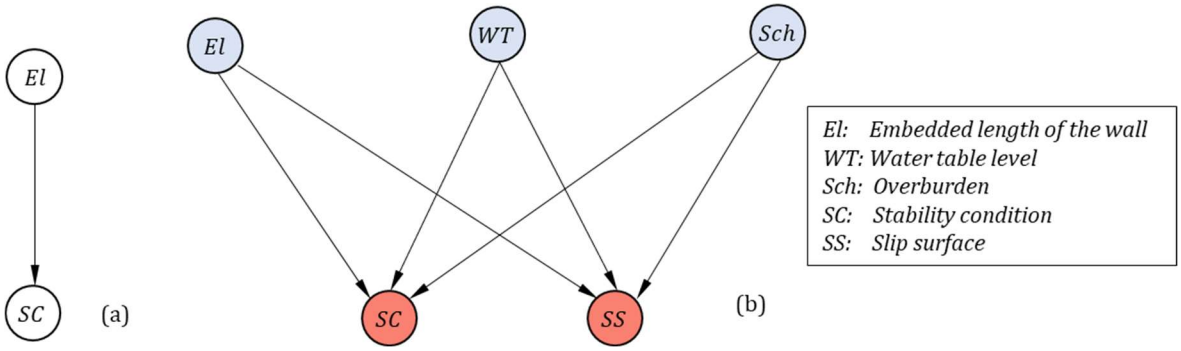


Figure 4-12. (a) A simple DAG to represent the influence of *El* on *SC*. (b) DAG for the ERTC7 excavation failure.

Eliciting of CPTs

As previously mentioned, semantic structures for constructing DAGs do not describe the strength of probability relationships between nodes. Therefore, probability relationships should be estimated using a combination of expert knowledge and geotechnical models. In the case of the ERTC7 exercise, CPTs for hypothesis nodes (blue circles of **Figure 4-12**a.) can be defined using prior expert knowledge. Since it is assumed that no prior information is available for the hypothesis nodes, the probability distribution for the domain of each node can be described by the discrete uniform distributions defined in **Table 4-4** and **Figure 4-6**. In other words, all states included in hypothesis nodes are equally probable before seeing any other data. **Figure 4-13** shows the initial states of the hypothesis nodes.

Regarding the evidence nodes *SC* and *SS*, their prior states (i.e., their states before seeing any information about stability conditions or slip surface length) depend on the probability distribution of their parent nodes. For example, the prior probability distribution of the *SC* node in **Figure 4-13** depends on every combination of *El, WT,* and *Sch*. Since each hypothesis node contains only a few states, the probability dependence of *SC* can be expressed via CPTs. **Figure 4-13** presents the complete BN of the ERTC7 excavation failure and the prior states of hypothesis and evidence nodes.

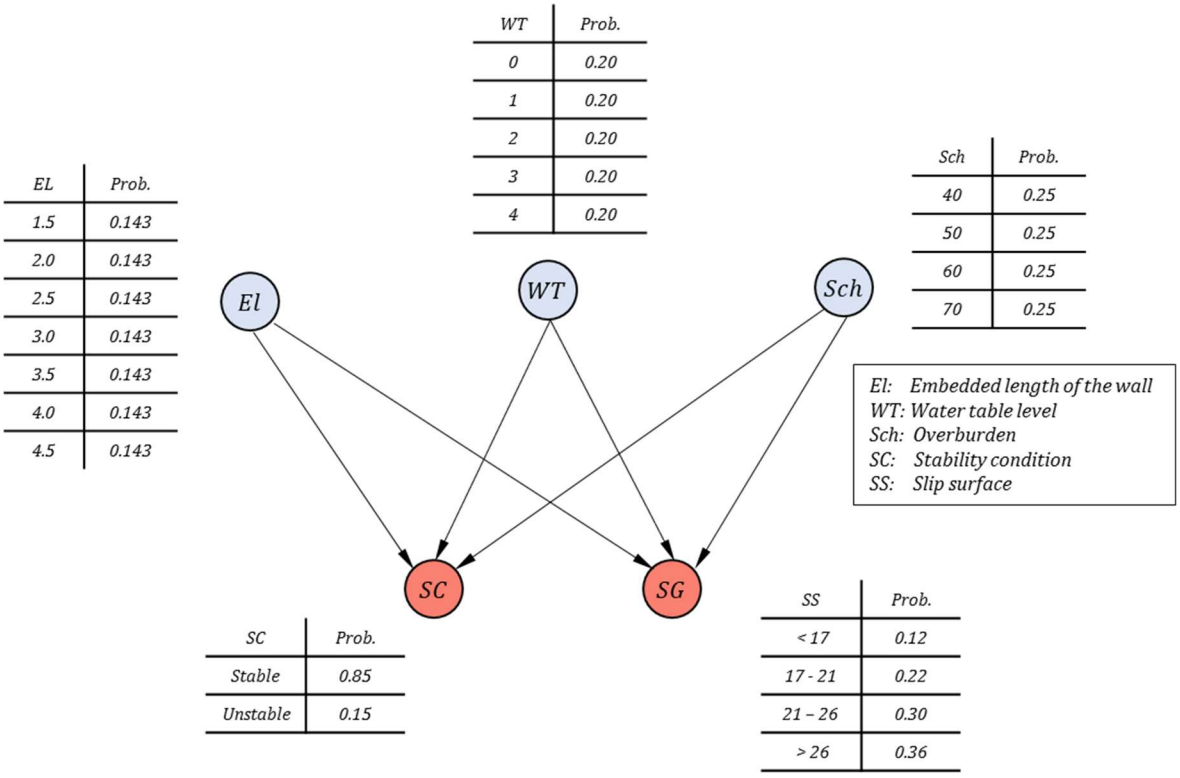


Figure 4-13. Bayesian network for the ERTC7 excavation failure. Prior states of hypothesis and evidence nodes

Hypotheses Comparison Via Probability Queries

The evidence in **Table 4-2** is included in the BN (**Figure 4-13**) to estimate the probability of the hypotheses formulated in Section 4.2.2. As mentioned in Chapter 2, BNs can perform inference processes by inserting pieces of evidence in some of their nodes. The inference process updates the probability of the nodes when new information is included. For the ERTC7 excavation failure, information in **Table 4-2** is included as realizations of evidence nodes.

The conditional probability of hypotheses H_0 to H_3 given the evidence, is estimated using *conditional probability queries* in the form $P(\mathcal{H}_i \mid SC = Unstable,\ SS < 26m)$, where $\mathcal{H}_i$ is a combination of states in hypothesis nodes that represents the hypothesis H_i. For example, $\mathcal{H}_1$, corresponds to the combination of states that represents the hypothesis H_1, or mathematically $\mathcal{H}_1 = \{El = 3.5,\ WT \geq 2.0,\ Sch \leq 50\}$. Therefore, the probability of observing the occurrence of hypothesis H_1 given that the excavation failed and the observed slip length is less than 26 m can be represented as $P(El = 3.5,\ WT \geq 2.0,\ Sch \leq 50 \mid SC = Unstable,$

$SS < 26$). The results of conditional probabilities for hypotheses H_0 to H_3 using the BN of **Figure 4-13** are presented in **Table 4-9**.

Table 4-9. Probabilities for hypotheses H_0 to H_3 given the evidence $SC = Unstable$ and $SS < 26\,m$.

Hypothesis	Description	$P(H_i \mid SC = Unstable, SS < 26\ m)$
H_0	Simple chance, random event, or "Act of God."	3.2 x 10^{-5}
H_1	Elevation of the water table *(WT)*	6.7 x 10^{-3}
H_2	Increase in the surcharge (*Sch)*	3.3 x 10^{-5}
H_3	Short embedded length of the wall *(El)*	2.3 x 10^{-2}

According to the results of probability queries shown in **Table 4-9**, the hypothesis H_3 better explains the excavation failure than hypotheses H_0 to H_2. In other words, the short embedded length of the diaphragm wall is the most probable cause of the excavation failure and its observed slip length. Not surprisingly, this result is similar to that found using the POR technique. In addition, the result confirms that BNs is a feasible technique to determine the most probable causes of geotechnical failures.

Additional Hypotheses comparison using the K Most Probable Explanation (K MPE) Algorithm

As described in Chapter 2, the K MPE algorithm is interested in obtaining the K most probable explanations, i.e., finding K configurations of hypotheses nodes consistent with the observed states in evidence nodes. To this end, all combinations of hypothesis nodes that led to the observed evidence are identified using an R routine (Annex A1). The combinations resulted in 86 additional hypotheses that could explain the ERTC7 failure. The probability of each hypothesis is estimated and organized in descending order.

Table 4-10 shows the first *K=20* hypotheses, and Annex A2 contains the probabilities for the rest of the hypotheses. Note that all the first *K=20* hypotheses include embedded lengths of the

diaphragm wall less than 3.5 m and water table levels varying between 0 m and 2 m. The *Sch* value seems irrelevant to explain the failure.

Table 4-10. $K = 20$ Most Probable Explanations (K MPE) for the ERTC7 Benchmark Exercise.

K-MPE	EL	WT	Sch	Prob.	K	EL	WT	Sch	Prob.
1	1.5	0	70	4.1×10^{-2}	11	1.5	1	40	3.1×10^{-2}
2	1.5	0	40	4.1×10^{-2}	12	1.5	1	60	3.1×10^{-2}
3	1.5	0	60	4.1×10^{-2}	13	2.5	0	70	2.3×10^{-2}
4	1.5	0	50	4.1×10^{-2}	14	2.5	0	40	2.2×10^{-2}
5	2	0	50	3.3×10^{-2}	15	2	1	70	2.2×10^{-2}
6	2	0	70	3.3×10^{-2}	16	2.5	0	60	2.2×10^{-2}
7	2	0	60	3.1×10^{-2}	17	2.5	0	50	2.1×10^{-2}
8	2	0	40	3.1×10^{-2}	18	2	1	50	2.1×10^{-2}
9	1.5	1	50	3.1×10^{-2}	19	1.5	2	40	2.0×10^{-2}
10	1.5	1	70	3.1×10^{-2}	20	2	1	40	2.0×10^{-2}

4.3 The Breitenhagen Levee Failure

The Breitenhagen levee failure described by Grubert (2013) and Kool et al. (2019) explores the applicability of the POR and BN techniques for hypothesis comparison as proposed in this thesis. The levee failure occurred near the Breitenhagen municipality in Saxony-Anhalt, Germany, during the Saala and Elbe Rivers floods of 2013. According to Grubert (2013), slope instability was caused due to high water pressures that had developed inside the levee.

4.3.1 Geometry, Geotechnical Conditions, and Previous Forensic Studies

A cross-section of the levee at the failure location and a simplified stratigraphy are presented in **Figure 4-14**. Note that for the average height of the levee (3.5m), the upstream and downstream slope inclinations are different even though the entire levee is composed of the same clayed (cohesive) material. Underneath this clayed material lies a sandy (cohesionless) soil at a depth of 5.5m measured from the levee's crest. The red dotted arc in **Figure 4-14**

represents the actual slip surface geometry observed during the slope instability described by Kool et al. (2019). All elevations in **Figure 4-14** are referenced to the German vertical datum (NHN - Normalhöhennull).

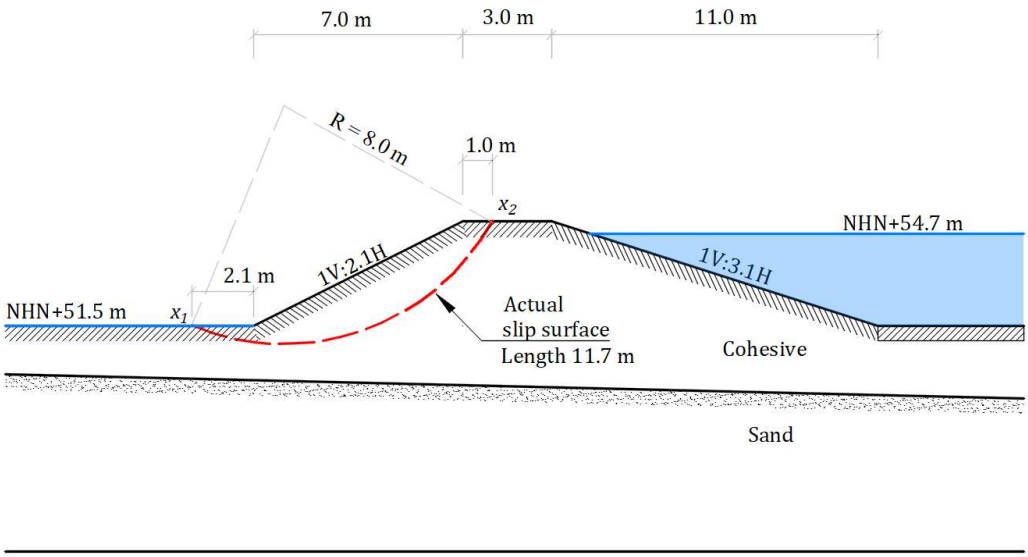


Figure 4-14. Cross-section of the Breitenhagen levee at the failure location and its simplified stratigraphy (Modified from Kool et al., 2019).

Upper and lower boundaries for geotechnical parameters were inferred by Kool et al. (2019) using the soil exploration results reported by Grubert (2013) and typical values reported in Dutch technical literature. The boundaries for clayed and sandy soils using Mohr-Coulomb and SHANSEP constitutive models are reproduced in **Table 4-11**.

Several authors have studied the Breitenhagen levee failure from a forensic perspective. For example, Grubert (2013) investigated the breach's causes by conducting a geotechnical exploration and slope stability analysis. In his study, the tree roots inside the levee were identified as the most probable cause of failure. Moreover, the author highlighted the contribution of the downstream slope angle and the existence of a conductive layer inside the levee as secondary causes of failure. Subsequent studies by Kool et al. (2019) and Kool et al. (2020) argue that locally weak clayed soils and a pond connection with an aquifer with high water pressures led to the levee failure.

Table 4-11. Lower and upper boundaries for clayed and sandy soils at the Breitenhagen levee failure. γ: total unit weight of the soil, c': effective cohesion, ϕ': effective shear resistance angle, S: shear strength ratio, m: strength increase exponent, POP: pre-overburden pressure.

Soil	Mohr-Coulomb			SHANSEP		
	$\gamma_{low/up}$ [$\mathbf{kN/m^3}$]	$c'_{low/up}$ [**kPa**]	$\phi'_{low/up}$ [º]	$S_{low/up}$ [−]	$m_{low/up}$ [−]	$POP_{low/up}$ [−]
Cohesive	15/21	0/15	15/34	0.23/0.49	0.50/0.98	0/150
Sand	19/22	0/0	32/40	-	-	-

4.3.2 Stage 1: Preliminary Steps. Failure hypotheses and Collected Evidence

The causes of failure identified in previous forensic studies (Grubert, 2013; Kool et al., 2019) are used as failure hypotheses. The hypotheses can be divided into two general groups. The first group includes three hypotheses related to changes in pore water pressure conditions inside the levee (see **Figure 4-15**). These hypotheses assume average values for soil strength parameters in drained and undrained conditions.

- H_1: Hypothesis H_1 states that the failure was caused by an unexpected elevation of the water table inside the levee. This elevation is attributed to a sustained high level of the Saala River (**Figure 4-15** b).

- H_2: Hypothesis H_2 suggests that the tree roots inside the levee created a highly conductive layer in which high pore-water pressures were developed (**Figure 4-15** c).

- H_3: Hypothesis H_3 affirms that a highly conductive layer was created through the sandy soil (**Figure 4-15** d) as a result of an early breach in the levee. A pond identified adjacent to the breach suggests a connection between it and the downstream side through a conductive layer (aquifer). According to hypothesis H_3 the failure was caused by high pore-water pressures below the levee aquifer.

- H_0: A base scenario with design soil parameters and average pore water conditions is included in the analysis for comparative purposes. The base scenario represents the

design values, or at least the expected geotechnical performance of the levee. All hypotheses are compared to the base scenario in order to verify their probability. Therefore, the base scenario becomes the null hypotheses H_0 when multiple comparisons are performed.

The second group of hypotheses is related to locally weak soil conditions. Mohr-Coulomb and SHANSEP soil constitutive models characterize drained and undrained soil behavior at the time of failure. In this case, several hypotheses can be established using the values of soil strength parameters. For example, the hypothesis H_4 can be formulated as follows: a combination of weak soil and an elevation of the water table caused the levee failure. Further hypotheses related to geotechnical soil conditions are developed in Section 4.3.5 (Hypotheses Comparison Via Probability Queries).

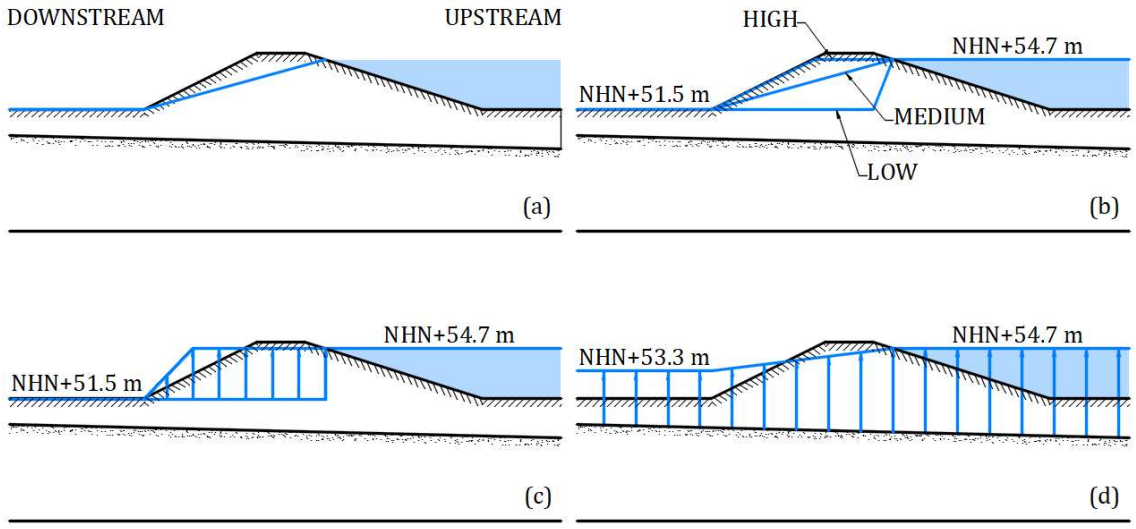


Figure 4-15. Pore-water pressure conditions for hypotheses H_0 to H_3. (a) H_0 base scenario model, (b) H_1 elevation of the water table inside the levee, (c) H_2 conductive layer due to tree roots inside the levee, (d) H_3 high pore-water pressures due to an aquifer in the sandy soil.

Concerning the collected evidence, Grubert (2013) identified the slope failure using visual inspection. The author also describes the consequences of the breach (e.g., floods) to further demonstrate the failure. Additionally, Kool et al. (2019) described the slip surface geometry using a set of photographs before, during, and after the breach. **Figure 4-14** depicts the

circular slip failure identified by Kool et al. (2019), and **Table 4-12** summarizes the collected information used as the main pieces of evidence.

Table 4-12. Collected evidence in the Breitenhagen levee failure.

Evidence	Description	Additional information
e_1	Levee failure	Unstable condition FoS<1.0. Consequences such as loss of levee continuity and floods were observed.
e_2	Circular slip failure	A circular slip failure with a total length of 11.7 m was inferred. Starting (x_1) and ending (x_2) points were identified at 2.1m from the toe and 1.0m from the crown, respectively. This slip surface is named as sg_6

4.3.3 Stage 2: Probabilistic Failure Model

Random Variables

All relevant variables and their randomness must be identified. The levee failure involves several geometric and geotechnical variables. In the Breitenhagen levee case, geometric variables do not exhibit huge variations. Likewise, the unit weight of soils has little influence on slope stability conditions. Therefore, these variables are defined as deterministic, and they are not included in the probabilistic analysis.

Contrary to geometry variables and unit weight of soils, geotechnical variables such as elevation of the water table, pore water pressures, and soil strength significantly influenced stability conditions. A slight variation in the values of these variables could lead to considerable changes in the stability condition and the slip surface geometry. Therefore, the variables presented in **Table 4-13** are defined as random variables. For the sake of POR and BN simplicity, all variables in **Table 4-13** are classified as discrete and described by uniform discrete distributions. Notice that they are also closely related to the hypotheses formulated in Section 4.3.2.

Table 4-13. Random variables for the forensic analysis of the levee failure.

Node (variable)	Description	Probability distribution	Domain (states)
WT	Water Table position – m	Discrete uniform	$[High, Medium, Low]$
CLT	High pore-water pressures in the conductive layer due to tree roots	Discrete uniform	$[Yes, No]$
PCA	High pore-water pressures in the aquifer (pond connection)	Discrete uniform	$[Yes, No]$
c′	Effective cohesion in the clayed soil (Mohr-Coulomb) – kPa	Discrete uniform	$[0, 1, 2, \dots, 15]$
ϕ′	Effective angle of shear resistance (Mohr-Coulomb) – degrees	Discrete uniform	$[15, 16, \dots, 33, 34]$
S	Shear strength ratio (SHANSEP) - /	Discrete uniform	$[0.23, 0.24, \dots, 0.49]$
m	Strength increase exponent -/	Discrete uniform	$[0.50, 0.54, \dots, 0.98]$
POP	Pre-overburden pressure – kPa	Discrete uniform	$[0{,}010, 20, \dots, 150]$
SC	Stability condition	Discrete uniform	$[Stable, Unstable]$
SG	Slip surface geometry (circular failure geometries)	Discrete uniform	$[sg_1, sg_2, \dots, SG_{21}]$

Probabilistic Failure Model (Limit Equilibrium)

Two probabilistic failure models are constructed for the Breitenhagen levee failure. The first model uses a drained soil behavior represented by the Mohr-Coulomb soil constitutive equations. In this case, the effective cohesion (c′) and the effective angle of shear resistance (ϕ') characterize the shear resistance. The second model uses the SHANSEP constitutive equations to describe an undrained soil behavior. The shear strength ratio *(S)*, strength increase exponent *(m)*, and pre-overburden pressure *(POP)* are the variables that represent the SHANSEP soil model. Both models are probabilistic and include the probability distributions presented in **Table 4-13**.

The models were obtained from a computational experiment in which one hundred thousand (100,000) slope stability calculations were performed using D-Stability (Meij & Deltares, 2020). In the experiment, variable values are drawn according to their probability distribution (see **Table 4-13**). The D-Stability calculations were customized in a python script to simplify the running process and data acquisition (Refer to Annex A3 and A4).

The convergence of the probability of failure $P(f)$ derived from the $N = 100{,}000$ computational experiments is shown in **Figure 4-16**. Note that after approximately 40,000 runs, $P(f)$ converges to 0.270 and 0.139 for drained and undrained models, respectively. Consequently, the representativeness of the models is guaranteed with $N = 100{,}000$ computational experiments.

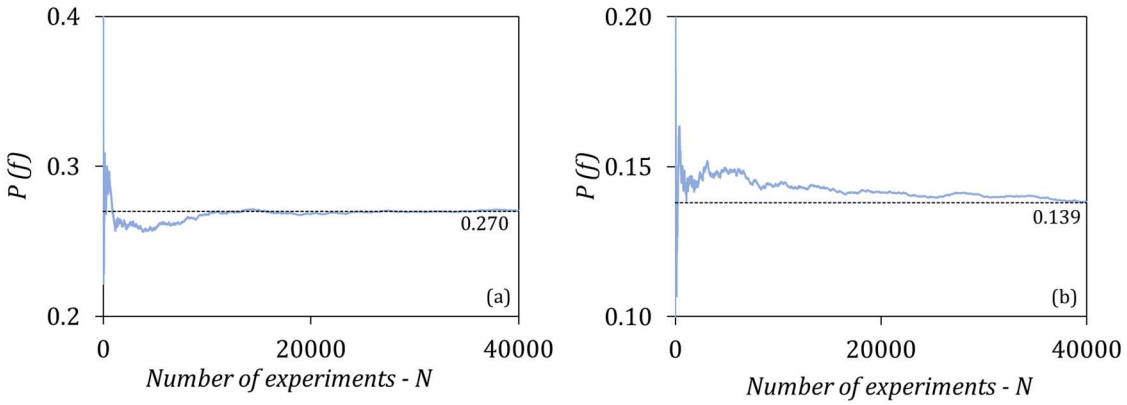


Figure 4-16. The Breitenhagen levee failure: Convergence plot for $P(f)$. (a) Drained model Mohr-Coulomb, (b) Undrained model SHANSEP.

4.3.4 Stage 3 (a). Probabilistic Hypotheses Comparison via POR Technique

The proposed hypotheses regarding the levee failure (H_1 to H_3) are compared against the base scenario (i.e., the proposed design H_0) using the POR technique. Similarly to the ERTC7 failure case, the levee failure analysis includes the following steps:

- Estimate the likelihood term (Equation 2-6). In other words, estimate the probability of observing the evidence e_i given that hypothesis H_i is true. The $N = 100{,}000$ results from stability calculations are used to estimate these probabilities.

- Define prior odds of H_i as $P(H_i/H_j) = O(H_{ij})$. In this a POR analysis $O(H_{ij})$ is defined as 1.0.

- Calculate the posterior odds to determine how many times better H_i explains the evidence than H_j (Equation 2-10).

- Comparison between all failure hypotheses. Results allows the forensic investigator to support or discredit some causes of failure.

Pairwise Bayesian Hypotheses Comparison

Pairwise comparison for hypotheses H_0 to H_3 is performed for the evidence e_1 and e_2. In addition, hypotheses H_4 to H_6 that combines weak soil strength parameters and different pore water pressure conditions are also included in the analysis. These additional hypotheses are proposed to evaluate the impact of weak soils on identifying the causes of failure. Likelihood values estimated from conditioning the evidence on each hypothesis are presented in **Table 4-14**, and posterior odds for evidence e_1 and e_2 when $O(H_{ij}) = 1.0$ are summarized in **Table 4-15**, **Table 4-16** and **Figure 4-17**.

Table 4-14. Likelihood values for H_0 to H_6. Evidence e_1 and e_1 & e_2

Hypothesis H_i	**Description**	Evidence e_1: Levee failure ($FoS \leq 1.0$) **Likelihood P(e_1 | H_i)**	Evidence e_1 & e_2:Levee failure and circular slip surface type sg_6 ($FoS \leq 1.0$ & $SG = sg_6$) **Likelihood P(e_1, e_2 | H_i)**
H_0	Design conditions	0.00	0.00
H_1	Elevation of the water table	8.16×10^{-2}	8.16×10^{-2}
H_2	High pore-water pressures due to tree roots	3.51×10^{-2}	3.51×10^{-2}
H_3	High pore-water pressures in the aquifer below the levee	2.51×10^{-1}	2.51×10^{-1}
H_4	Elevation of the water table inside the levee and weak soils	1.00×10^{0}	7.84×10^{-1}
H_5	High pore-water pressures due to tree roots and weak soils	7.50×10^{-1}	2.95×10^{-1}
H_6	High pore-water pressures in the aquifer below the levee + weak soils	9.84×10^{-1}	0.00

Table 4-15. Pairwise Bayesian hypothesis comparison $H_i - H_j$. Posterior odds for a fixed value $O(H_{ij}) = 1.0$ and evidence $e_1 = FoS \leq 1.0$.

$H_i - H_j$		j					
		H_1	H_2	H_3	H_4	H_5	H_6
i	H_0	*Inf	*Inf	*Inf	*Inf	*Inf	*Inf
	H_1	1.0	0.4	3.1	12.3	9.2	12.1
	H_2	-	1.0	7.1	28.4	21.4	28.0
	H_3	-	-	1.0	4.0	3.0	3.9
	H_4	-	-	-	1.0	0.8	1.0
	H_5	-	-	-	-	1.0	1.3
	H_6	-	-	-	-	-	1.0

*Inf: stands for infinity. $P(H_0) = 0$: the null hypothesis H_0 does not generate the the evidence. In other words, geotechnical and kinematic conditions of H_0 cannot develop a failure and a slip geometry like the sg_6.

Table 4-16. Pairwise Bayesian hypothesis comparison $H_i - H_j$. (a) Posterior odds for a fixed value $O(H_{ij}) = 1.0$, evidence $e_1 = FoS \leq 1.0$ and $e_2 = (SG = sg_6)$

$H_i - H_j$		j					
		H_1	H_2	H_3	H_4	H_5	H_6
i	H_0	*Inf	*Inf	*Inf	*Inf	*Inf	*Inf
	H_1	1.0	0.4	0.0	9.6	3.6	0.0
	H_2	-	1.0	0.0	22.3	8.4	0.0
	H_3	-	-	1.0	Inf	Inf	Inf
	H_4	-	-	-	1.0	0.4	0.0
	H_5	-	-	-	-	1.0	0.0
	H_6	-	-	-	-	-	1.0

*Inf: stands for infinity. $P(H_0) = 0$: the null hypothesis H_0 does not generate the the evidence. In other words, geotechnical and kinematic conditions of H_0 cannot develop a failure and a slip geometry like the sg_6.

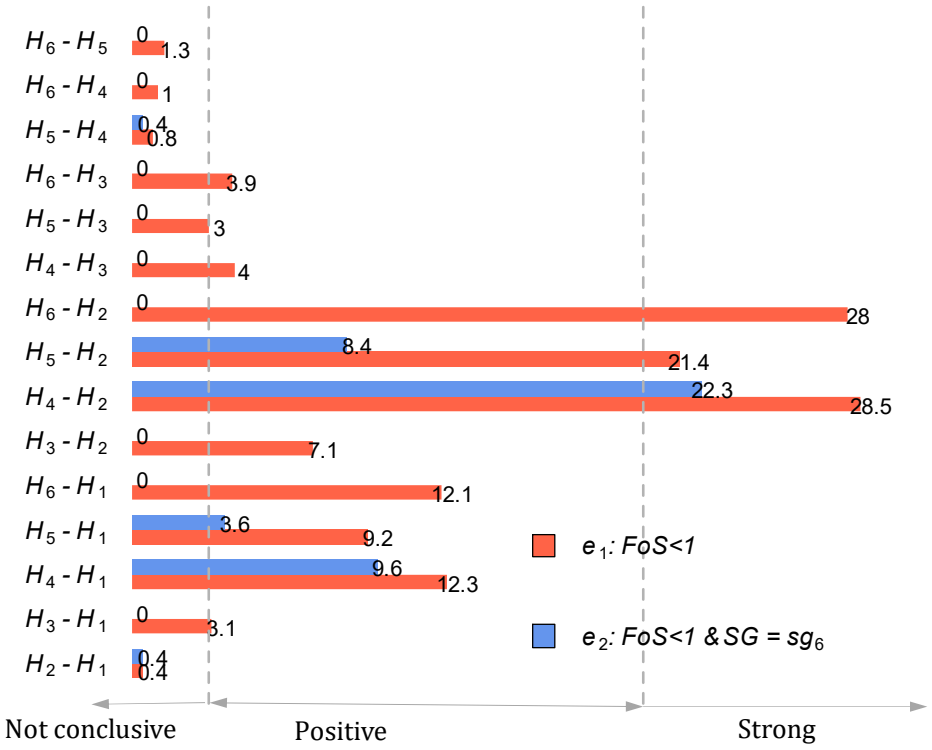


Figure 4-17. Prior and posterior odds values. Comparison summary of H_i *vs* H_j for $O(H_{ij}) = 1.0$.

The pairwise comparison using the POR technique led to the following conclusions:

- If only evidence e_1 is included in the POR analysis, the posterior odds show that hypotheses H_4, H_5 and H_6 are more likely than the hypothesis H_0 to H_3. However, when additional evidence e_2 is considered, posterior odds are drastically reduced with the exception of H_4. In other words, an elevation of the water table and weak soils (hypothesis H_4) is the most likely explanation of the levee failure.

- "Inf" values in **Table 4-15** and **Table 4-16** for row H_0 suggest that hypothesis H_0 does not have the geotechnical and kinematic conditions leading a failure with the observed zslip geometry sg_6.

- "Inf" and "zero" values in **Table 4-16** also suggest that in some cases, hypothesis H_1, H_2 and H_3 cannot explain the failure and the observed slip geometry sg_6.

- Finally, when all available evidence e_1 and e_2 is included in the POR analysis, only pairwise comparisons for hypothesis H_4 remain in the positive and strong sectors of evidence scale (refer to **Figure 4-17** and **Figure 2-2**).

4.3.5 Stage 3 (b). Probabilistic Hypotheses Comparison via Bayesian Networks

Node classification: Hypothesis and evidence

In the case of the Breitenhagen levee failure, *WT, CLT, PCA, c′, ϕ′, S, m*, and *POP* are selected as hypotheses nodes because they are related to the hypotheses formulated in Section 4.3.2. On the other hand, the slope stability condition (*SC*) and the slip geometry (*SG*) are used as evidence nodes because their information can be directly observed and used as evidence (**Figure 4-20**). For example, the *SC* node involves two states (*stable and unstable*) that an experienced engineer can easily identify through visual inspection. For the *SG* node, starting and ending locations are identifiable right after the failure, and commonly circular slip surfaces are assumed in clayed soils. Therefore, several circular surfaces may define the *SG* domain. **Figure 4-18** presents an example of two out of 21 potential slip surfaces defined for the levee failure.

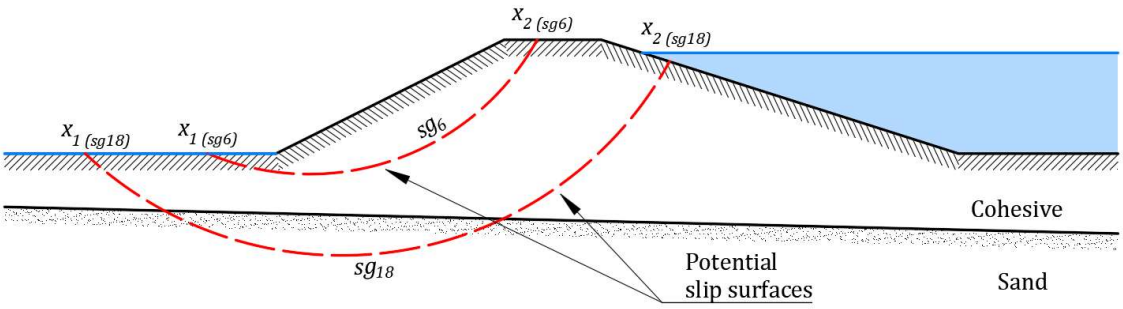


Figure 4-18. Example of two potential slip surface geometries for the Breitenhagen levee failure.

Direct Acyclic Graphs (DAG) or Causal Graphs

As mentioned in previous examples, causality can be inferred from (i) physical relationships or (ii) semantic substructures that reflect standard knowledge. For example, the influence of c' and ϕ' on *SC* for drained conditions can be represented by the DAG shown in **Figure 4-19**a. This relationship is inferred from Janbus's equations which describes the influence of soil strength on slope stability. The relationships between *S*, *m*, *POP,* and *SC* for undrained analysis and their corresponding DAG are constructed using similar physical relationships.

Common geotechnical knowledge can also be used for inferring causality. The semantic substructure of the hypothesis H_1 (*failure was caused by an unexpected WT elevation inside the levee*) can be translated into a DAG, as shown in **Figure 4-19**b. Although the influence of *WT* on *SC* and *SG* is described in the hypothesis, the sentence does not indicate the physical or mathematical relationships between them. Consequently, additional models (mathematical or physical) are needed to characterize the influence. The influence of hypotheses H_2 and H_3 on *SC* and *SG* can be inferred using similar reasoning.

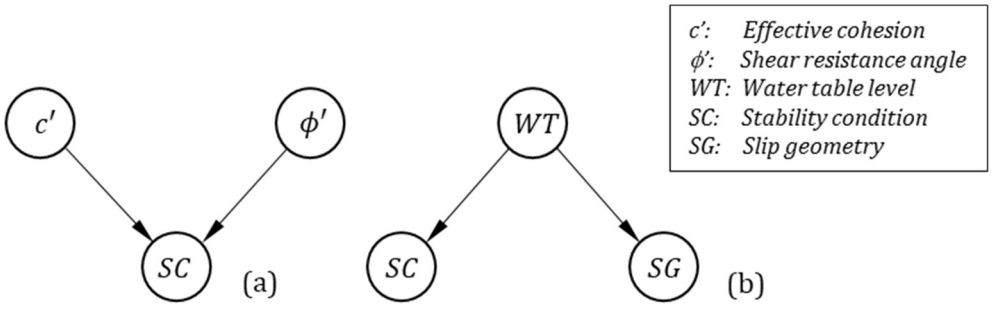


Figure 4-19. Simple DAGs. (a) Influence of c′ and ϕ′ on SC. (b) Influence of WT (H_1) on SC and SG

Figure 4-20 presents the two DAG for the Breitenhagen levee failure inferred from physical and semantic structures. The DAG of **Figure 4-20**a represents the levee failure analysis using the Mohr-Coulomb soil constitutive model for drained conditions. The DAG of **Figure 4-20**b shows an undrained condition using the SHANSEP model. In both figures, red circles correspond to evidence nodes, and blue circles represent hypothesis nodes.

Eliciting of CPTs

After constructing the DAGs, the strength of causality between nodes is defined via probability relationships. In the case of the DAGs shown in **Figure 4-20**, the hypothesis nodes (blue circles) do not depend on other nodes; thus, their probability distributions can be described by the discrete uniform distributions. **Figure 4-21** displays the CPTs for the hypothesis nodes constructed from the information of probability distributions of **Table 4-13**.

The prior states of evidence nodes (*SC* and *SG*) depend on the probability distribution of their parent nodes. For example, the prior probability distribution of the *SC* node in **Figure 4-20**a depends on every single combination of *WT, CLT, PCA,* c', and ϕ' values. All CPT values for *SC* and *SG* were obtained from the probability models described in Section 4.3.3 (e.g., **Table 4-17**). The *SC* and *SG* values were collected from the computational experiments and summarized through the CPTs shown in **Figure 4-21** and **Figure 4-22**.

The DAGs of **Figure 4-20** and their corresponding CPTs constitute the BNs used in the Breitenhagen failure analysis. The resulting BNs and their prior probabilities are presented in **Figure 4-23**.

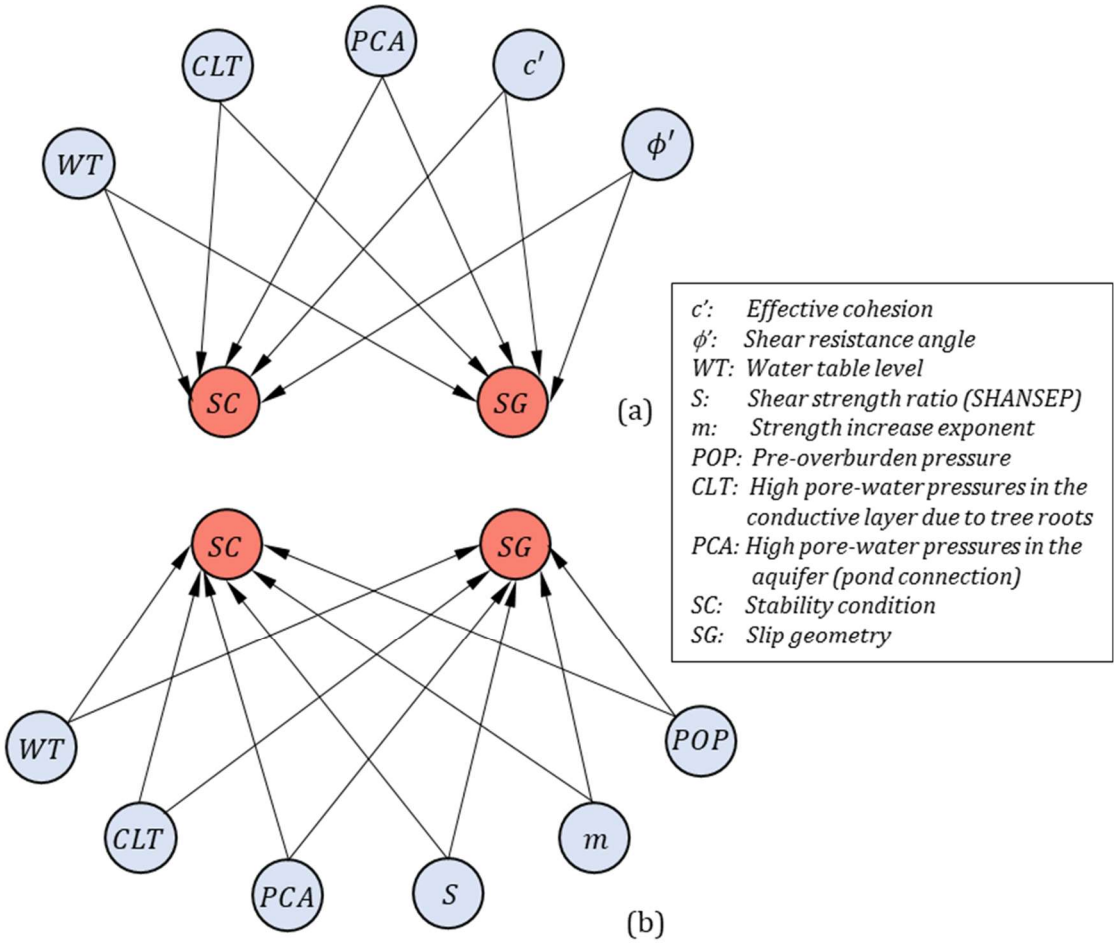


Figure 4-20. DAGs for the levee failure analysis (a) Drained conditions using a Mohr-Coulomb model, and (b) Undrained conditions using the SHANSEP model. Red and blue circles represent evidence and hypothesis nodes, respectively.

Table 4-17. Partial segment of the CPT for the SC node (the complete CPT contains 209 additional matrix slices)

CLT = No, ϕ'= 15, PCA = Yes, WT = High																	
		c′															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SC	Stable	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	Unstable	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0

CLT = Yes, ϕ'= 15, PCA = Yes, WT = High																	
		c′															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SC	Stable	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	Unstable	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0

CLT = No, ϕ'= 16, PCA = Yes, WT = High																	
		c′															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SC	Stable	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	Unstable	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0

CLT = Yes, ϕ'= 16, PCA = Yes, WT = High																	
		c′															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SC	Stable	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	Unstable	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0

CLT = No, ϕ'= 17, PCA = Yes, WT = High																	
		c′															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SC	Stable	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	Unstable	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0

CLT = Yes, ϕ'= 17, PCA = Yes, WT = High																	
		c′															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SC	Stable	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	Unstable	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0

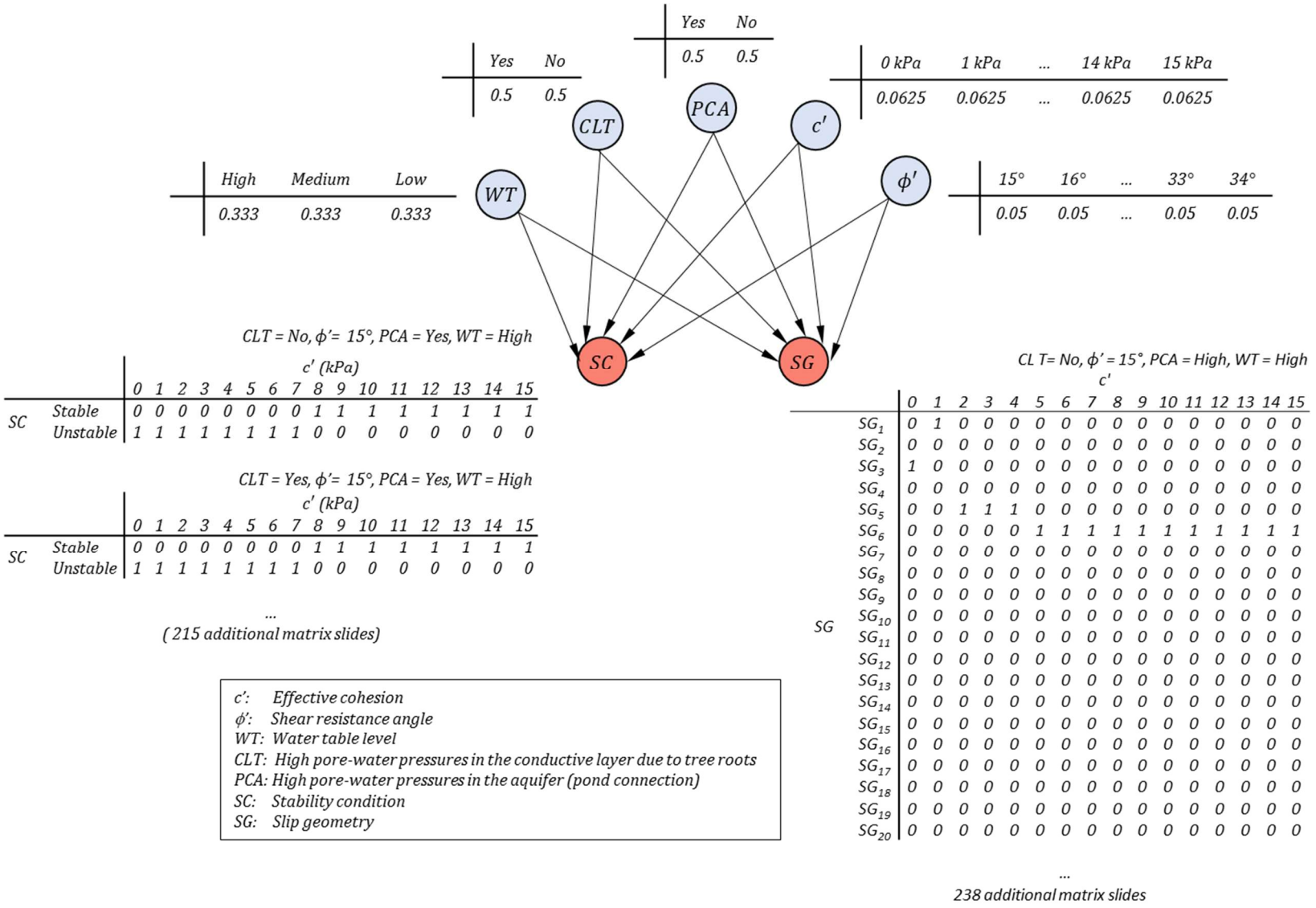


Figure 4-21. Bayesian network for the Breitenhagen levee failure using drained conditions and a Mohr-Coulomb soil constitutive model.

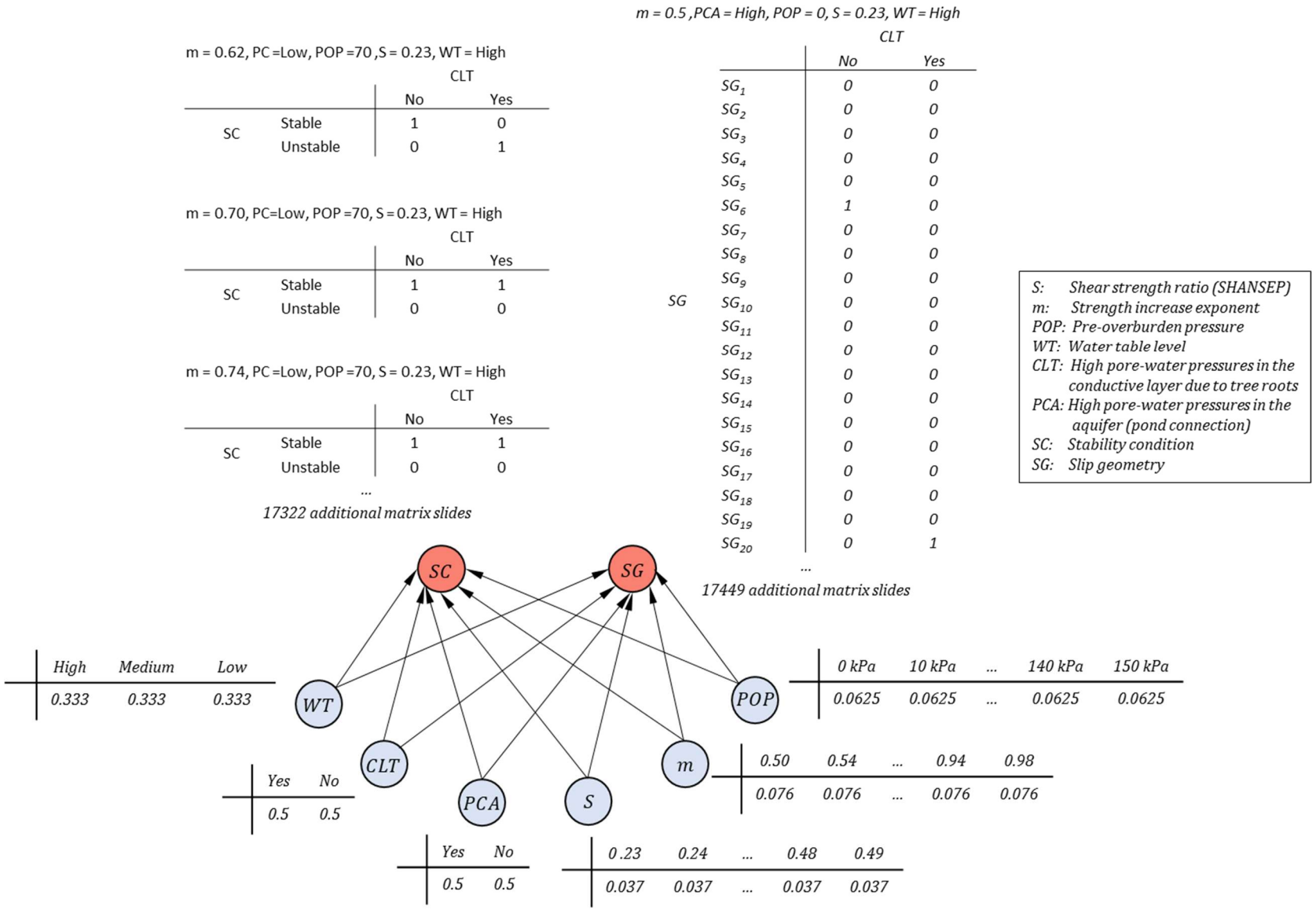


Figure 4-22. Bayesian network for the Breitenhagen levee failure using undrained conditions and SHANSEP constitutive model.

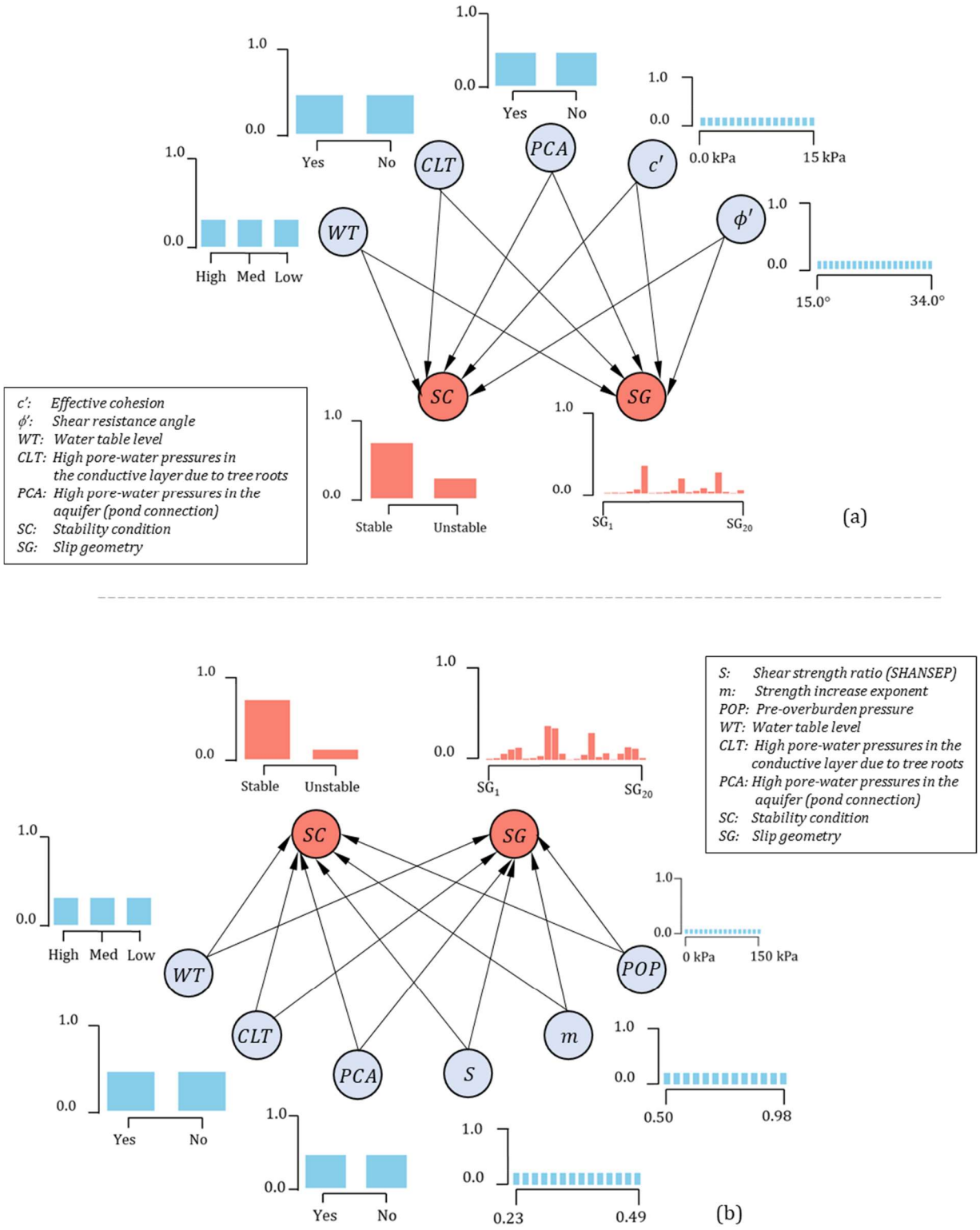


Figure 4-23. Bayesian network for the Breitenhagen levee failure (prior probabilities). (a) Drained condition (Mohr-Coulomb), (b) Undrained condition (SHANSEP).

Hypotheses Comparison Via Probability Queries

To assess the probability of the competing hypotheses H_1 to H_3, the evidence presented in **Table 4-12** is included in the BNs of **Figure 4-23**. The conditional probability of hypotheses H_1 to H_3 given the evidence is estimated using *conditional probability queries*. The combination of states that represent hypotheses H_1 to H_3 are shown in **Table 4-18**. **Table 4-18** also presents additional hypotheses H_4 to H_6 that combines weak soil strength parameters and different pore water pressure conditions. These additional hypotheses are proposed to evaluate the impact of including weak soil strength values on identifying causes of failure. The results of conditional probabilities queries for hypotheses H_1 to H_6 are presented in **Table 4-19**.

Table 4-18. Combination of states in hypothesis nodes for representing failure hypothesis H_1 to H_6.

Hypothesis	Description	Pore water pressure conditions			Drained analysis		Undrained analysis		
		WT	***CLT***	***PCA***	**c′ [kPa]**	**ϕ′ [°]**	***S*** [–]	***m*** [–]	**POP [kPa]**
H_1	Elevation of the water table	High	No	No	4-8	21-28	0.28-0.34	0.90-0.98	10-50
H_2	High pore-water pressures due to tree roots	Medium or Low	Yes	No	4-8	21-28	0.28-0.34	0.90-0.98	10-50
H_3	High pore-water pressures in the aquifer below the levee	Medium or Low	No	Yes	4-8	21-28	0.28-0.34	0.90-0.98	10-50
H_4	Elevation of the water table inside the levee and weak soils	High	No	No	≤ 4	≤ 21	≤ 0.28	≤ 0.90	≤ 10
H_5	High pore-water pressures due to tree roots and weak soils	Medium or Low	Yes	No	≤ 4	≤ 21	≤ 0.28	≤ 0.90	≤ 10
H_6	High pore-water pressures in the aquifer below the levee + weak soils	Medium or Low	No	Yes	≤ 4	≤ 21	≤ 0.28	≤ 0.90	≤ 10

Table 4-19. Probabilities for hypotheses H_1 to H_6 given the evidence $SC = Unstable$ and $SG = sg_6$.

Hypothesis	Description	$P(H_i \mid SC = Unstable, SG = sg_6)$	
		Drained	Undrained
H_1	Elevation of the water table	1.7 x 10^{-2}	4.0 x 10^{-3}
H_2	High pore-water pressures due to tree roots	1.2 x 10^{-2}	0.0
H_3	High pore-water pressures in the aquifer below the levee	0.0	0.0
H_4	Elevation of the water table and weak soil	1.1 x 10^{-1}	4.2 x 10^{-2}
H_5	High pore-water pressures due to tree roots and weak soils	8.7 x 10^{-2}	1.9 x 10^{-2}
H_6	High pore-water pressures in the aquifer below the levee + weak soil	0.0	0.0

According to the probability results of **Table 4-19**, the hypothesis H_4 better explains the failure than the rest of the hypotheses formulated in Section 4.3.2. That is to say, the elevation of the water table inside the levee combined with weak soils are the most probable causes of the Breitenhagen levee failure and its observed slip surface. This result is similar to that found using the POR technique.

Additional Hypotheses comparison using the K Most Probable Explanation (K MPE) Algorithm

The configurations of hypotheses nodes consistent with the observed states in evidence nodes resulted in 241 additional hypotheses (H_7 to H_{248}) that could explain the Breitenhagen levee failure. The probability of each hypothesis is estimated and organized in descending order. **Table 4-20** shows nine additional hypotheses (H_7 to H_{16}) and their probabilities estimated using the K MPE algorithm. The probability results for hypotheses H_{17} to H_{248} are presented in Annex A5.

Notice that the $K = 2$ most probable explanations combine H_1 with low soil strength values in both drained and undrained conditions. The $K = 10$ results also indicate that the failure is not

related to the hypothesis H_3. In conclusion, the combination of a WT elevation due to the sustained high level in the Saala River and the low strength of the levee's soil is the most likely cause of the Breitenhagen levee failure.

Table 4-20. $K = 10$ Most Probable Explanations (K MPE) for the Breitenhagen levee failure.

K-MPE	H_i	Drained Condition (Mohr – Coulomb)						Undrained Condition (SHANSEP)						
		WT	CLT	PCA	c' (kPa)	ϕ' (°)	$P(H_i \mid e)$	WT	CLT	PCA	S	m	POP (kPa)	$P(H_i \mid e)$
1	H_7	High	No	No	1	30	4.33×10^{-3}	High	No	No	0.37	0.74	0	1.49×10^{-3}
2	H_8	High	No	No	1	26	4.30×10^{-3}	High	No	No	0.43	0.54	10	1.48×10^{-3}
3	H_9	Medium	Yes	No	4	21	4.29×10^{-3}	High	Yes	No	0.27	0.58	70	1.42×10^{-3}
4	H_{10}	High	No	No	1	25	4.28×10^{-3}	High	Yes	No	0.41	0.7	0	1.40×10^{-3}
5	H_{11}	Medium	Yes	No	3	26	4.27×10^{-3}	High	Yes	No	0.27	0.82	20	1.39×10^{-3}
6	H_{12}	High	No	No	3	25	4.27×10^{-3}	High	Yes	No	0.35	0.9	0	1.37×10^{-3}
7	H_{13}	High	No	No	1	21	4.26×10^{-3}	High	No	No	0.31	0.7	10	1.35×10^{-3}
8	H_{14}	Medium	Yes	No	3	25	4.26×10^{-3}	High	No	No	0.25	0.9	10	1.31×10^{-3}
9	H_{15}	Medium	Yes	No	3	21	4.25×10^{-3}	High	Yes	No	0.41	0.9	10	1.31×10^{-3}
10	H_{16}	Medium	Yes	No	2	23	4.25×10^{-3}	Medium	Yes	No	0.23	0.54	20	$1.31x\ 10^{-3}$

4.4 Summary

Chapter 4 presents the proposed Bayesian methodology for decision support in FGE. In general, the methodology includes the formulation of multiple hypotheses regarding the cause of a geotechnical failure. The hypotheses are contrasted against the collected evidence using two Bayesian techniques: posterior odds ratio (POR) and Bayesian Networks (BN). The main aspects of the proposed methodology and the examples that confirm its validity can be summarized as follows:

- The proposed Bayesian methodology for decision support in FGE consists of three main stages: (i) Preliminary steps, (ii) construction of a probabilistic failure model, and (iii) probabilistic hypotheses comparison. The first stage focuses on collecting evidence and formulating credible hypotheses about the causes of geotechnical failures. The second stage includes defining relevant random variables and developing a probabilistic failure model. The last stage is the core of the proposed Bayesian

methodology and includes the POR and BN techniques for comparing hypotheses. The result of this comparison is the selection of one or more hypotheses as the most likely cause of failure.

- The following aspects are crucial for successfully applying the proposed methodology: (i) The purpose of collecting evidence is to validate failure hypotheses. The evidence and their analysis should be able to validate or disapprove hypotheses about the causes of failure. Therefore, the collected evidence should be unbiased and objective. (ii) All hypotheses should be based on the predictable behavior of soil/rock materials and expected external/internal forces acting on the geotechnical structure. Semantic expressions (geotechnical jargon) can be used to formulate credible hypotheses. In addition, all credible hypotheses must be able to be translated into a probabilistic model. (iii) Random variables should be selected based on the hypotheses, the influence of each variable on the soil/rock behavior, and the characteristics of the performance variables. (iv) A reliable geotechnical failure model should be capable of including all failure hypotheses defined in the preliminary stage.

- A modified version of the ERTC7 benchmark exercise (Schweiger, 2006) is presented to describe and validate the POR and BNs techniques. The embedded length of the wall is deliberately defined to be short in order to induce a failure in the ERTC7 exercise. Then, both techniques are used to validate their ability to detect the cause of failure. Each stage of the proposed methodology is applied to the modified version of the ERTC7. The results show that the POR and BN techniques predicts the short embedded length of the wall as the most probable cause of excavation failure.

- The applicability of the POR and BN techniques for hypothesis comparison are also explored through the Breitenhagen levee failure (Grubert, 2013; Kool et al., 2019). Four hypotheses related to pore water pressure conditions and soil behavior are formulated as causes of failure. The results show that the POR and BN techniques predict the most probable cause of a failure. Moreover, the results are comparable to those found in previous forensic studies.

5. Determining the Causes of an Excavation Failure. The Green Office Project: A Real Application of Bayesian Methodology

This chapter presents an application of the proposed Bayesian methodology for supporting decisions regarding the causes of geotechnical failures. Posterior odds ratio (POR) and Bayesian networks (BN) techniques described in Chapter 4 have been applied to an excavation failure in Bogotá, Colombia (Unal, 2012). The chapter is divided into two sections. The first section describes the general characteristics of the project and the geotechnical conditions under which the excavation failure was developed. The second section examines the failure using the proposed Bayesian methodology and determines the most probable causes that led to the excavation failure.

5.1 General Description and Geotechnical Conditions

5.1.1 General description

The Green Office Project is a six-level building with three basement levels located on the northeast side of Bogotá, Colombia (**Figure 5-1**). During the construction activities in December 2011, an excavation failure occurred. Heave of the bottom of the excavation, large settlements on 11th street, and significant horizontal movements in the diaphragm wall were observed and monitored. The large settlements impacted vehicular and pedestrian traffic on 11th street and caused the total closure of the road for at least seven months. Although no injuries to workers or neighbors were reported, safety conditions at the workplace and traffic security were significantly impacted.

Figure 5-1. Location of the Green Office project on the northeast side of Bogotá, Colombia. (Modified from Google Earth, 2022).

5.1.2 Geotechnical Characterization of the Site.

Geotechnical Characterization of the Original Design

According to the documentation collected during the forensic investigation, the original geotechnical design included seven borings with depths ranging between 12 m and 49 m. Several Vane Shear Tests (VST), Standard Penetration Tests (SPT), and laboratory tests were performed on disturbed and undisturbed soil samples to characterize the geotechnical behavior of the soils. Unfortunately, the design documents do not provide information from boring logs, field tests, and laboratory tests. Moreover, a detailed geotechnical characterization of the subsoil was not delivered.

The stratigraphy from the original geotechnical design is described in **Table 5-2**a. The stratum M1 is a brown to dark brown soft clay with intercalations of thin sand layers and peat. Stratum M2 is a dark brown very soft clay that extends to 36.0 m depth below the original surface grade. Stratum M3 is a gray medium-dense sand. Finally, a gravelly medium to dense sand with thin

peat layers is located up to 49 m in depth. According to the water levels observed at borehole locations, the original design assumed the water table at a depth of 3.0m.

Geotechnical Characterization Performed for the Forensic Investigation

Subsurface conditions for the forensic investigation were evaluated by drilling six (6) boreholes, two (2) Cone penetrations Tests (CPTu), and six (6) Down Hole tests to depths ranging from 40 m to 50 m. **Figure 5-2** shows the location of these exploration activities. A total of 288 disturbed and undisturbed soil samples were recovered from the boreholes using Shelby tubes (thin-walled, open-tube samplers) to conduct physical and mechanical laboratory tests required for analyses. A summary of the geotechnical properties obtained from laboratory and field tests is shown in **Figure 5-3** and **Figure 5-4**. **Table 5-1** presents the average values of the geotechnical parameters.

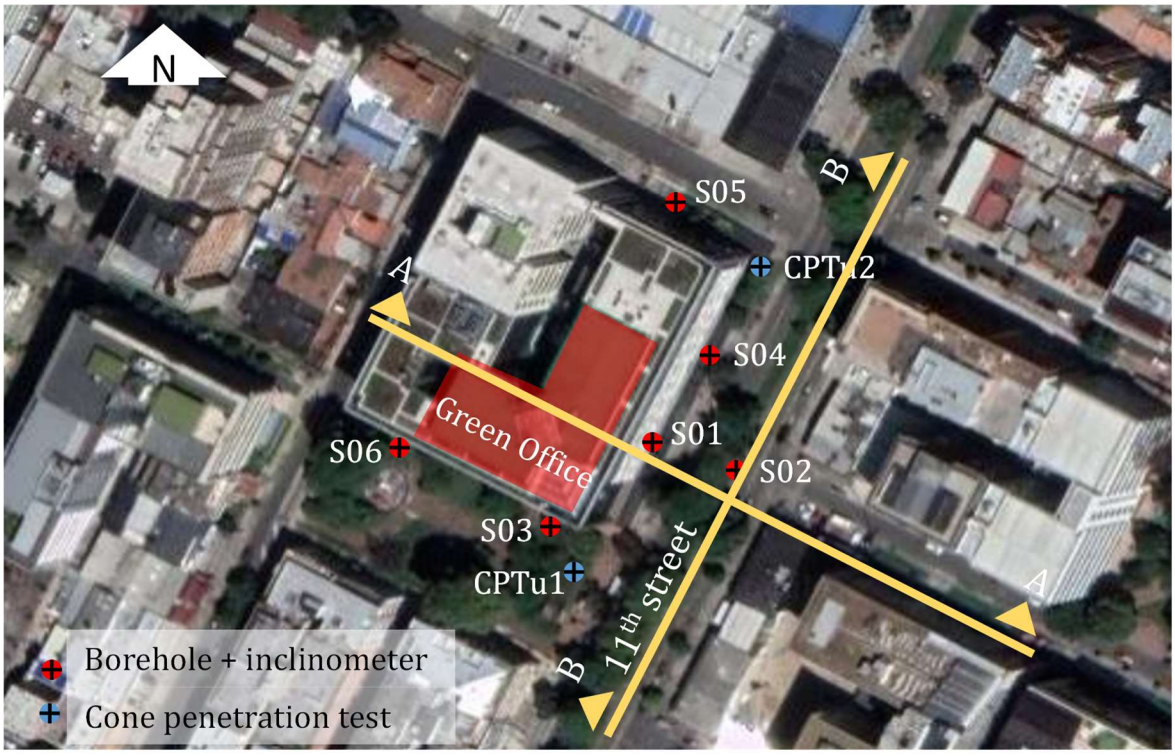


Figure 5-2. Location of the six (6) boreholes and two (2) Cone penetrations Tests (CPTu) developed during the forensic investigation (Modified from Google Earth, 2022 and Unal, 2012).

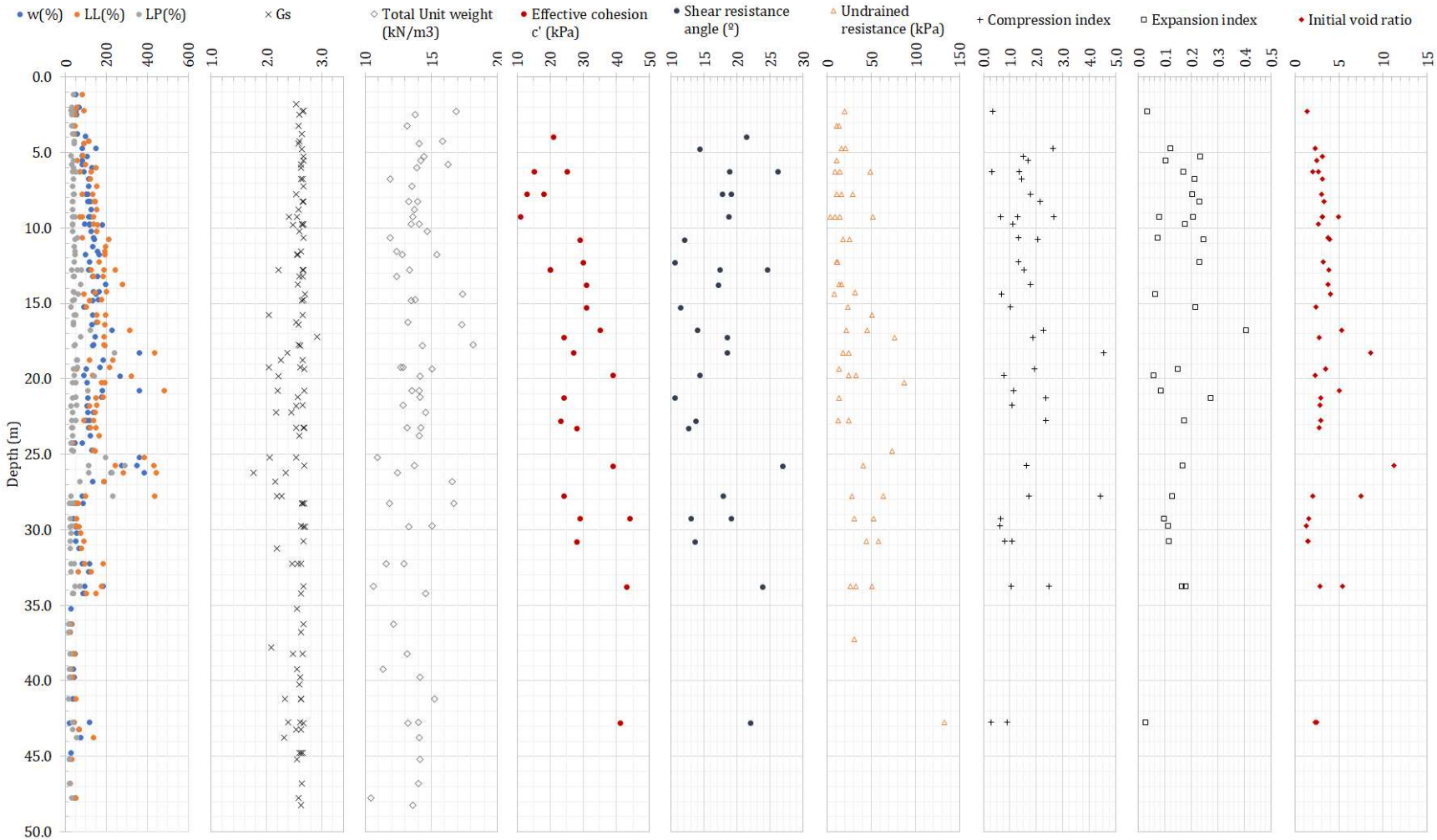


Figure 5-3. Laboratory results for the forensic investigation. Green Office Project (Unal, 2012).

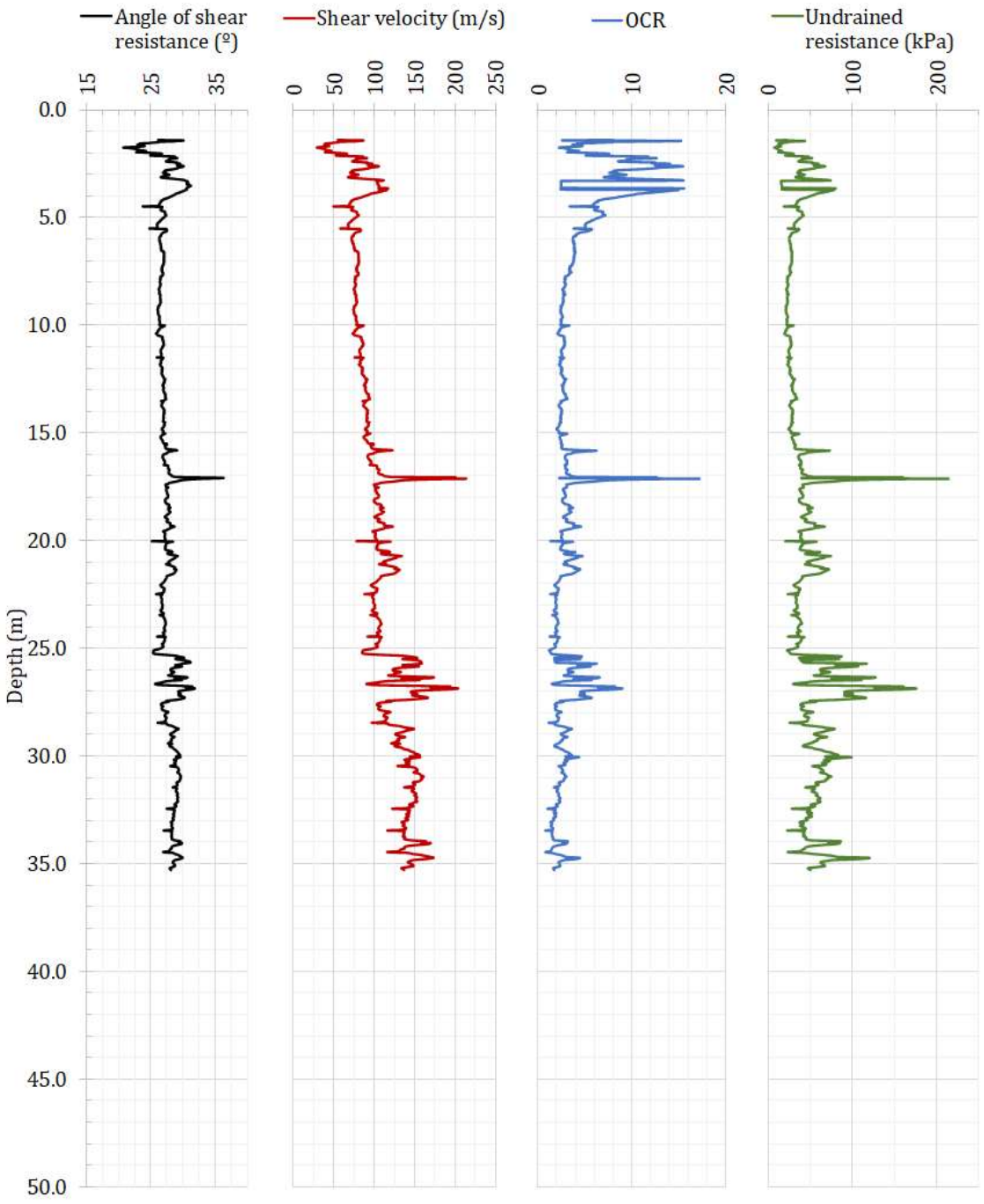


Figure 5-4. Field results from CPTu tests for the forensic investigation. Green Office Project (Unal, 2012).

Table 5-1. Average values of geotechnical parameters.

Stratum	Soil constitutive model	γ_{sat} (kN/m^3)	γ_{uns} (kN/m^3)	c' (kN/m^2)	ϕ' (°)	λ^* (−)	κ^* (−)	e_{ini} (−)
M1	Soft soil	14.8	13.8	5.0	24.0	0.188	0.047	2.0
M2	Soft soil	13.4	12.4	6.3	19.3	0.137	0.045	2.8
M3	Mohr-Coulomb	17.8	16.8	0.0	34.8	-	-	-
M4	Mohr-Coulomb	18.5	17.5	0.0	42.0	-	-	-

γ_{sat}: Saturate unit weight, γ_{uns}: unsaturated unit weight, c': effective cohesion, ϕ': effective angle of shear resistance, λ^*: modified compression index, κ^*: modified swelling index, e_{ini}: initial void ratio.

Comparison of Geotechnical Characterizations. Original Design vs. Forensic Investigation

The stratigraphy inferred from the six borings of the forensic investigation is compared with the original geotechnical design (**Table 5-2**). Notice that differences in depth and soil descriptions were found in both stratigraphies. No geotechnical properties were compared because the information from the original design was unavailable. A simplified stratigraphy inferred from the forensic investigation is presented in **Figure 5-5**.

Table 5-2. Comparison of the stratigraphy obtained from borings in the original geotechnical design (a) and forensic investigation (b).

(a) Original geotechnical design		
Stratum	**Approximate Depth Range of Stratum (m)**	**Description**
M1	0 – 4	Brown to dark brown clay with intercalations of thin sand layers and peat
M2	4 – 36	Dark brown, very soft clay.

(b) Forensic investigation		
Stratum	**Approximate Depth Range of Stratum (m)**	**Description**
M1	0 – 6	Inorganic brown soft clay of high plasticity (MH - CH)
M2	6 – 35	Inorganic brown silts and clays of high compressibility (MH – CH)

(a) Original geotechnical design		
Stratum	**Approximate Depth Range of Stratum (m)**	**Description**
M3	36 – 38	Gray medium-dense sand.
M4	38 - 49	Gravelly medium to dense sand with some thin peat layers.

(b) Forensic investigation		
Stratum	**Approximate Depth Range of Stratum (m)**	**Description**
M3	35 – 40	Dense gray sand (SP)
M4	40 – 50	Possibly a colluvial soil composed of gravel and sand.

5.1.3 Structure, Foundation, and Construction Process

The Green Office Project is an L-shaped building 23.5 m high (six floors above the ground surface) and 11 m deep (three basements). The structure at the basements consists of a portal frame that includes circular columns of 0.8 m in diameter, which are spaced 10.0 m in both directions, and rectangular beams with dimensions 0.4 x 0.7 m. The beams support several I-shaped steel beams and a composite metal floor deck. The structure above the ground surface consists of I-shaped steel columns supporting a composite metal floor deck. Each column is supported by three piles connected by a pile cap. The piles are 29.0 m deep and 0.8 m in diameter. Additionally, a 17 m deep and 0.4 m thick diaphragm wall provides lateral support to the 11 m deep excavation for the basements. A simplified scheme of the structure and diaphragm wall is shown in **Figure 5-5**.

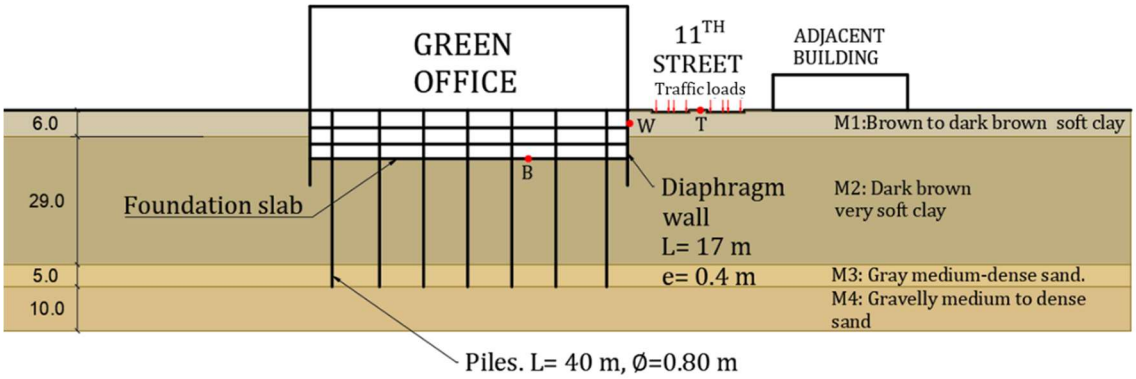


Figure 5-5. Stratigraphy and simplified scheme of the Green Office Project.
(Cross section A-A from Figure 5-2).

According to the construction logbook, the deep excavation and diaphragm wall included the following construction steps:

i. Construction of the piles from the -11.0 m level to the -40.0 m level.
ii. Construction of the diaphragm wall panels: this step included the construction of guided walls, trench excavation, steel placing, and concrete casting.
iii. Construction of temporary caisson piles from the 0.0 m level to the -11.0 m level and concrete beams for the first floor. According to the design, the concrete beams act as struts and provide additional lateral support. At this step, the construction of the superstructure started.
iv. Excavation of the whole first basement down to the -4.0 m level.
v. Construction of concrete beams for the first basement.
vi. Excavation of the second basement down to the -7.70 m level. At this point, construction managers decided to excavate areas of 20 m x 20 m. In each area, the beams and the second basement slab were constructed.
vii. Excavation of the third basement: following the same procedure as the second basement.
viii. Construction of foundation slabs in every 20 m x 20 m area.

5.1.4 Monitoring System and Failure Description

Description of the monitoring system

Eight inclinometers and five Casagrande piezometers were installed during the construction activities. Six inclinometers were located within the diaphragm wall, and two inclinometers were installed in the adjacent area. The total length of the inclinometers ranged between 17 m and 26 m, whereas the piezometers were 12 m in length. **Figure 5-6** shows the location of the inclinometers and piezometers installed before the construction stage.

The forensic investigation included the installation of six additional inclinometers at the locations of boreholes shown in **Figure 5-2**. These additional inclinometers aimed to verify the measurements and inclination rates provided by the constructor. Additionally, during the drilling activities for the forensic investigation, several water table levels were recorded. Finally, a conventional topographic survey was carried out on the adjacent area to verify settlements and heaves reported by the constructors.

Figure 5-6. Location of inclinometers and Casagrande piezometers installed before the construction stage (Modified from Google Earth 2022).

Failure Process and Measurements from the Monitoring System

During the construction and forensic investigation, a topographic survey was carried out to identify the critical cross-section A-A' shown in **Figure 5-2** and **Figure 5-5**. The chronological analysis of the A-A cross-section revealed a continuous settlement process. At the early stages of the process, small settlement values due to the first basement excavation (-4.0 m level) were observed. Then, large settlements up to 0.90 m were measured adjacent to the excavation for basements two and three. In addition, heave of the base of the excavation was recorded during construction activities for the first basement, and significant heave values up to 0.4 m were measured for basements two and three.

The measurements from the 14 inclinometers were used to identify critical deformations on the diaphragm wall and the surrounding area. Inclinometers located at positions S01 and S02, shown in **Figure 5-2**, revealed large horizontal displacements. As expected, the location of the critical cross-section A-A' coincides with the location of the critical inclinometers. In other words, large settlements on 11th street, significant heave values at the base of the excavation,

and large horizontal displacements in the surrounding area indicate a failure process through cross-section A-A'.

Regarding the elevation of the water table (WT), **Figure 5-22** presents the monthly rainfall recorded by a nearby weather station from 2007 to 2011, and **Figure 5-23** shows the variation of the WT during the construction process. Additionally, **Table 5-3** presents the values of the water table levels recorded during the exploration work for the forensic investigation stage. In both cases (construction and forensic stages), there is a positive correlation between antecedent rainfall and the WT levels.

Table 5-3. Levels of the water table during the forensic investigation stage (January 2012).

Borehole	Total depth(m)	Water table levels (m)
S01	50	-1.75
S02	40	-1.60
S03	50	-2.30
S04	45	-1.70
S05	45	-1.60
S06	45	-0.50

5.2 Application of the Bayesian Methodology

5.2.1 Stage 1: Collected Evidence and Hypotheses Formulation

Collected Evidence

Table 5-4 summarizes the evidence collected from design documents, construction logbooks, and geotechnical characterization during the forensic investigation. The evidence is used to develop the geotechnical model and evaluate the failure hypotheses. Additional evidence is presented and discussed in Section 5.2.3.

Table 5-4. Summary of collected evidence: geometry, loads, monitoring, groundwater conditions, and failure mechanism.

Category	Symbol	Value	Units	Evidence Description
Geometry	L	17.0	m	The total length of the diaphragm wall
	E	6.0	m	Embedded length of the diaphragm wall
	T	0.4	m	Diaphragm wall thickness
	D_T	11.0	m	The total depth of excavation
	D_1	4.0	m	Depth of excavation for basement 1
	D_2	7.7	m	Depth of excavation for basement 2
	D_3	11.0	m	Depth of excavation for basement 3
	B	72.0	m	Total width of the excavation
	H_B	23.5	m	Height of the building at failure
Loads	q_B	21.0	kN/m^2	Estimated load of the superstructure at failure
	q_F	8.0	kN/m^2	Estimated load of the foundation slab at failure
	q_b	6.0	kN/m^2	Estimated load of each basement slab at failure
	q_T	90	kN	Traffic Loads (point loads)
	q_{AB}	20	kN/m^2	Estimated load of the adjacent building (2 floors)
Measurements from monitoring	ρ_{max}	0.98	m	Maximum settlement observed on 11th street
	i_{max}	0.19	m	Maximum inclination observed in the diaphragm wall
	h_{max}	0.50	m	Maximum heave observed at the base of the excavation
Water Table	WT	-3.0	m	Level of the water table at the time of failure
Failure mechanism	L_{CR}	19.0	m	Distance from the diaphragm wall to the main crown crack
	SS_s	-	-	Observed slip surface shape: circular

Hypotheses about the causes of failure

Since the first evidence of failure, several stakeholders, such as owners, construction contractors, public utility companies, local government authorities, and the affected community, have shown an interest in determining the causes that led to large deformations and the subsequent excavation failure. Stakeholders formulated numerous hypotheses about

the causes of failure. For example, building owners and construction contractors stated that the excavation failure occurred only as a result of an unexpected elevation in the water table level due to damage to a sewer pipe near the building. On the other hand, public utility companies and local government authorities claimed that the failure was due to design shortcomings. In particular, they argued that the thickness and the embedded length of the diaphragm wall were inadequate to support lateral earth pressures and a possible elevation of the water table to the 0.0 level. Some argued, even further, that the original design was not even appropriate for supporting lateral pressures with a water table at the -3.0 level. In summary, the hypotheses formulated by the stakeholders are described below:

- Hypothesis H_1: A diaphragm wall 0.4 m thick and 17.0 m long combined with a water table at the -3.0 m level led to the excavation failure and the displacements recorded at points T, W, and B (refer to **Figure 5-5**). The hypothesis H_1, proposed by public utility companies, local government authorities, and the community, implies that the original design was unable to support lateral earth pressures.

- Hypothesis H_2: A diaphragm wall 0.8 m thick and 25.0 m long combined with a water table at the -3.0 level led to the excavation failure and the displacements recorded at points T, W, and B (refer to **Figure 5-5**). The hypothesis H_2 is a counterfactual design condition included only for comparison purposes. It provides information about what would have happened if a more robust design, such as a diaphragm wall 0.8 m thick and 25.0 m long, had been implemented.

- Hypothesis H_3: A diaphragm wall 0.4 m thick and 17.0 m long combined with a water table at the ground surface led to the excavation failure and the displacements recorded at points T, W, and B (refer to **Figure 5-5**). The hypothesis H_3 proposed by the owner and construction contractors implies that the elevation of the water table alone led to the excavation failure.

- Hypothesis H_4: A diaphragm wall 0.8 m thick and 25.0 m long combined with a water table at the ground surface led to the excavation failure and the displacements recorded at points T, W, and B (refer to **Figure 5-5**). Hypothesis H_4 is also a

counterfactual design condition. It provides information about what would have happened if a more robust design had been implemented.

The values of the random variables for hypotheses H_1 to H_4 are summarized in **Table 5-5**.

Table 5-5. Green Office Project. Summary of values for failure hypotheses.

Variable	Hypothesis H_i			
	H_1	H_2	H_3	H_4
Total length of the wall W_L (m)	17.0 m	25.0 m	17.0 m	25.0 m
Thickness of the wall W_T (m)	0.4 m	0.8 m	0.4 m	0.8 m
Water table position WT (m)	-3.0 m	-3.0 m	0.0 m	0.0 m

5.2.2 Stage 2: Random Variables and Probabilistic Failure Model

According to the criteria presented in Section 4.1.2, the selected random variables for hypotheses comparison are the total length of the wall L, the thickness of the wall T, and the water table WT elevation. For simplicity, L and T are merged into one new variable LT with two states $LT = \{LT_1, LT_2\}$. The state LT_1 is the original geotechnical design (LT_1: $L = 17.0\ m$, $T = 0.4\ m$) and the state LT_2 is the robust counterfactual design (LT_2: $L = 25.0\ m$, $\mathrm{T} = 0.8\ m$) suggested by an experienced geotechnical designer. **Figure 5-7**a shows the probability mass function (PMF) for the wall variable LT in which both states $LT_1$and LT_2 are equally probable.

Similar to LT, two equally probable states are assigned to the variable water table elevation $WT = \{WT_{0.0}, WT_{-3.0}\}$. Each element represents the water table at the 0.0 m and -3.0 m levels, respectively. **Figure 5-7**b shows the PMF for the WT variable. For the sake of simplicity of the probability model, $LT = \{LT_1, LT_2\}$ and $WT = \{WT_{0.0}, WT_{-3.0}\}$ comprise the total sample space of the model.

In order to construct a complete probability failure model for the Green Office Project, the forensic analysis requires additional random variables. In the case of the strength limit state, variables $c'$and ϕ', for strata M1 and M2 are defined as random variables described by

truncated normal distributions with the mean and coefficient of variation (COV) values presented in **Table 5-6** (Phoon & Kulhawy, 1999a). For strata M3 and M4, $c'$and ϕ' values are defined as deterministic due to the considerable depth of the strata and their little influence on the excavation failure (refer to **Table 5-1**). Deterministic values also describe the unit weights of soils for all strata.

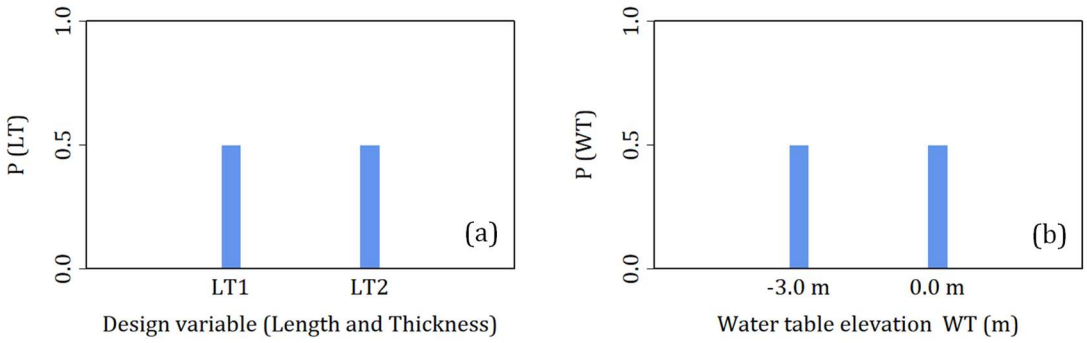


Figure 5-7. Probability mass functions (pmf) for (a) design variable and (b) Water table elevation.

In the case of service limit state analysis, variables λ^*, κ^*, and e_{ini} for strata M1 and M2 are defined as random variables due to their influence on deformations of the excavation. Truncated normal distributions also describe these random variables with the mean and COV values of **Table 5-1** (Phoon & Kulhawy, 1999a). Deformation variables for strata M3 and M4 are described by the deterministic values presented in Annex A7. Probability density functions for all random variables are shown in **Figure 5-8**.

Table 5-6. Probability functions for random variables of the Green Office Project.

Component	Variable	Symbol	Unit	Variable type	Probability distribution	Parameters*
Stratum M1	Effective cohesion	c'_{M_1}	kPa	Continuous	Normal truncate	Mean: 5.0 kPa COV = 10% LB =3.5 kPa UB = 6.5 kPa
	Effective angle of shear resistance	ϕ'_{M_1}	°	Continuous	Normal truncate	Mean: 24.0 ° COV = 10% LB =16.8 ° UB = 31.2 °

Component	Variable	Symbol	Unit	Variable type	Probability distribution	Parameters*
	Modified compression index	$\lambda^*_{M_1}$	-	Continuous	Normal truncate	Mean = 0.188 COV = 10% LB =0.132 UB =0.244
	Modified swelling index	$\kappa^*_{M_1}$	-	Continuous	Normal truncate	Mean: 0.047 COV = 10% LB = 0.033 UB = 0.061
	Initial void ratio	$e_{ini_{M_1}}$	-	Continuous	Normal truncate	Mean: 2.0 COV = 7% LB = 1.6 UB = 2.4
Stratum M2	Effective cohesion	c'_{M_2}	kPa	Continuous	Normal truncate	Mean: 6.3 kPa COV = 10% LB =4.4 kPa UB = 8.2 kPa
	Effective angle of shear resistance	ϕ'_{M_2}	°	Continuous	Normal truncate	Mean: 19.3 ° COV = 10% LB =13.5 ° UB = 25.1 °
	Modified compression index	$\lambda^*_{M_2}$	-	Continuous	Normal truncate	Mean = 0.137 COV = 10% LB =0.096 UB =0.178
	Modified swelling index	$\kappa^*_{M_2}$	-	Continuous	Normal truncate	Mean: 0.0457 COV = 10% LB = 0.0320 UB = 0.0594
	Initial void ratio	$e_{ini_{M_2}}$	-	Continuous	Normal truncate	Mean: 2.8 COV = 10% LB = 1.9 UB = 3.6
Diaphragm Wall	Total length and thickness	LT	-	Categorical	Categorical	k = 2 $LT = \{LT_1,\ LT_2\}$ $P(LT_1) = 0.5$ $P(LT_2) = 0.5$
Water Table Level	Water Table	WT	-	Categorical	Categorical	k = 2 $LT = \{LT_1,\ LT_2\}$ $P(WT_1) = 0.5$ $P(WT_1) = 0.5$

*COV: Coefficient of variation, LB: lower bound, UP: Upper bound, k: number of categories.

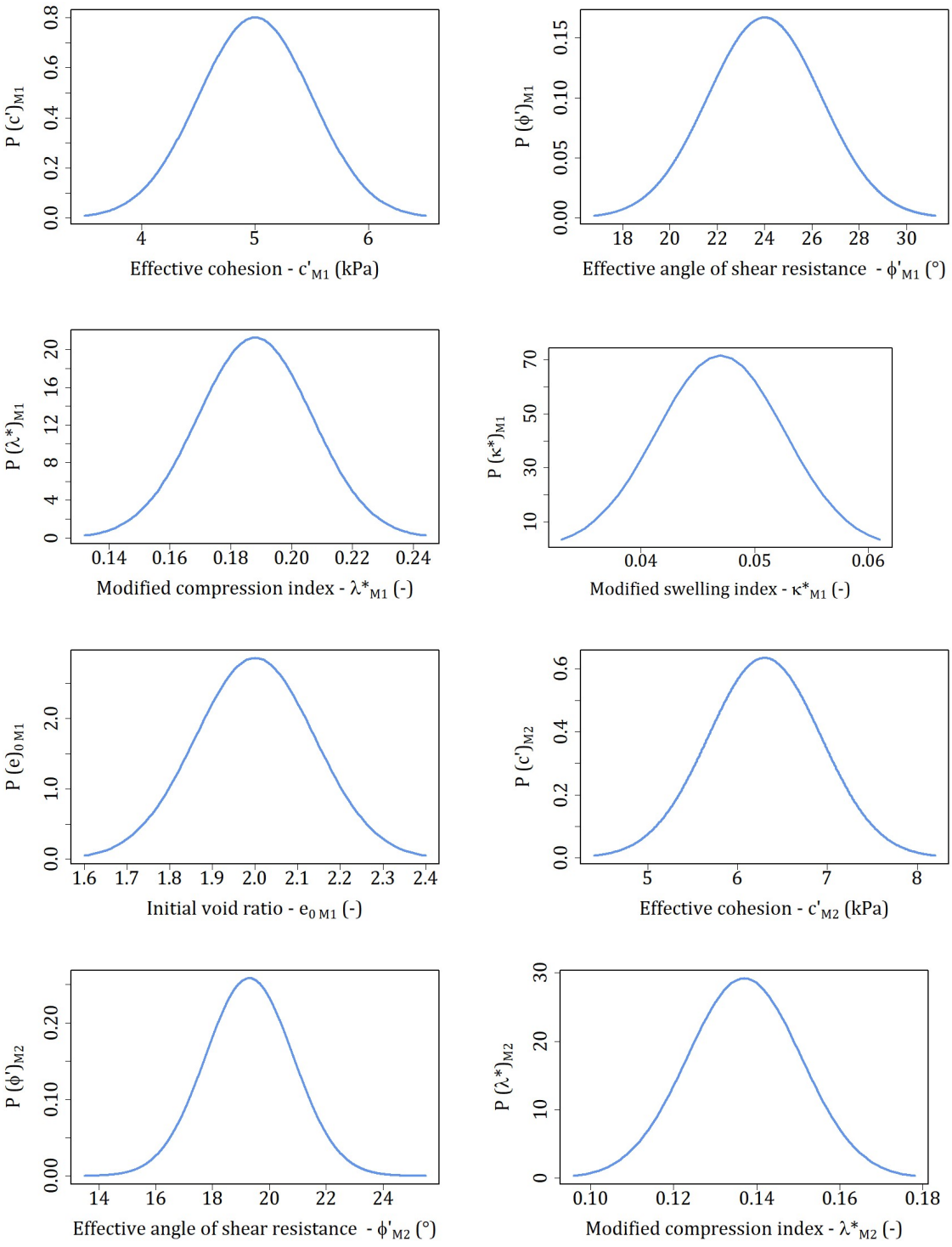


Figure 5-8. Probability density functions (pdf) for random variables of the Green Office Project.

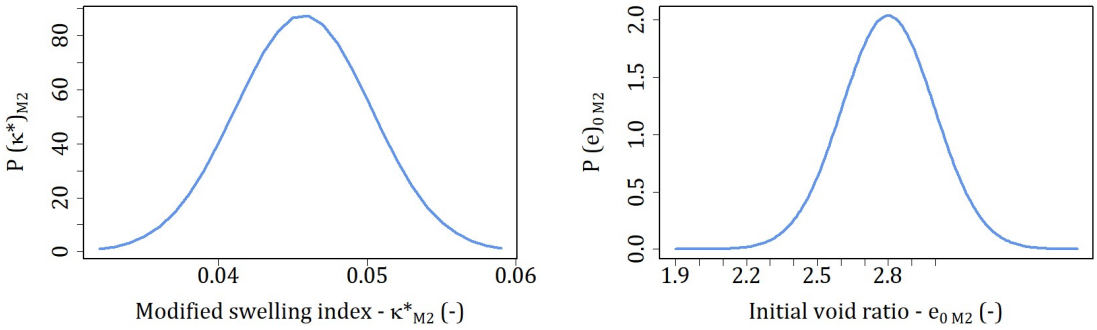


Figure 5-8 (continued). Probability density functions (pdf) for random variables of the Green Office project.

Two probabilistic failure models are constructed to determine the causes of the excavation failure. The first model uses limit equilibrium models via Janbu's corrected equations to analyze the strength limit state. The limit equilibrium model compares the available strength in the soil against the stresses imposed by the loads or excavation processes via the Factor of Safety (FoS). A FoS value smaller than 1.0 indicates an unstable condition, whereas a FoS value greater than 1.0 implies a stable condition. The probabilistic model for the strength limit state implements $N = 8{,}000$ computational experiments of the geotechnical model presented in **Figure 5-5** with the random variables presented in **Table 5-6**. The Python script presented in Annex A6 estimates the FoS through Slide v5.0 (Rocscience Inc., 2006). **Figure 5-9** shows the FoS histograms for designs LT_1 and LT_2, and their corresponding fitted lognormal probability distributions. **Figure 5-10** presents the convergency plot for the FoS and the standard deviation of the FoS.

The second probabilistic model uses a Finite Element Model (FEM) to evaluate the service limit state. A Python routine implements $N = 8{,}000$ computational experiments to calculate deformations using Plaxis V20 (Bentley Systems, 2020). Deformations are estimated on points *T, W,* and *B,* located on 11th street, the diaphragm wall, and at the bottom of the excavation, respectively (refer to **Figure 5-5** for points location).

The FEM comprised the soil stratigraphy and geotechnical parameters outlined in **Table 5-1**, the structural components along with the point/linear loads specified in **Table 5-4**, the geometry illustrated in **Figure 5-5**, and the construction sequence detailed in section **5.1.3**. The model employed 1121 finite elements and 9472 nodes, with increased refinement around

the excavation area to achieve more precise deformation outcomes. All components of the FEM were constructed using the Python Application Programming Interface (API) included in Plaxis, and the repetitive calculations were implemented in the Python script presented in Annex A.7. **Figure 5-11** displays the flowchart that served as the foundation for the Python script. This script was employed to derive the deformation results of the model by incorporating the random variables listed in **Table 5-6**.

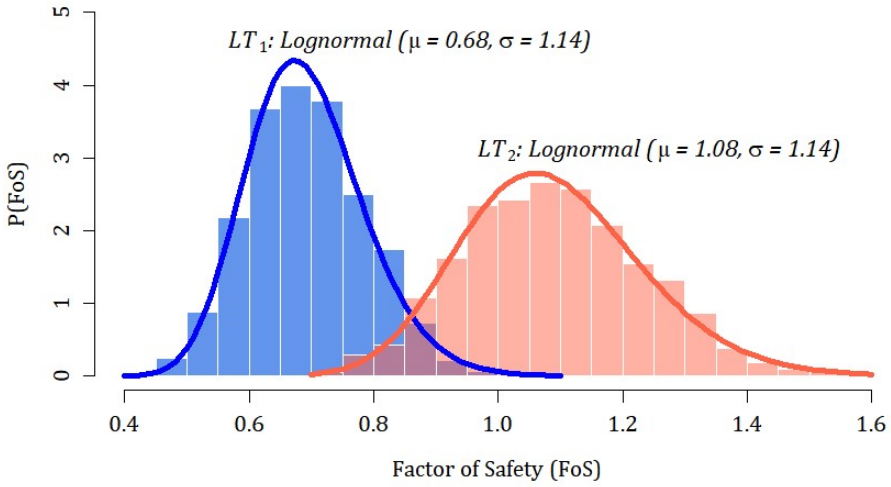


Figure 5-9. Green Office Project: Histograms of the Factor of Safety (FoS).

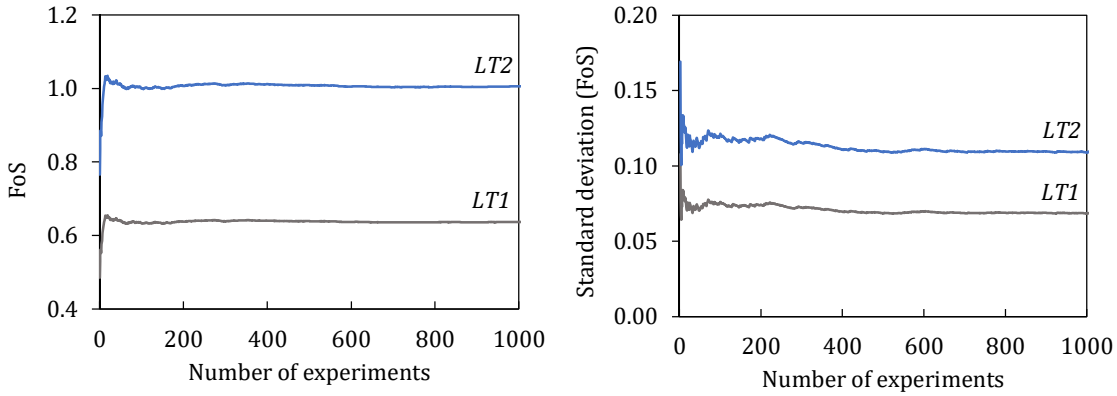


Figure 5-10. Convergence plots for FoS and standard deviation (FoS).

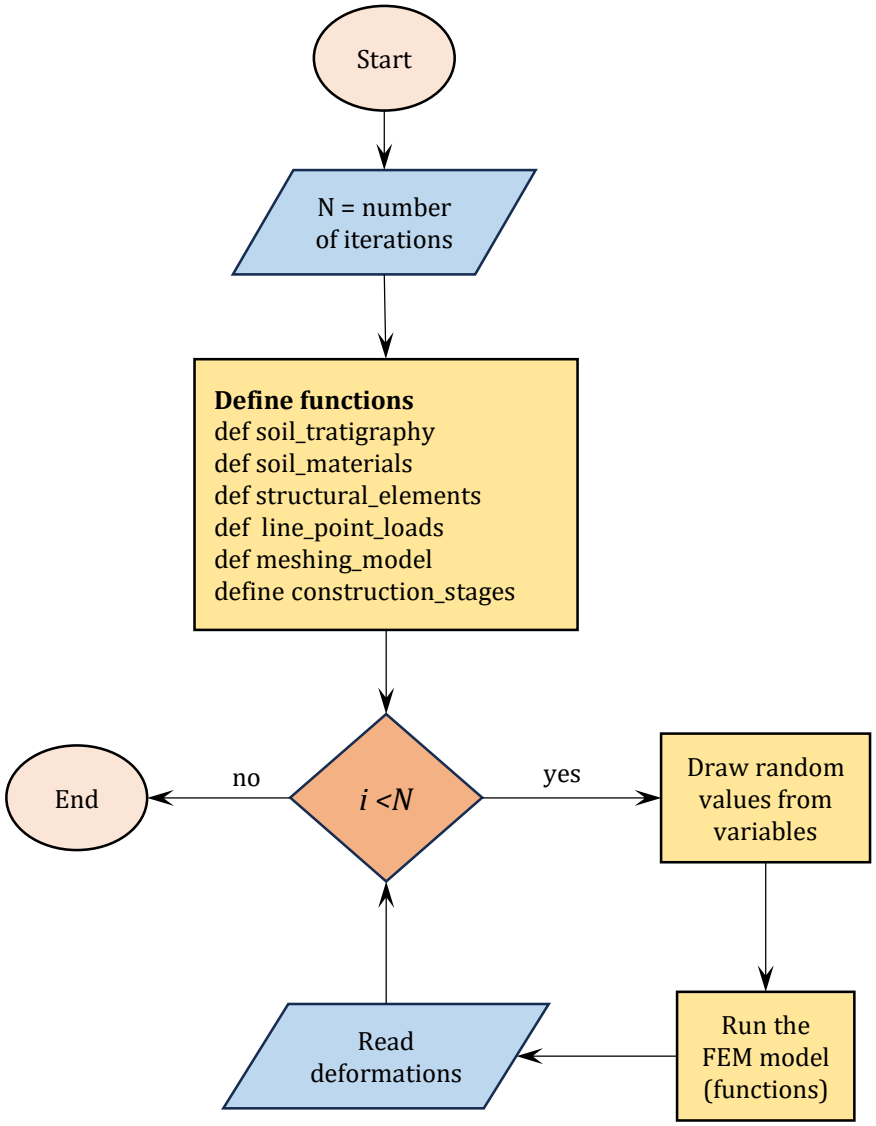


Figure 5-11. Flow chart used to construc the Python script (Green office FEM).

5.2.3 Stage 3. Bayesian Hypotheses Comparison

Hypotheses H_1 to H_4 are compared using the Bayesian Methodology proposed in Chapter 4. The posterior odds ratio (POR) and the Bayesian network (BN) technique are applied to the Green Office Project following the procedure described in Section 4.1.3. and **Figure 4-1**.

Posterior Odds Ratio

The results from the computational experiments are used to calculate the POR based on the steps described in Section 4.13. The first step estimates the probability of observing the evidence given that a hypothesis is assumed to be true (i.e., the likelihood term in Equation

2-6). **Table 5-7** presents different amounts of evidence to verify their influence on the likelihood term and the failure hypotheses. For example, Equation **5-1** shows several likelihood terms related to evidence e_3. The first term is read as the probability of simultaneously observing an unstable condition, a settlement on 11th street greater than 0.30 m, an inclination of diaphragm wall greater than 0.20 m, and a heave at the base of the excavation higher than 0.30 m, given that hypothesis H_1 is true (i.e., H_1: a diaphragm wall 0.4 thick and 17.0 m long combined with a water table at the -3.0 level led to the excavation failure and the displacements recorded at points T, W, and B.). **Table 5-8** shows the calculated likelihood values for all H_i and $e_k$combinations.

Table 5-7. Different amounts of evidence for the Green Office forensic analysis.

Amount of evidence (e_k)	**Description**
e_1	$FoS \leq 1.0$ (Unstable excavation)
e_2	$FoS \leq 1.0$, $\rho > 0.3$ m (Simultaneously observing an unstable condition of the excavation and a settlement on 11th street greater than 0.3 m)
e_3	$FoS \leq 1.0$, $\rho > 0.3$ m , i > 0.2 m , h > 0.3 m (Simultaneously observing an unstable condition of the excavation, a settlement on 11th street greater than 0.3 m, an inclination of diaphragm wall greater than 0.2 m, and a heave at the bottom of excavation greater than 0.3 m)

$$L_{(e_{3|H_1})} = P(e_3 \mid H_1)$$
$$L_{(e_{3|H_2})} = P(e_3 \mid H_2)$$
$$L_{(e_{3|H_3})} = P(e_3 \mid H_3)$$
$$L_{(e_{3|H_4})} = P(e_3 \mid H_4) \quad (5\text{-}1)$$

Table 5-8. Estimated Likelihood values for all H_i and e_k combinations.

Amount of evidence ($\mathbf{e_k}$)	**Likelihood**			
	$\boldsymbol{P(e_k \mid H_1)}$	$\boldsymbol{P(e_k \mid H_2)}$	$\boldsymbol{P(e_k \mid H_3)}$	$\boldsymbol{P(e_k \mid H_4)}$
e_1	1.0×10^{0}	9.3×10^{-2}	1.0×10^{0}	4.8×10^{-1}
e_2	1.0×10^{0}	2.0×10^{-3}	1.0×10^{0}	4.8×10^{-1}
e_3	1.0×10^{0}	2.0×10^{-4}	4.7×10^{-1}	2.0×10^{-4}

The second step of the posterior odds technique defines the prior odds for each pair of hypotheses. The prior odds are initially specified as $P(H_i/H_j) = O(H_{ij}) = 1.0$ (or 1:1) for all hypotheses combinations. In other words, all hypotheses are equally likely to be true before any data or evidence is observed. A 1:1 prior odds value guarantees no bias for any of the hypotheses formulated as causes of failure. However, different prior odds values can be evaluated to verify the influence of any bias about the causes of failure. For the Green Office Project, the prior odds values of 1:1 (hypotheses H_i and H_j are equally likely), and 10:1 (hypothesis H_i is ten times more probable than hypothesis H_j) are evaluated.

The third and final step is comparing pairs of hypotheses through **Equation 5-2**. The comparison includes all possible hypotheses combinations, the differing amounts of evidence established in **Table 5-7**, and the prior values defined above. The results of the Bayesian hypotheses comparison are summarized in **Figure 5-12** and **Table 5-9**.

$$\frac{P(H_i \mid e_k)}{P(H_j \mid e_k)} = \frac{P(e_k \mid H_i)}{P(e_k \mid H_j)} \frac{P(H_i)}{P(H_j)} \tag{5-2}$$

Table 5-9. Bayesian comparison for hypotheses pairs $(\mathrm{H}_i - \mathrm{H}_j)$ for different amounts of evidence (e_1, e_2, e_3). Prior odds values 1.0 (1:1) and 10 (10:1). Color code identifies the posterior odds interpretation according to the Kass & Raftery (1995) criteria in Table 2-1 and Figure 2-2. Color convention: red: "Not worth more than a bare mention," orange: positive, yellow: strong, green: very strong.

(a) Prior odds 1.0 (1:1)
Evidence e_1

H_j \ H_i	H_1	H_2	H_3
H_1	-	-	-
H_2	10.8	-	-
H_3	1.0	0.09	-
H_4	2.1	0.2	2.1

(b) Prior odds 1.0 (1:1)
Evidence e_2

H_j \ H_i	H_1	H_2	H_3
H_1	-	-	-
H_2	500	-	-
H_3	1.0	0.02	-
H_4	2.1	0.04	2.1

(c) Prior odds 1.0 (1:1)
Evidence e_3

H_j \ H_i	H_1	H_2	H_3
H_1	-	-	-
H_2	5000	-	-
H_3	2.1	0.0004	-
H_4	5000	1.0	2385

(d) Prior odds 10 (10:1)
Evidence e_1

H_j \ H_i	H_1	H_2	H_3
H_1	-	-	-
H_2	108	-	-
H_3	10.0	0.9	-
H_4	21.0	2.0	21.0

(e) Prior odds 10 (10:1)
Evidence e_2

H_j \ H_i	H_1	H_2	H_3
H_1	-	-	-
H_2	5000	-	-
H_3	10.0	0.2	-
H_4	21.0	0.4	21.0

(f) Prior odds 10 (10:1)
Evidence e_3

H_j \ H_i	H_1	H_2	H_3
H_1	-	-	-
H_2	50000	-	-
H_3	21.0	0.004	-
H_4	50000	10.0	23850

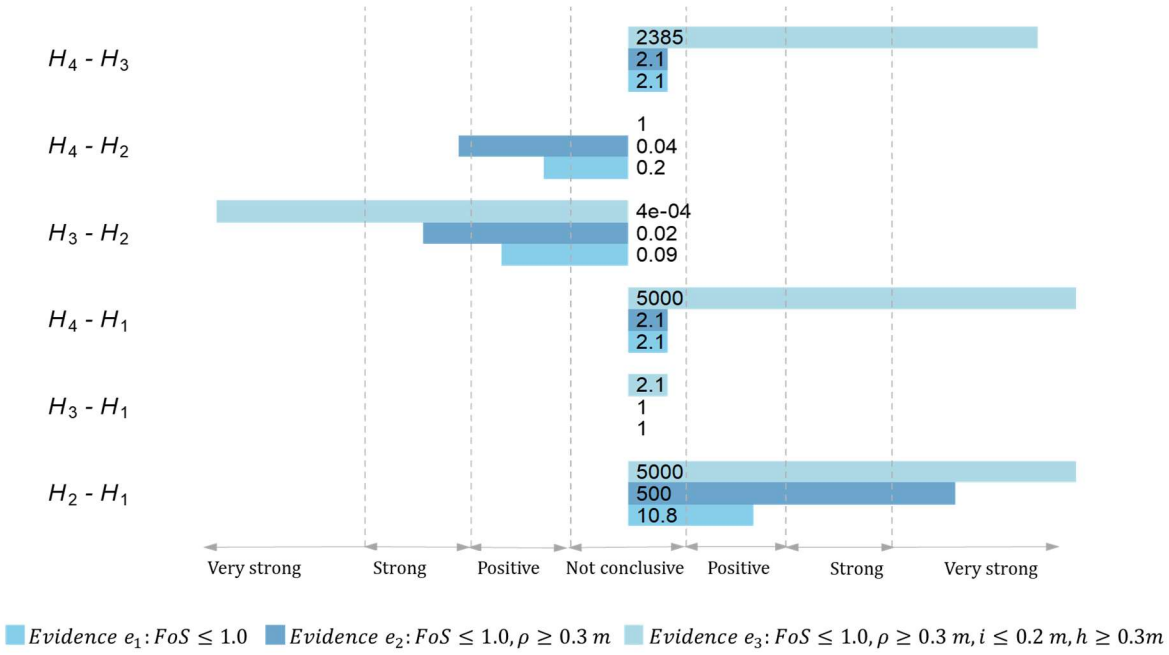


Figure 5-12. Pairwise Bayesian comparison summary for H_i *vs* H_j. Prior odds $O(H_{ij}) = 1.0$ (1:1).

From the analysis of **Figure 5-12** and **Table 5-9**, the following conclusions are drawn:

For prior odds $O(H_{ij}) = 1.0$ (1:1)

- When only the amount of evidence e_1 is included in the analysis, the posterior odds results show that a failure due to hypothesis H_1 is 10.8 times more likely than a failure due to hypothesis H_2 (refer to **Table 5-9**(a)). In other words, there is positive evidence in favor of a failure due to shortcomings of the geotechnical design combined with a water table at the -3.0 m level (H_1 vs H_2). Similarly, there is positive evidence in favor of a failure due to shortcomings of the geotechnical design combined with a water table at the ground surface when compared to the suggested design (H_3 vs H_2). However, there is no conclusive evidence suggesting that shortcomings of the design and a water table at the ground surface caused the excavation failure (H_3 vs H_4).

- When the amount of evidence e_2 is included, results show that hypothesis H_1 is 500 times more likely than hypothesis H_2 (H_1 vs H_2, refer to **Table 5-9**b). That means that

shortcomings in the design combined with a water table at -3.0 level are the most probable cause of excavation failure. Similar conclusions can be drawn when hypotheses H_3 and H_4 are compared to hypothesis H_2. In these cases, shortcoming of the design and a water table at the ground surface are the most likely causes of failure.

- When all available evidence is included in the analysis (i.e., evidence e_3 is included, refer to **Table 5-9**c), conclusions about the most probable causes of failure reaffirm the results obtained for evidence e_2. In conclusion, comparison of hypotheses H_1 - H_2 and H_3 - H_4 shows that shortcomings in the design (i.e., hypotheses H_1 and H_3) are the most probable causes of excavation failure.

The above results show the relevance of including all available evidence in forensic analyses. For example, evidence e_1 shows slight support for hypotheses H_1 as the cause of failure. However, when all available evidence e_3 is included, hypotheses H_1 and H_3 are strongly favored.

For prior odds $O(H_{ij}) = 10$ (10:1)

When prior odds are defined as $O(H_{ij}) = 10$ (10: 1), similar conclusions to those obtained in the $O(H_{ij}) = 1.0$ analysis can be inferred. Although the strength of the evidence is different (notice the variation in color codes), conclusions about the most probable causes that led to the excavation failure are similar to $O(H_{ij}) = 1.0$. In other words, for the Green Office Project, the evidence and the geotechnical model are sufficient to reach the same conclusions, even if there are biases in favor of some hypotheses.

Bayesian Networks

A BN for the forensic analysis of the Green Office excavation failure is proposed in this subsection. The BN is constructed based on the steps described in Section 4.1.3 and the properties presented in Section 2.2.7. The advantage of the BN approach is its ability to expand forensic analyses by including additional nodes (variables) related to evidence, expert opinion, and common sense. In addition, BNs allow for the exploration of several hypotheses not included in the first hypotheses formulation. The results from the computational numerical experiments are also used in the BN.

The forensic analysis of the Green Office Project uses geotechnical idioms to construct small fragments of DAGs (refer to Section 2.2.7). The DAG fragments are joined to form the final DAG used in the forensic analysis. The process of constructing the DAG fragments, joining them together, and using the BN for hypotheses comparison is described below.

DAG fragment 1: Stability condition of the excavation due to the water table level and design variables

The DAG for representing the relationship between the stability of the excavation (SC), water table elevation (WT) and the design variables (LT), uses a simple cause-consequence idiom (refer to Section 2.2.7). The interpretation of the DAG presented in **Figure 5-13**a is straightforward and is based on the accepted knowledge that WT and LT have a direct influence on SC. However, since the influence of WT and LT on SC is estimated via computational experiments, the factor of safety (FoS) is included in the DAG of **Figure 5-13**b as a mediating variable between input and output nodes.

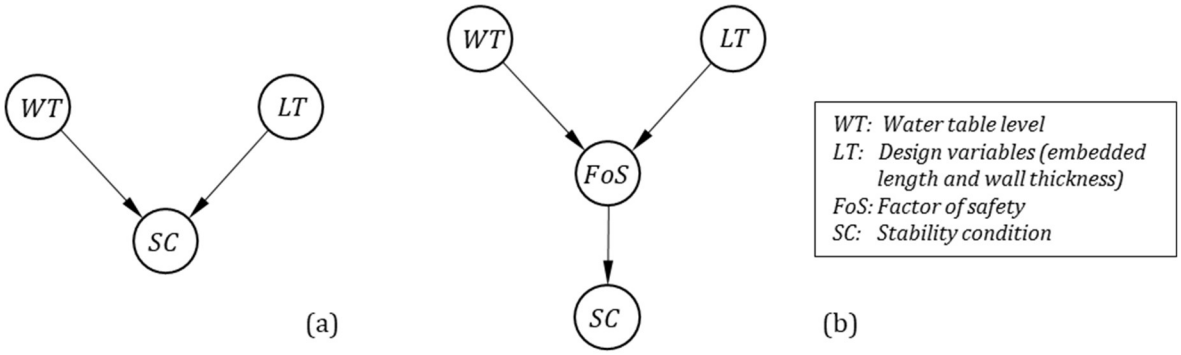


Figure 5-13. DAG fragment for the cause-consequence idiom between the stability condition of the excavation (SC), the position of the water table (WT), and the design variables (LT). (a) Simple DAG, (b) Final version of the DAG fragment with the FoS as a mediating node.

DAG fragment 2: Nodes indicating the stability condition of the excavation

The stability condition is itself an abstract concept, which is challenging to measure or observe. However, several measurable variables can provide information about the stability condition (SC). The measurement idiom and the indicator nodes described in Section 2.2.7 are used to characterize the causal relationships between (SC) and the following indicator variables: expert opinion (Eo), the heave (h) of the base of the excavation measured at point B (**Figure 5-5**), the inclination of the wall (i) measured at point W, and the settlement (ρ) measured at

point T on 11th street. The causal relationship between SC and its indicators is shown in **Figure 5-14.**

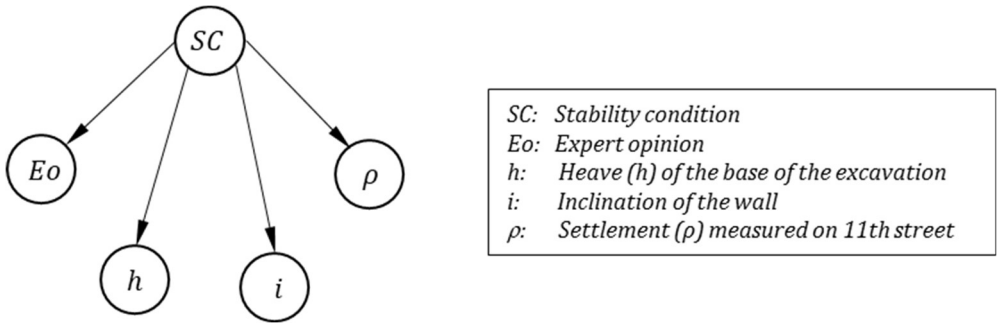


Figure 5-14. DAG fragment for the measurement idiom between the stability condition of the excavation (SC), and its indicators nodes: expert opinion (Eo), the settlement on the 11th street (ρ) measured at point T , the inclination of the wall (i) measured at point W, and the heave (h) measured at point B.

DAG fragment 3: Causes of variation in the water table and their indicators

One advantage of BNs over the prior odds ratio technique for comparing hypotheses is their ability to include additional nodes (i.e., variables) related to hypotheses, measurements, observations, and standard engineering practice. In the case of the Green Office Project, additional failure causes related to the water table elevation were formulated during the development of the forensic analysis. Owners and constructor contractors argued that damage to a sewer pipe near the project caused an unexpected elevation in the water table to 0.0 level. The sewerage network was inspected, and some minor damage, such as leaks and cracks, were identified.

On the other hand, local authorities, utility companies, and the affected community claimed that the excavation failure resulted from shortcomings in the geotechnical design. They argued that the geotechnical design should have considered a potential elevation in the water table to ground surface, given contractors' experience in nearby excavations and the quick response of the groundwater level to antecedent rainfall. In addition, they considered that the length and thickness of the diaphragm wall were inadequate to support lateral earth and hydrostatic pressures.

According to the hypotheses formulated by owners and local government entities, a variation in the water table elevation can be explained by two causes: antecedent rainfall (Ar) and damage to sewer pipelines (Dm) . **Figure 5-15a** shows the DAG fragment for this simple cause-effect idiom. Each cause (i.e., Ar and Dm) is represented in distinct nodes because they are not mutually exclusive. This means that the presence of one cause does not negate the possibility of the other. Since cause nodes Ar and Dm cannot be directly observed, indicator nodes are included. For node Ar, rainfall (Ra) node is included as an indicator node to identify whether significant antecedent rainfall was recorded in nearby weather stations. In the case of node Dm, inspection node (In) is used to indicate whether significant damage to a sewer pipeline was identified during inspection works. In addition, the indicator node Pz is included in the DAG to consider water table measurement using piezometers. The complete DAG fragment related to the causes of water table elevation is presented in **Figure 5-15b**.

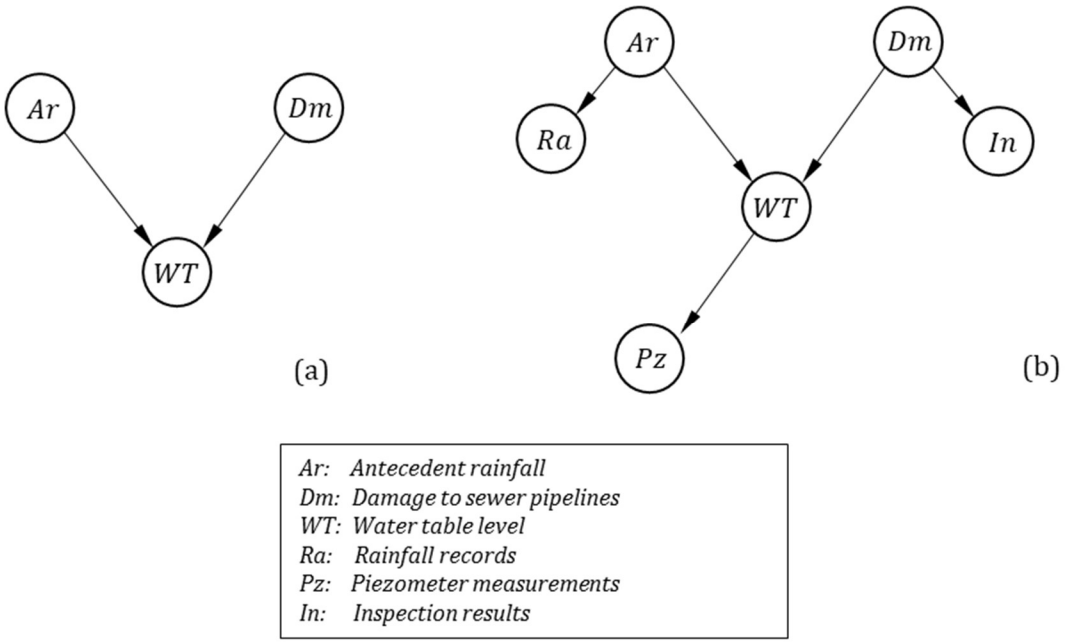


Figure 5-15. DAG fragments for causes of water table elevation. (a) DAG fragment from a simple cause-effect idiom. (b) Expanded DAG fragment with indicator nodes.

Final DAG

Figure 5-16 presents the complete DAG for the forensic analysis of the Green Office excavation failure. The complete DAG is obtained by joining the DAG fragments shown in **Figure 5-13b**, **Figure 5-14**, and **Figure 5-15b**. Red, blue, and green nodes in **Figure 5-16**represent evidence, hypothesis, and mediating nodes, respectively.

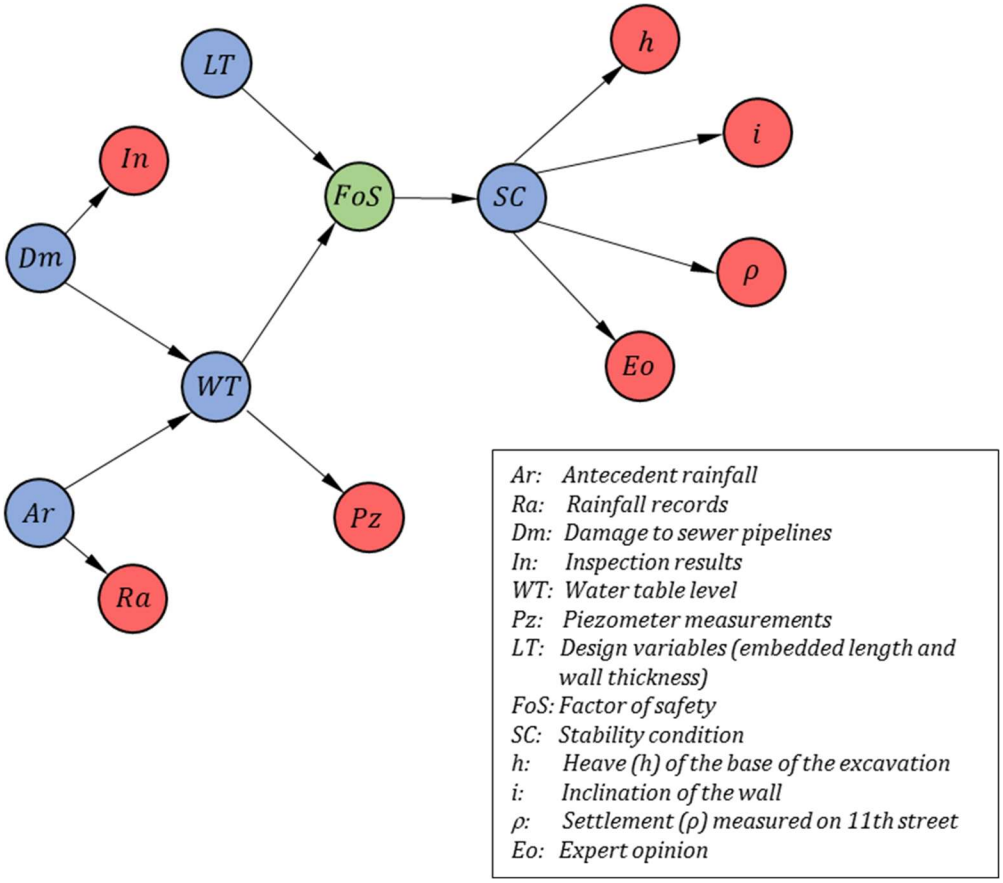


Figure 5-16. DAG for the forensic analysis of the Green Office excavation failure. Red, blue, and green nodes represent evidence, hypothesis, and mediating nodes, respectively.

Conditional Probability Tables (CPT)

The strength of causality relationships between nodes is defined by the conditional probability tables (CPTs) shown in **Figure 5-17**. The data for each CPT was obtained as follows:

- CPTs for the indicator nodes ρ, i, and h are constructed based on the results from the $N = 8000$ computational experiments described in Section 5.2.2. Each probability value is obtained by conditioning the indicator node on the state values of the SC node.

- In the case of node Eo, expert opinion is employed. Since not all expert opinions are identical, and some may disagree as to whether the excavation is stable, a probability of 80% is assigned to "*do not observe failure,*" given that the excavation is stable ($i.e., SC = stable$). Consequently, a probability of 20% is assigned to "*observe failure,*"

given that the excavation is stable. On the other hand, when the excavation is unstable (i.e., $SC = unstable$), experts are more likely to agree with the assessment. Therefore, a probability of 98% is assigned to "observe failure," and only 2% is assigned to "do not observed failure," given that the excavation is unstable. For practicality, the above probabilities are assigned using only one expert opinion. However, extensive studies can be conducted on expert opinions about stability conditions.

- The CPT for SC node is constructed from the FoS values. From numerical analysis, a $FoS \leq 1.0$ value indicates an unstable condition, whereas $FoS > 1.0$ denotes a stable excavation condition. For the convenience of analysis, it is assumed that all $FoS \leq 1.0$ values lead to an unstable condition and all $FoS > 1.0$ values represent stable conditions. The CPT of the SC node shown in **Figure 5-17** represents this situation.

- The CPT of the FoS is constructed based on the results from the $N = 8000$ computational experiments. Probability values are conditioned on the WT and LT values, as shown in **Figure 5-17**.

- In the case of LT, Dm, and Ar, their CPTs are defined based on prior knowledge. Since no prior knowledge is available, all states have the same probability values (**Figure 5-17**).

- The CPT values for indicator node Pz consider the accuracy of the piezometer and the error measurement rate. For simplicity, the values shown in **Figure 5-17** are assumed. These values are related to false negative rates reported in similar tests (Kruschke, 2015). CPT values for Ra and In nodes are defined similar to Pz.

- The CPT values of the WT node are conditioned on Dm and Ar states. These values are inferred from one expert opinion. For example, the probability of the water table at the 0.0 m level, given the simultaneous occurrence of antecedent rainfall and damage to sewer pipes, is 0.99. In contrast, the probability of the water table at the 0.0 m level, given that rainfall did not occur but damage to sewer pipelines occurred, is 0.15.

Bayesian Network and hypotheses comparison

Figure 5-18 presents the BN for the forensic analysis and the initial state in all their nodes. The initial state represents the probability of the node's states before any evidence is included in the BN. As described in Chapter 2, BNs can update the probability of their nodes by including evidence as realizations. Realizations involve assigning specific values to evidence nodes to verify how the rest of the nodes are updated.

The evidence shown in **Table 5-7** is included in the BN of **Figure 5-18** to compare the hypotheses about the cause of failure described in Section 5.2.1. The twelve probability queries of **Table 5-10** are used to include evidence and update the node's states of the BN. Probability queries ask questions about a hypothesis's probability given some evidence. For example, the probability query for the hypothesis H_3 and evidence e_2 can be read as the *probability of observing simultaneously a diaphragm wall 0.4 thick and 17.0 m long and a water table at the ground surface, given that a failure and a settlement on 11th street greater than 0.3 m are observed*. **Table 5-10** presents the numerical result of the twelve probability queries.

Table 5-10. Probability queries for hypotheses H_1 to H_4 and evidence e_1 to e_3.

Hypothesis H_i	**Evidence** e_k	**Probability query**	$P(H_i \mid e_k)$
H_1	e_1	$P(LT_1, WT_{-3.0} \mid Eo = observe\ failure)$	3.29×10^{-1}
	e_2	$P(LT_1, WT_{-3.0} \mid Eo = observe\ failure, \rho > 0.3\ m)$	3.65×10^{-1}
	e_3	$P(LT_1, WT_{-3.0} \mid Eo = observe\ failure, \rho > 0.3\ m, i > 0.2, h > 0.3)$	3.65×10^{-1}
H_2	e_1	$P(LT_2, WT_{-3.0} \mid Eo = observe\ failure)$	9.10×10^{-2}
	e_2	$P(LT_2, WT_{-3.0} \mid Eo = observe\ failure, \rho > 0.3\ m)$	3.40×10^{-2}
	e_3	$P(LT_2, WT_{-3.0} \mid Eo = observe\ failure, \rho > 0.3\ m, i > 0.2, h > 0.3)$	3.40×10^{-2}
H_3	e_1	$P(LT_1, WT_{0.0} \mid Eo = observe\ failure)$	3.64×10^{-2}
	e_2	$P(LT_1, WT_{0.0} \mid Eo = observe\ failure, \rho > 0.3\ m)$	4.04×10^{-1}
	e_3	$P(LT_1, WT_{0.0} \mid Eo = observe\ failure, \rho > 0.3\ m, i > 0.2, h > 0.3)$	4.03×10^{-1}
H_4	e_1	$P(LT_2, WT_{0.0} \mid Eo = observe\ failure)$	2.15×10^{-1}
	e_2	$P(LT_2, WT_{0.0} \mid Eo = observe\ failure, \rho > 0.3\ m)$	1.96×10^{-1}
	e_3	$P(LT_2, WT_{0.0} \mid Eo = observe\ failure, \rho > 0.3\ m, i > 0.2, h > 0.3)$	1.97×10^{-1}

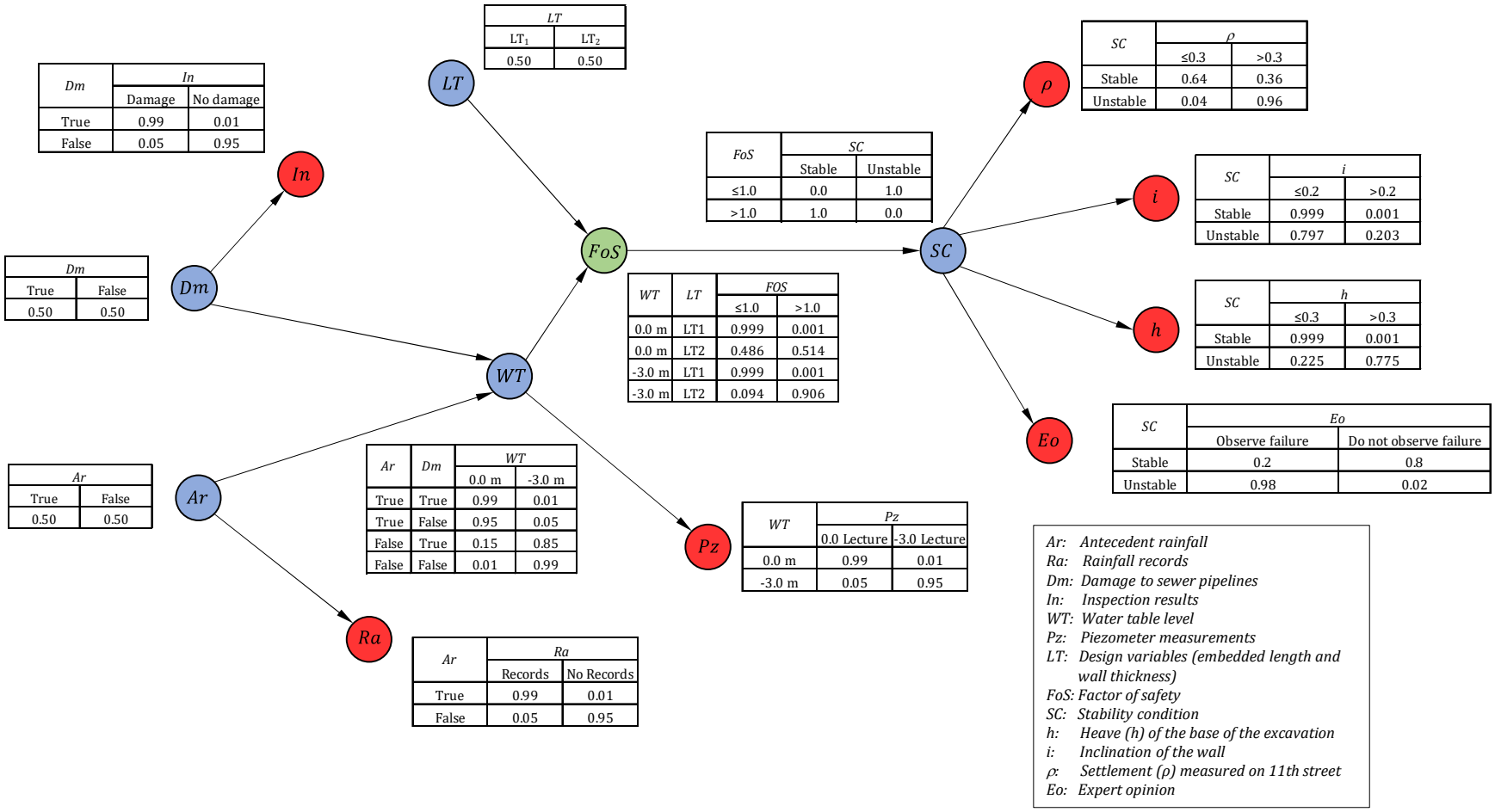


Figure 5-17. DAG for the forensic analysis of the Green Office excavation failure and its conditional probability tables (CPTs).

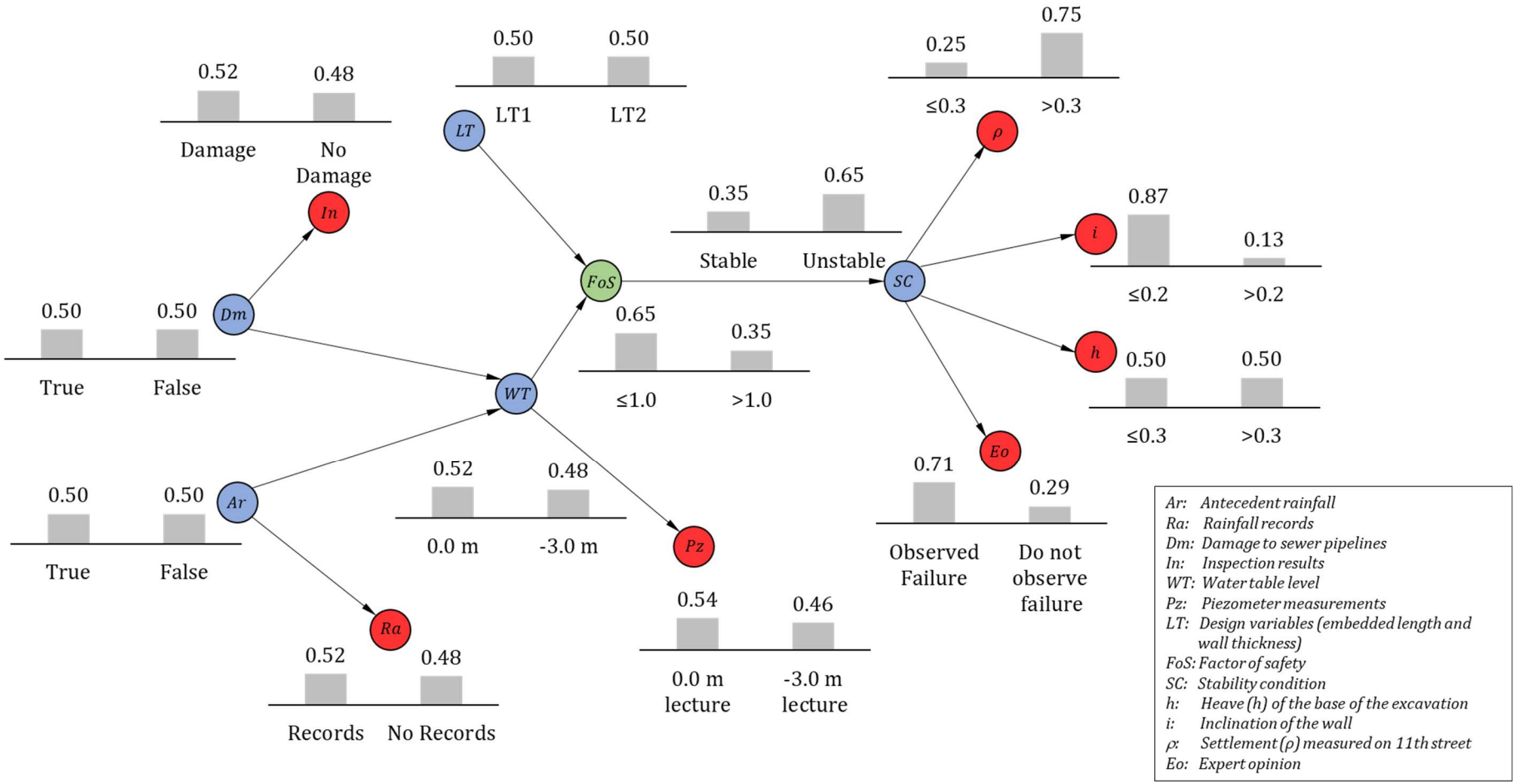


Figure 5-18. BN for the forensic analysis of the Green Office excavation failure. Initial state (i.e., evidence not included in the nodes).

Unlike the posterior odds technique for hypotheses comparison, BNs may include additional nodes (variables) related to hypotheses or evidence. In the case of the Green Office Project, the BN includes variables related to the water table level (Pz, Ra, AR), stability condition (SC, FoS), and several indicator nodes (ρ, i, h, Eo). These additional nodes allow evidence to be more easily included. Furthermore, the influence of the amount of evidence on the updating process can be observed graphically in the BN. The following analysis present four examples regarding the influence of evidence and its impact on failure hypotheses.

- Case No. 1: Assume that only two pieces of evidence are available: a piezometer lecture at the ground surface and no significant damage reported during the sewer pipe inspection ($Pz = 0.0\ m\ lecture, In = No\ damage$). **Figure 5-19** compares the initial BN (no evidence included) and the updated BN (evidence included). Notice that by including these two pieces of evidence, the probability of the "Unstable" state in the SC node increased from 0.65 to 0.73, and the water table WT as a cause of failure increased from 0.53 to 0.95. However, given the d-separation properties of the BN, the states of the design node LT remain constant ($LT_1 = LT_2 = 0.5$). Consequently, the two pieces of evidence do not provide information about the causes of failure related to the design variable LT.

- Case No. 2: In addition to the evidence provided in Case No. 1, an expert is consulted on the stability of the excavation. After a field visit, the expert concludes that the excavation is in a condition of failure. This additional piece of evidence is included in the BN as $Eo = observe\ failure$. **Figure 5-20** compares the updated BN of Case No. 1 and the updated BN with the additional evidence (Case 2). The resulting BN shows a minimal increase in the "0.0 m" state of the WT node but reveals a significant increase in LT_1, and a reduction in LT_2. In other words, the expert opinion about the stability condition favors the design LT_1 as a probable cause that led to the excavation failure.

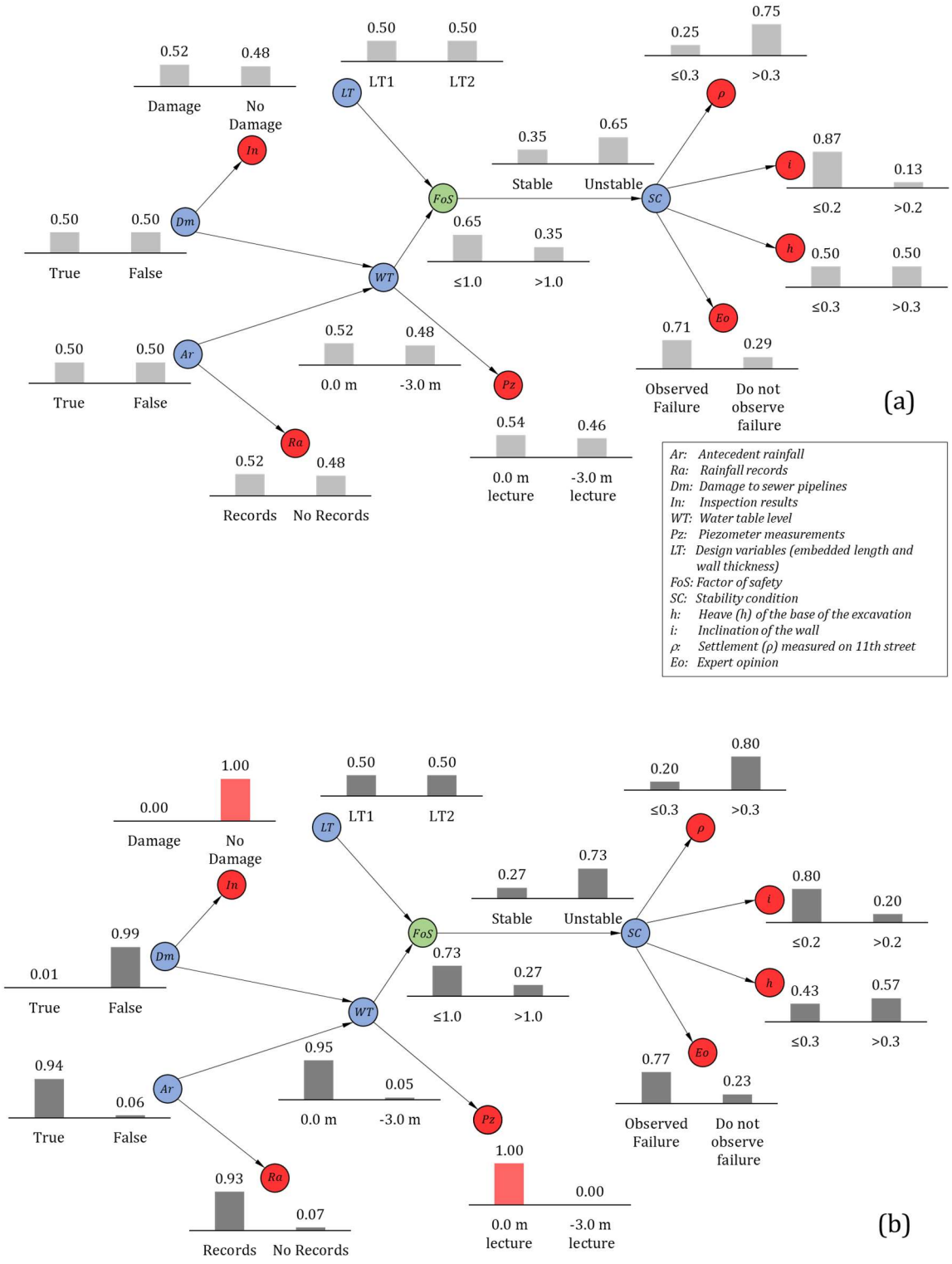
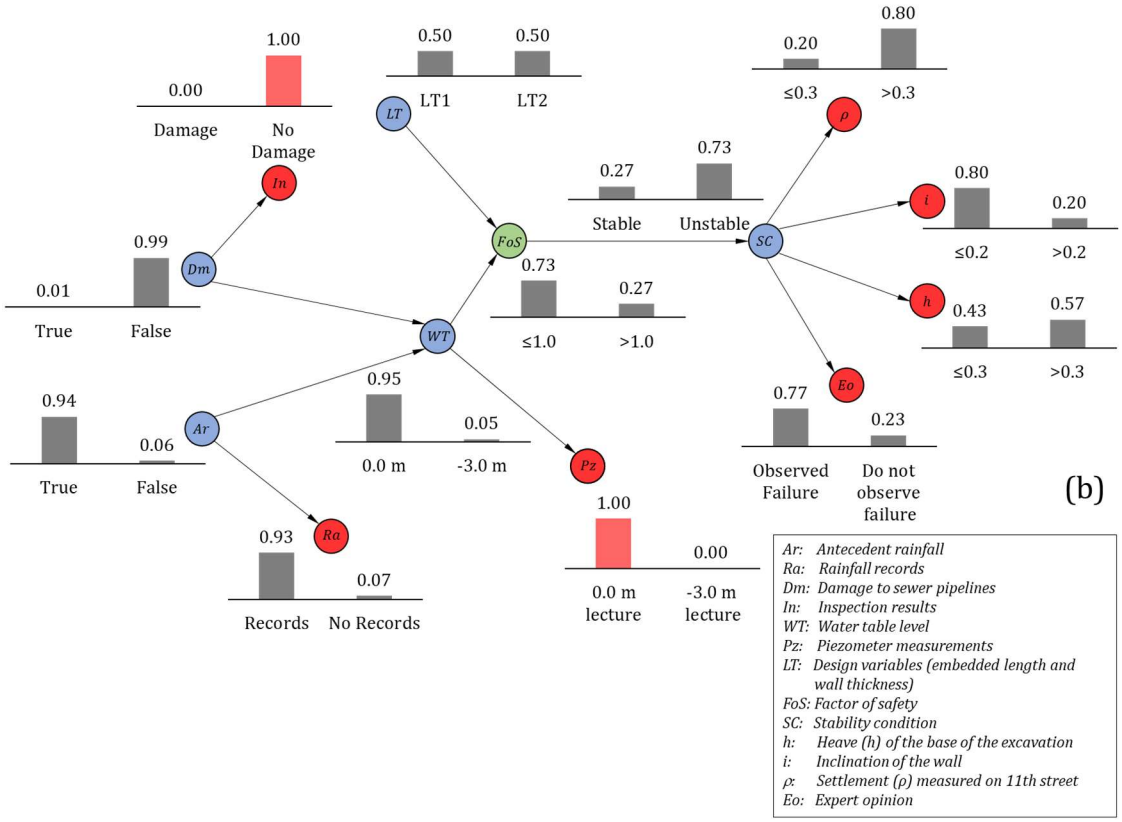


Figure 5-19. Case No 1. Comparison between (a) the initial BN and (b) the updated BN for evidence: $Pz = 0.0\ m\ lecture, In = No\ damage$.



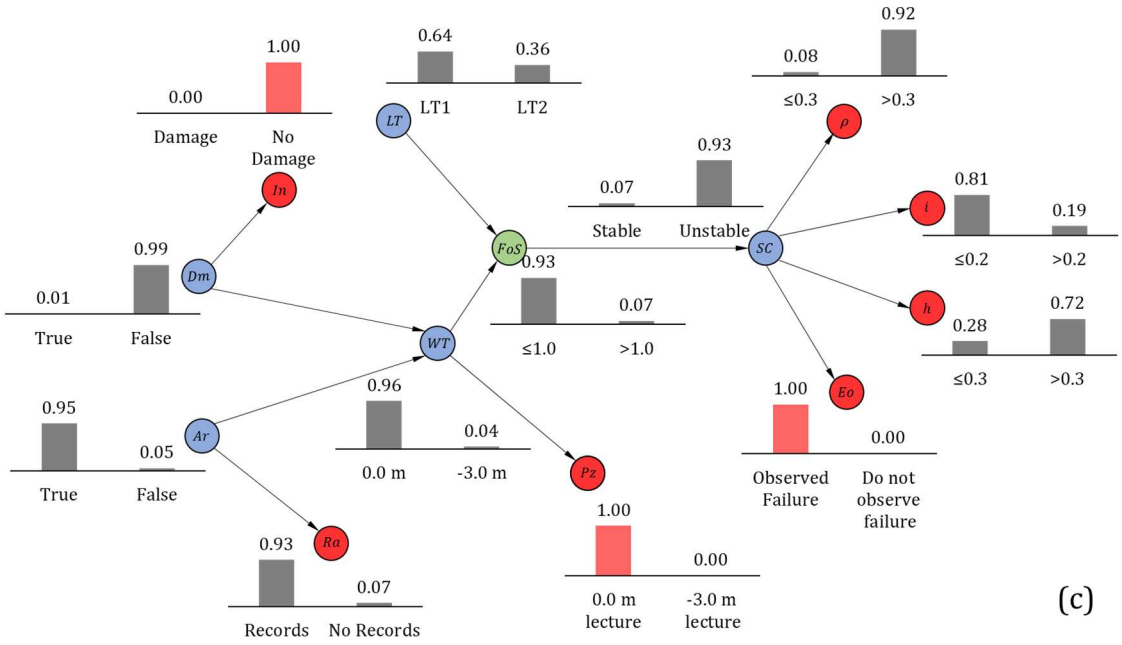


Figure 5-20. Case No 2. Comparison between (b) the BN of Case No. 1 and (c) the updated BN for evidence: $Pz = 0.0\ m\ lecture, In = No\ damage, Eo = observe\ failure$.

- Case No. 3: Assume now the case with the following pieces of evidence: (i) a piezometer reading shows a water table at -3.0 m level, (ii) there are no extraordinary rain records during previous months near the excavation, (iii) damage to sewer pipelines is found, (iv) settlement measurements on 11th street exceed 0.30 m, (v) the inclinometer near the diaphragm wall shows an inclination greater than 0.20 m, (vi) the heave of the base of the excavation is higher than 0.3 m, and (vii) the expert concludes that the excavation is under failure. **Figure 5-21** compares the initial and updated state of the BN after including the above seven pieces of evidence. Notice that under these conditions, the BN indicates that the design LT_1 (i.e., a diaphragm wall 17 m long and 0.4 thick) is the most probable cause of excavation failure. In addition, although damage to sewer pipelines is found, the BN indicates that the elevation of the water table to the ground surface is not a probable cause of failure.

- Case No. 4: Case No. 4 includes the actual evidence collected during the forensic investigation of the Green Office Project. The evidence includes the following information:

  Monthly rainfall data during 2011 compared with previous years

  Figure 5-22 shows the monthly rainfall data provided by IDEAM (Colombian weather service) from 2007 to 2011. Precipitation data from May to October 2011 shows lower values than the same months of previous years. However, in November 2011 (one month before the failure), the precipitation was extraordinarily high (292.4 mm). Consequently, this information is included as evidence in the BN as $Ra = Records$.

  Piezometer measurements near the excavation failure area

  Figure 5-23 presents the water table level measured in piezometers PZ1 and PZ5 from January 2011 to January 2012. Piezometer PZ1 shows a stable water table level from June 2011 to September 2011, but by October 2011 and November 2011, a considerable elevation to the -3.0 m level has been recorded. Then, by December 2011, a sudden decrease in the level was observed, possibly explained by the damage to the piezometer caused by the excavation failure. The information provided by the piezometers is included in the BN as $Pz = -3.0\ m\ reading$.

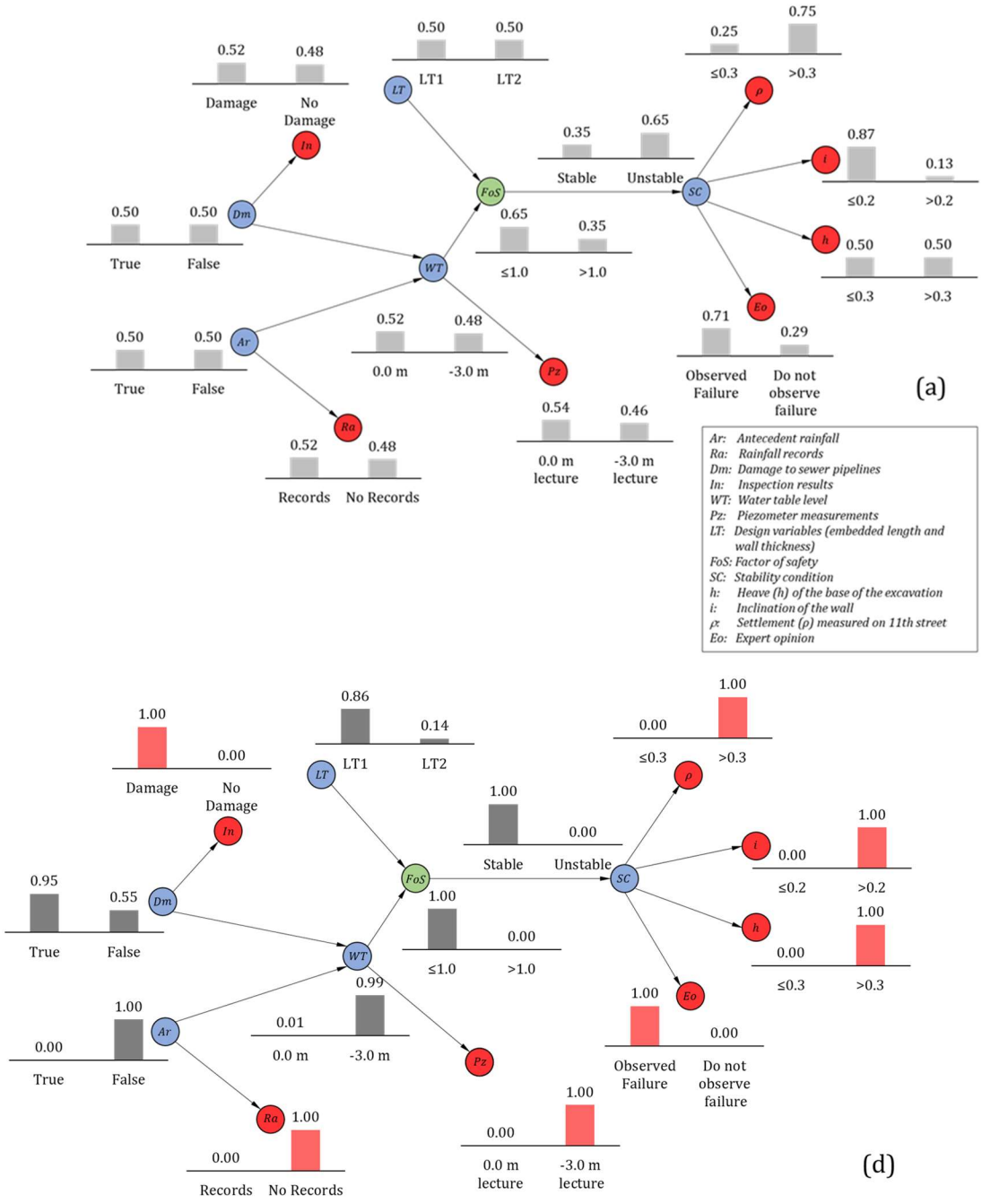


Figure 5-21. Case No 3. Comparison between (a) the initial BN and (d) the updated BN for evidence: $Pz = -3.0\ m\ reading, Ra = No\ records,\ In = damage, \rho > 0.3\ m, i > 0.2\ m, h > 0.3\ m, Eo = observe\ failure$.

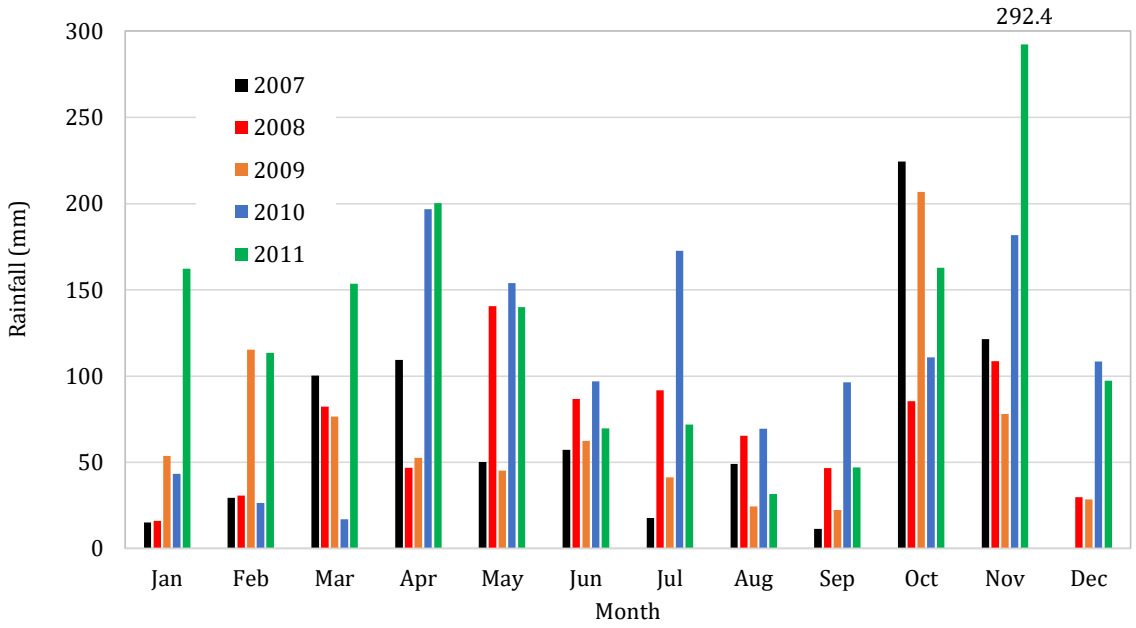


Figure 5-22. Monthly rainfall from 2007 to 2011 in the project area (Unal, 2012).

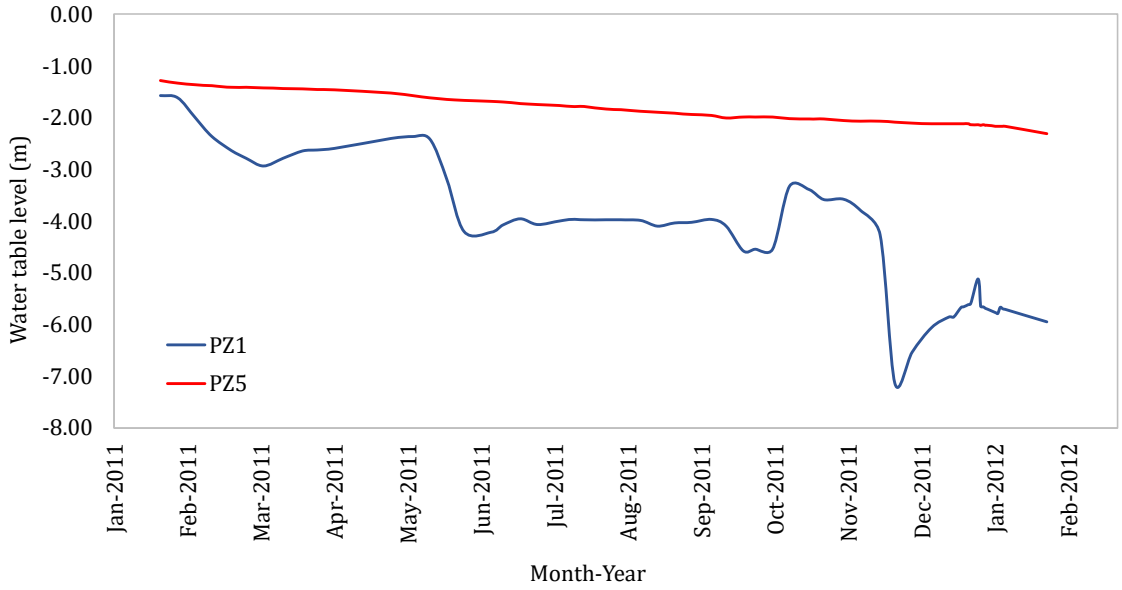


Figure 5-23. Variation in the water table level according to piezometers PZ 1 and PZ5 (Unal, 2012).

Inspection and diagnosis of aqueduct and sewer pipelines

The Bogota Aqueduct and Sewerage Service (EAAB) inspected the aqueduct and sewer pipelines near the project area to detect any damage. The inspection did not find any damage to aqueduct pipelines. However, several sewer pipelines were clogged, and minor leaks were detected. The evidence provided by this information is included in the BN as $In = Damage$.

Topographic survey: 11th street and at the bottom of the excavation.

A topographical survey was carried out to verify the magnitude of the settlements on 11th street. **Figure 5-24** presents the longitudinal section B-B (refer to **Figure 5-2**) of the Green Office Project, where a differential settlement of 0.90 m on 11th street was observed. **Photo 5-1** demonstrates the settlement's magnitude and impact on vehicular and pedestrian traffic. The evidence provided by this information is included in the BN as $\rho > 0.30\ m$.

A topographical survey at the bottom of the excavation was not carried out during the construction stage. However, the construction logbook reveals a heave during the excavation activities for the third basement. The evidence provided by this information is included in the BN as $h > 0.30\ m$.

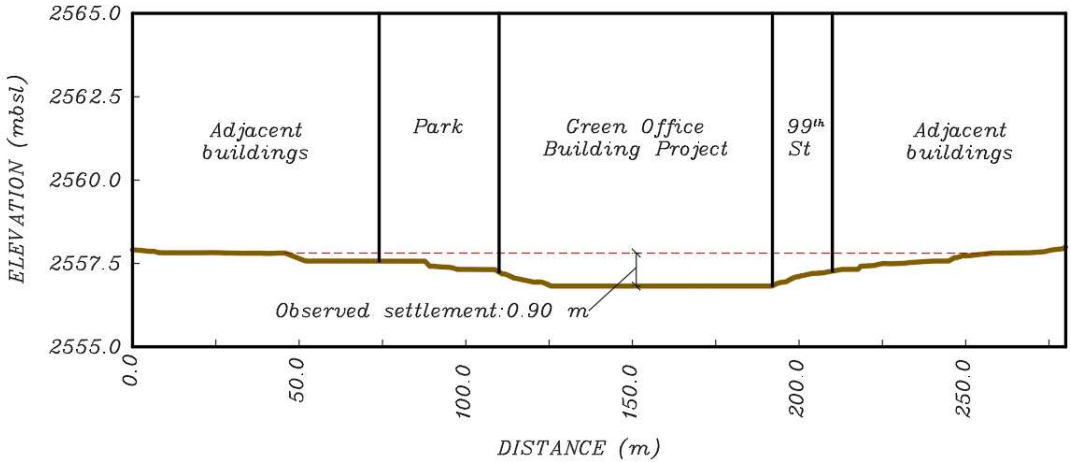


Figure 5-24. Longitudinal section B-B (refer to Figure 5-2). Maximum settlement measured on 11th street.

Photo 5-1. Settlement on 11th street. (a) impact on the vehicular roadway, and (b) impact on the pedestrian walkway. Photographs from Caracol Radio (2012) and Radio Santafe (2012).

Inclinometer measurements adjacent to the diaphragm wall

Inclinometers I6, I7, and I8 were the closest to the excavation failure area (refer to **Figure 5-6**). The accumulated measurements of inclinometers I7 and I8 from October 2010 to January 2012 are plotted in **Figure 5-25**. Although inclinometer I6 was damaged in October 2011, it recorded an accumulated inclination of 196 mm up to that date. The damage to the inclinometer may be due to the excavation failure. Given that no conclusive evidence is gathered from inclinometer measurements, the information is included in the BN as $i < 0.20\ m$.

Figure 5-26 presents the updated state of the BN after the inclusion of the actual evidence collected during the forensic investigation (Case No. 4). The analysis of the BN of **Figure 5-26** and its comparison with the BN of Cases 1 to 3 (**Figure 5-19** to **Figure 5-21**), leads to the following conclusions:

- The six pieces of actual evidence presented in Case no. 4 indicate that the design LT_1 (diaphragm wall 17 m long and 0.4 m thick) combined with a water table at the -3.0 m level, was the likely cause of the excavation failure. A closer inspection of the updated BN (**Figure 5-26**) reveals that the states $LT = LT_1$ and $WT = -3.0\ m$ are far more likely than the states $LT = LT_2$ and $WT = 0.0\ m$. The results from the conditional probability queries presented in **Table 5-11** support this conclusion. Note that the

combination of states $LT = LT_1$ and $WT = -3.0\,m$ is more likely than other combinations.

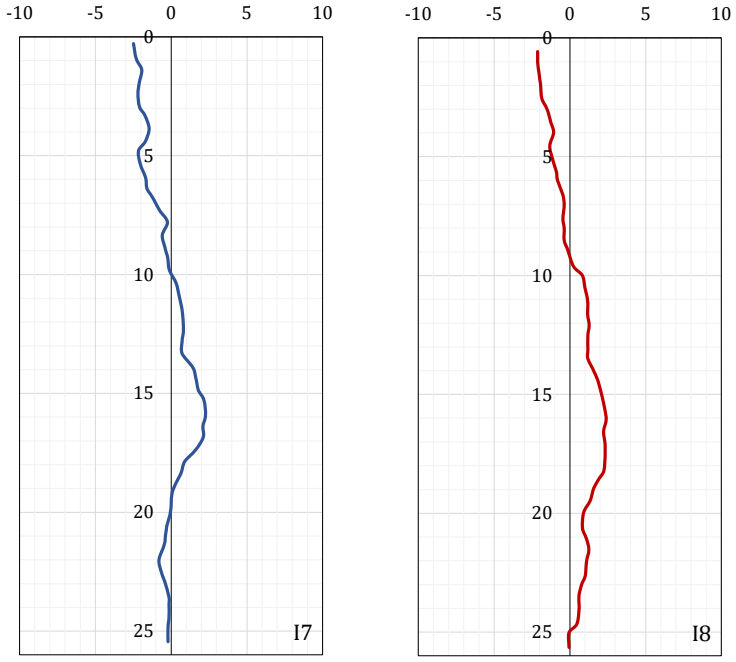


Figure 5-25. Accumulated inclination measured in inclinometers I7 and I8 (Unal, 2012).

- Although damage to an adjacent sewer pipeline was detected, the updated BN of **Figure 5-26** indicates no impact on $WT = 0.0\,m$. Therefore, damage to a sewer pipeline is unlikely to explain the excavation failure.

- Even though unusual rainfall was recorded during the two months before the failure, the piezometer readings remain below the -3.0 m level on average. Consequently, the unusual rainfall and the likely elevation of the water table to the ground surface are unlikely to explain the excavation failure.

- The evidence provided by nodes ρ, h, and i is decisive in determining the stability of the excavation. Even though the inclination measurements i do not provide conclusive evidence about the inclination in the diaphragm wall, settlements ρ recorded on the 11th and the heave h identified at the base of the excavation are enough to demonstrate the failure of the excavation. In addition, although the expert opinion node Eo is not used as evidence, the state $Eo = observe\ failure$ is updated to 98%.

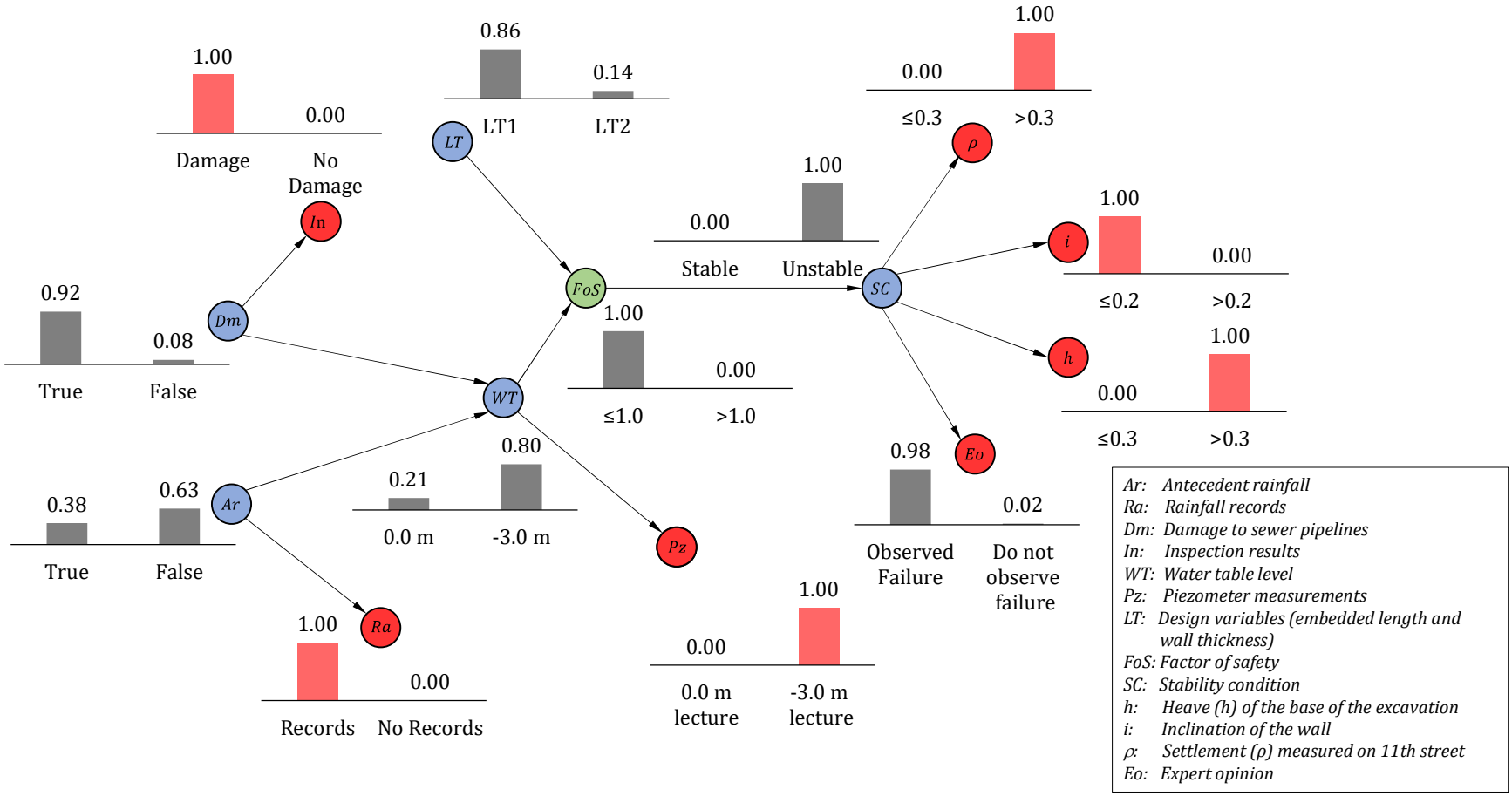


Figure 5-26. Updated BN after the inclusion of actual evidence:

$$Ra = Records, Pz = -3.0\ lecture, In = Damage, \rho > 0.3\ m, i < 0.2, h > 0.3.$$

Table 5-11. Probability queries for the actual evidence of the Green Office failure.

Probability query	$P(H_i \mid e_k)$
$P(LT_1, WT_{0.0\,m} \mid Ra = Records, Pz = -3.0\ reading, In = Damage, \rho > 0.3\ m, i < 0.2, h > 0.3)$	0.138
$P(LT_1, WT_{-3.0\,m} \mid Ra = Records, Pz = -3.0\ reading, In = Damage, \rho > 0.3\ m, i < 0.2, h > 0.3)$	0.727
$P(LT_2, WT_{0.0\,m} \mid Ra = Records, Pz = -3.0\ reading, In = Damage, \rho > 0.3\ m, i < 0.2, h > 0.3)$	0.067
$P(LT_2, WT_{-3.0\,m} \mid Ra = Records, Pz = -3.0\ reading, In = Damage, \rho > 0.3\ m, i < 0.2, h > 0.3)$	0.068

K Most Probable Explanation – K MPE

All combinations of hypotheses nodes in Figure 5-16 that led to the observed evidence ($Ra = Records, Pz = -3.0\ lecture, In = Damage, \rho > 0.3\ m, i < 0.2, h > 0.3$) are identified through an R routine (refer to Annex A8). The routine detected the 17 combinations shown in Table 5-12. According to this table, the most probable explanation of the excavation failure is a water table elevation at the -3.0 m level and a diaphragm wall 17 m long and 0.4 thick. Notice that the first K=3 explanations include LT_1 as a probable explanation. Explanation K=4 includes LT_2 as a probable explanation, but its probability is two orders of magnitude less likely than K=3.

Table 5-12. K=17 Most probable explanations of the Green Office excavation failure.

K	***Dm*** **(Damage)**	***Ar*** **(Antecedent Rainfall)**	***WT*** **(Water Table)**	***LT*** **(Design)**	***FoS*** **(Factor Of Safety)**	***SC*** **(Stability Condition)**	***Eo*** **(Expert opinion)**	***Prob.***
1	FALSE	TRUE	-3.0m	LT1	<1.0	Unstable	Observe_ failure	0.5378
2	TRUE	TRUE	0.0m	LT1	<1.0	Unstable	Observe_ failure	0.1338
3	TRUE	TRUE	-3.0m	LT1	<1.0	Unstable	Observe_ failure	0.1227
4	TRUE	TRUE	0.0m	LT2	<1.0	Unstable	Observe_ failure	0.0656
5	FALSE	TRUE	-3.0m	LT2	<1.0	Unstable	Observe_ failure	0.0456
6	FALSE	FALSE	-3.0m	LT1	<1.0	Unstable	Observe_ failure	0.0346

7	TRUE	FALSE	-3.0m	LT1	<1.0	Unstable	Observe_ failure	0.0275
8	TRUE	TRUE	-3.0m	LT2	<1.0	Unstable	Observe_ failure	0.0126
9	TRUE	FALSE	0.0m	LT1	<1.0	Unstable	Observe_ failure	0.0069
10	TRUE	FALSE	0.0m	LT2	<1.0	Unstable	Observe_ failure	0.0040
11	FALSE	FALSE	-3.0m	LT2	<1.0	Unstable	Observe_ failure	0.0038
12	TRUE	FALSE	-3.0m	LT2	<1.0	Unstable	Observe_ failure	0.0029
13	FALSE	TRUE	0.0m	LT1	<1.0	Unstable	Observe_ failure	0.0009
14	FALSE	TRUE	-3.0m	LT2	>1.0	Stable	Do_not_observe_ failure	0.0005
15	TRUE	TRUE	0.0m	LT2	>1.0	Stable	Do_not_observe_ failure	0.0002
16	FALSE	TRUE	0.0m	LT2	<1.0	Unstable	Observe_ failure	0.0002
17	FALSE	TRUE	-3.0m	LT2	>1.0	Stable	Observe_ failure	0.0002

5.2.4 Comparison of POR and BN Techniques

POR technique

The analysis of the Green Office failure using the POR technique revealed that initial shortcomings in the design were the most probable cause of the excavation failure. This conclusion is derived from the probabilistic comparison of several hypotheses formulated by the stakeholders. Even though several prior odd values were analyzed to check the sensitivity

of the analysis, the POR technique disproves several hypotheses and shows that the evidence is conclusive in favor of design flaws.

The POR analysis was carried out by increasing the amount of evidence. In the first case, only the evidence (e_1) related to the stability of the excavation was used. The results showed that this evidence alone did not provide conclusive results for any hypothesis. When additional evidence related to settlements and inclination measurements was included (e_2 and e_3), the analyses indicated that design flaws were the most likely cause of the excavation failure. These results show the importance of including all available evidence in forensic analyses.

BN technique

The analysis of the failure using BNs allowed the inclusion of additional variables. These variables are related to qualitative evidence, such as expert opinion (*Eo*), antecedent rainfall (*Ar*), and damage to sewer pipelines (*Dm*) that could not be included in the POR technique. The inclusion of qualitative variables strengthened the analysis and reaffirmed design flaws as the most likely cause of failure.

As in the POR technique, evidence was gradually included in the BN. The results showed that including a few pieces of evidence does not yield conclusive results. However, including all available evidence in the BN allows for determining the most likely cause of failure with high reliability. Furthermore, given the BN's characteristics, the hypothesis comparison process is traceable and reproducible.

Main Differences

The main difference between POR and BN techniques lies in the type and amount of information that can be included in the analysis. Although both techniques employ similar geotechnical and information models, the POR technique is simpler in its construction and interpretation. However, it is limited by the type of information that can be included. For example, the direct inclusion of qualitative variables in Bayes' theorems is not possible with the POR technique.

On the other hand, although constructing a reliable BN is challenging, the inclusion of additional variables expands the number of initial hypotheses and allows for more complex

queries. For example, in the studied case, qualitative evidence related to expert opinion and inspection of sewerage networks included some aspects of the hypotheses that were not included in the POR analysis.

5.3 Summary

Chapter 5 presents an actual case study of the proposed Bayesian methodology for supporting decisions regarding the causes of geotechnical failures. The case study is a forensic geotechnical analysis of a deep excavation failure that occurred in the Green Office Project in Bogotá, Colombia. The general characteristics of the project and the geotechnical conditions under which the excavation failed are described. The main aspects of the case study and its results can be summarized as follows:

- The Green Office Project is a six-level building with three basement levels located on the northeast side of Bogotá, Colombia. During construction activities, an excavation failure was observed. Heave of the bottom of the excavation, large settlements on 11th street, and significant horizontal movements in the diaphragm wall were monitored.

- The failure sequence included the following steps: (i) small settlements values were observed during the excavation for the first basement (-4.0 m level), (ii) large settlements up to 0.90 m adjacent to the excavation were measured during construction activities for basements two and three (-11.0 m level), (iii) large horizontal displacements were recorded in the surrounding area and the diaphragm wall.

- Stakeholders formulated some hypotheses about the causes of failure. Building owners and construction contractors stated that excavation failure occurred as a result of an unexpected elevation in the water table level due to damage to a sewer pipe near the building. Utility companies and local government authorities claimed that the failure was due to design shortcomings. In particular, the thickness and the embedded length of the diaphragm wall were inadequate to support lateral pressures.

- Two probabilistic failure models were constructed in order to apply the proposed Bayesian methodology and determine the causes of the excavation failure. The first

model analyzes the strength limit state through limit equilibrium via Janbu's corrected equations. The second model used a finite element analysis to evaluate the service limit state and to calculate deformations.

- The POR and BN techniques described in Chapter 4 were applied to the execution failure of the Green Office Project. The POR technique compared the competing hypotheses for several prior odds ratio values and different amounts of evidence. In all cases, the results identified that the most likely cause of the excavation failure was a shortcoming in the diaphragm wall design (thickness and length) combined with an elevation of the water table.

- The BN for the forensic analysis of the Green Office Project was constructed from small pieces of cause-effect relationships. Each piece was translated into a simple DAG. Then, all simple DAGs were joined together to create a large DAG that accounts for hypotheses and evidence. Probability relationships (i.e., CPT) were estimated from the limit equilibrium and finite element probabilistic models. The actual evidence collected during the forensic investigation was included in the BN. Unlike the POR technique, the BN identified only the shortcomings of the diaphragm wall design as the most likely cause of failure. The difference in results between the two techniques may be explained by the amount of evidence used in each case. For example, the evidence related to antecedent rainfall and damage to sewer pipelines was not included in the POR analysis, given the limitations of this Bayesian technique. However, the results from both techniques are comparable and support the thickness and short embedded length of the diaphragm wall as the most likely cause of the excavation failure.

6. Discussion

6.1 Key Findings and Implications

This section discusses key findings and implications of the Bayesian methodology for forensic geotechnical analysis. Improvement in estimating failure causes using the proposed techniques is examined.

6.1.1 Improvement in Estimating Failure Causes

The Bayesian methodology proposed in this thesis demonstrates that probabilistic methodologies constitute an essential tool that may help establish the most probable cause of a geotechnical failure. Bayesian inference via *posterior odds ratio (POR)* and *Bayesian networks (BN)* were the probabilistic tools used in this research due to their ability to simulate the abductive reasoning used in forensic engineering. The examples presented in Chapter 4 and the case study analyzed in Chapter 5 have demonstrated that the proposed methodology identifies the most probable cause of failure even when evidence is scarce.

Authors such as Kool et al. (2019) recognized that conclusions about the causes of a geotechnical failure sometimes seem biased, and therefore findings can be questionable. The Bayesian methodology overcomes this bias by testing several hypotheses against the available evidence. Since the testing process relies on a verifiable probability framework, hypotheses about the causes of failure can be compared probabilistically. For example, the method can estimate how many times better hypothesis H_i explains a failure than hypothesis H_j. In addition, the *K most probable explanation (KMPE)* algorithm provides the top K explanations of the causes of failure in a BN.

The ERTC7 benchmark in Section 4.2 presents an example of the improvement in estimating failure causes. In this exercise, a short embedded length of the wall was fixed as the cause of

an excavation failure. The excavation was then analyzed, assuming that the cause of failure was not known in advance. To this end, several hypotheses were formulated: (i) the elevation of the water table, (ii) an unforeseen increase in the surcharge, and (iii) the short embedded length of the wall. Each hypothesis was tested against the evidence using the POR technique. The results and the probabilistic analysis showed that the POR technique found the short embedded length of the wall as the most likely cause of failure.

The Breitenhagen levee failure presented in Section 4.3 provides an additional example of the improvement in estimating failure causes. In this case, the BN technique was used as an expert system. The hypotheses about failure causes were included as nodes in the BN in order to represent the failure model probabilistically. Different hypotheses regarding pore water pressure and geotechnical conditions were evaluated by including evidence and querying the BN. In addition, the KMPE algorithm was used to estimate the most probable causes of the levee failure. The analysis showed that the results provided by the BN are comparable to other forensic techniques (Kool et al., 2019; Kool et al., 2020).

6.1.2 Beyond Traditional Back Analysis

As described in Chapter 3, back analysis is the preferred tool in forensic geotechnical engineering (FGE). Hwang (2016) points out that back analysis must be used cautiously due to limitations, such as model complexity and numerical analysis. He argues that back analyses outcomes should only be interpreted by expert analysts. Furthermore, back-analysis tools should be chosen based on the collected evidence, laboratory tests, and soil/rock behavior. In practice, the standard procedure to construct the back analysis model is based on the "as-built" conditions of the structure (Babu & Singh, 2016). Then, the outcomes from the back analysis are compared against the collected evidence using deterministic criteria.

As in the standard procedure, the methodology proposed in this thesis constructs geotechnical models based on the characteristics of the materials and the collected evidence. However, the method recognizes the uncertainty of failure conditions and expands back analysis to include several failure hypotheses. Compared with standard back analyses, where only a few computational experiments are performed, the Bayesian methodology performs thousands of experiments. The outcomes from these experiments are translated into POR or BN, which can be queried as an expert system (refer to Section 2.2.9). The answers to the queries are

probabilistic, and comparisons between hypotheses are objective. The latter aspect is particularly important in the sense that the methodology is traceable and replicable. Consequently, it provides more reliable conclusions about the causes of geotechnical failures than traditional methods and simple back analysis. This aspect constitutes the key contribution of this research to forensic geotechnical engineering.

6.1.3 Two Bayesian Techniques for Hypotheses Comparison

To date very few papers have been published on the role of Bayesian analysis in FGE (e.g., Kool et al., 2020). In fact, no research on the use of Bayesian networks in FGE has been published, except for a preliminary approach carried out by Xu & Zhang (2016). The Bayesian methodology described in Chapter 4 represents a considerable breakthrough in the available techniques that support decisions about the causes of geotechnical failures. Furthermore, the possibility of changing the way that geotechnical engineers make decisions about the causes of failures is an exciting contribution to existing knowledge and potentially a vast area of research.

The method described in Chapter 4 explores two Bayesian tools for comparing hypotheses: posterior odds ratio (POR) and Bayesian networks (BN). Whereas the POR technique uses the basic Bayes' theorem (refer to Equation 2-6) to compare two competing hypotheses, BNs expand the theorem through the chain rule (refer to Section 2.2.6). In the case of the POR, probabilities are estimated directly from Bayes' theorem. Thus, calculations are simple, but results are sensitive to prior odds values. On the other hand, BNs overcome prior odds values influence, but their construction is more complicated than the POR. Further discussion of prior odds' influence on posterior probabilities and Bayesian network construction is given in Section 6.2.

Posterior Odds Ratio (POR)

Although the POR technique for comparing failure hypotheses has been used extensively in criminal investigations and forensic science (e.g., Taroni et al., 2014; Neil et al., 2019), to date there have been no reports in FGE. The methodology proposed in this thesis adapts the POR technique to the characteristics of geotechnical failures, such as material uncertainty, soil constitutive models, back analyses, failure scenarios, and numerical methods. The methodology represents an enhancement in FGE as it surpasses the conventional methods in

deriving conclusions about the causes behind geotechnical failures. For example, Kool et al. (2020) propose a Bayesian procedure based on an event tree for hindcasting levee failures. However, their procedure does not compare hypotheses using any criteria. In contrast to the Kool et al. procedure, Section 4.2 presents an example of the use of the POR technique and the criteria for selecting the most probable cause of failure based on Jeffreys's (1961) scale and Kass & Raftery's (1995) modified scale.

Bayesian Networks (BN)

The BN technique proposed in Section 4.1 and the results from examples in Section 4.3 and Chapter 5 suggest that BNs can help to identify the most probable cause of a geotechnical failure. Combining computational experiments and expert opinion through BNs results in less arbitrary conclusions. Furthermore, the KMPE algorithm can list the most probable causes that explain a failure. Similar results have been achieved in criminal investigations and forensic engineering. For example, Garbolino & Taroni (2002); Taroni et al. (2004); Biedermann et al. (2005); Kwan et al. (2008); Biedermann & Taroni (2012); and Holický et al. (2013) present several examples regarding BN applications for interpreting evidence and assessing failures. However, BNs have not been extensively employed in FGE, possibly due to the complexity of geotechnical failures and the poor training of geotechnical engineers in Bayesian statistics (Baecher, 2017).

6.1.4 Use of Standard Knowledge and Geotechnical Jargon

The proposed Bayesian methodology applies the DAG construction procedure suggested by Kjærulff & Madsen (2013). This procedure is based on depicting small pieces of cause-effect relationships through simple DAGs and then joining them together to represent large and complex geotechnical failure models. Geotechnical jargon and standard geotechnical knowledge can assist in identifying causal relationships between variables. The most challenging tasks in the proposed methodology are: (i) defining the pieces of cause-effect relationships based on the proposed failure hypotheses and (ii) eliciting the probabilistic relationships (i.e., the CPTs). In other words, each DAG should reflect the cause-effect relationship between variables (nodes), and the CPTs should account for the probabilities between variables. Experienced geotechnical engineers with abductive reasoning skills should complete these tasks.

In summary, the Bayesian method for comparing hypotheses can be used to support decisions regarding the causes of geotechnical failures. Since the method is based on a probabilistic framework, the decision-making process is traceable, biases are considerably reduced, and arbitrary conclusions can be avoided. Furthermore, the method can deal with multiple pieces of evidence from different sources and estimates the influence of the amount of evidence on selecting the most probable cause of failure.

6.2 Limitations

The following paragraphs discuss the weaknesses and limitations of the proposed Bayesian methodology and emphasize the need for further research in Bayesian statistics for FGE.

6.2.1 General Limitations

One of the limitations of the proposed methodology is the difficulty of translating geotechnical models into Bayesian networks or posterior odds ratios. This difficulty arises from the complexity of some models, in which many input variables may be needed. In order to keep the complexity to a minimum, it is necessary to keep the variables to those that may significantly influence the results. Therefore, before creating a probabilistic model, it is necessary to evaluate the influence of each variable on the geotechnical behavior and assign probability functions only to those that may have influence. Local and global sensitivity analyses can be used to examine the impact of uncertainties in input and output variables.

The lack of information associated with geotechnical failures may be another limitation in applying the proposed Bayesian methodology. Since geotechnical failures are rare and unique, the available information is often scarce or non-existent. Therefore, each forensic geotechnical analysis requires unique data from the site investigation, laboratory tests, design documentation, and as-built records. Although standard geotechnical knowledge can be used to analyze failures, each analysis should use its own geometry, geomechanical properties, and failure mechanism information. Sometimes this information is difficult to acquire due to economic restraints or practical limitations such as inaccessibility to failure sites or inability to collect evidence. In any case, all available information should be used to formulate hypotheses and evaluate the causes of failure.

A complete forensic analysis of geotechnical failures using the proposed Bayesian methodology may be challenging due to the many different failure hypotheses that can be formulated. In fact, hypotheses from the stakeholders should only be considered if they are credible and verifiable using current geotechnical and probabilistic knowledge. In other words, forensic analyses cannot include hypotheses that cannot be represented by mathematical, probabilistic, or geotechnical models. Falsifiability, refutability, and unfalsifiability fallacy concepts (Bowen, 2018) are crucial aspects to consider when formulating hypotheses about causes of failure. An additional aspect is that theoretical models of some geotechnical phenomena are still under research. For example, the unsaturated, thermodynamic, and chemical behavior of soils are active research areas. Consequently, the cause-effect relationships are not yet fully understood, and solutions must rely on empirical relationships or expert opinion.

6.2.2 Posterior Odds Ratio (POR) Technique

Regarding the posterior odds ratio technique described in Section 4.3.1, two drawbacks were identified: (i) sensitivity to prior odds and (ii) limitations to including expert knowledge. As demonstrated in Section 4.2 and Section 5.2.3, prior odds values can significantly influence the posterior odds ratio values. The main challenge consists in selecting appropriate values for the prior odds ratio. For example, when two failure hypotheses are equally likely, the forensic investigator should assign 1.0 to the prior odds ratio. Conversely, when a hypothesis is unlikely, high (or low) values such as 100 or 1000 (0.01 or 0.001) should be assigned. Assigning values to the prior odds ratio is a subjective task that requires expertise and knowledge of similar geotechnical failures. In case of doubt, when failure hypotheses are well formulated and seem equally likely, all geotechnical forensic analyses should use 1.0 for the prior odds ratio.

Including expert, empirical or common knowledge in the POR technique may be difficult if a mathematical model cannot represent this knowledge. Moreover, combining expert opinion and geotechnical models without a probabilistic framework is virtually impossible. In the case of the proposed Bayesian methodology, expert opinion should be included in the likelihood term of Bayes' theorem (Equation 2-9). For example, an expert can define a geotechnical failure using a deformation value based on his/her expertise. However, another expert can define a geotechnical failure using different deformation values. In order to include all available expert

opinions, the geotechnical failure should be defined probabilistically using a likelihood term. Consequently, likelihood values should be defined through surveys (i.e., statistical data) or epistemic probabilities (Kadane & Schum, 1998).

6.2.3 Bayesian Networks (BN) Technique

The Bayesian network technique described in Section 4.3.2 is subject to at least the following limitations: discretization of continuous variables, translating expert knowledge into probability distribution functions (PDF), construction of DAG structures, and computational demand. Each limitation is discussed below.

Discretization of continuous variables

Most variables used in geotechnical engineering to describe soil/rock properties are continuous. BNs can deal with continuous variables, although their use is limited to probability distribution functions such as normal or beta. Therefore, the standard practice discretize the variables and assign them a discrete domain. Friedman & Goldszmidt (1996) pointed out that discretization implies a tradeoff between two aspects. On the one hand, discretization leads to an efficient representation of cause-effect relationships between variables, especially when those relationships are highly nonlinear. In this way, it produces models that can be used for decision-making, as in the case of FGE. On the other hand, all discretization processes create an unavoidable loss of information that may affect the results. Therefore, a useful BN for FGE should be able to model hypotheses of geotechnical failures without losing accuracy and representativeness.

Discretization of a continuous variable is still an unsolved problem. Although several techniques, such as static discretization, dynamic discretization (Neil et al., 2007), and credal sets (Antonucci, 2018) are available, there is not yet a discretization tool for solving the problem of information loss without affecting computational cost. BNs used in FGE are no exception to discretization problems. For example, in the forensic analyses described in Sections 4.3 and 5.2.3, the number and width of the intervals were defined based on the model's accuracy and computational cost. Several trials were required to identify the number and width of intervals for each geotechnical variable.

In some cases, geotechnical variables are easy to discretize. The factor of safety (FoS) and some deformation variables are good illustrations of a straightforward discretization using standard geotechnical knowledge. For example, in the forensic analysis discussed in Chapter 5, the FoS was divided into groups ≤ 1.0 and > 1.0. The group ≤ 1.0 represents all geotechnical cases in which failure occurs, while the group ≥ 1.0 includes all stable cases. In addition, deformation variables such as settlement and horizontal displacement are discretized using the criteria specified in local engineering codes (e.g., settlement values higher than 0.3 m and horizontal displacements higher than 2% of the excavation depth are defined as serviceability failures).

Translating expert knowledge into probability distribution functions

As mentioned in Section 2.2.8, a BN can include human expert knowledge to support decision-making in FGE. Including expert knowledge is helpful when there are significant concerns regarding the uncertainty of cause-effect relationships between variables. However, translating expert knowledge into useful PDFs or CPTs can be challenging. Uusitalo (2007) breaks down this challenge into two main areas of difficulty: (i) classical statistics used by experts mainly focus on confidence intervals and point estimates, such as averages and standard deviations. Consequently, most human experts are not trained in Bayesian statistics and updating probability concepts. (ii) human experts commonly assign probabilities using large databases. Therefore, assigning probabilities to knowledge when information is unavailable is a complex task.

The latter reason is of particular interest in FGE because geotechnical failures are considered rare events. Large databases with similar geotechnical failures are uncommon. Moreover, when these databases are available, they should be used cautiously since each failure is unique, and general trends cannot be easily inferred. Heuristic reasoning can help allocate probabilities to rare events and supports the elicitation of CPT. The probabilities assigned to the Eo (expert opinion) variable in the Green Office example of Section 5.3.8 illustrate the use of heuristic reasoning in FGE. Although heuristic reasoning can support some BN tasks, it may lead to biased outcomes that could invalidate the forensic analysis. Boutang & De Lara (2015) provide good examples of some of the biases behind them.

Construction of DAG structures

In many forensic geotechnical problems, the DAG structure that reflects the cause-effect relationships between variables can be constructed using physical-mathematical models.

Since these models have been validated through experimental or real-scale prototypes, and the geotechnical community widely accepts them, the resulting DAG implies little uncertainty. For example, the DAGs constructed in Sections 4.3 and 5.2.3 use physical-mathematical models. Therefore, no uncertainty is considered during their construction.

In some forensic problems, identifying causality and constructing useful DAGs can be challenging (Uusitalo, 2007). Forensic studies that involve interaction between environmental and geotechnical variables (e.g., slope stability, volumetric changes due to climate variation, and weathering) are good illustrations of problems where constructing useful DAGs is difficult. Structuring multidisciplinary teams and using a standard reliability language could help develop useful DAGs and consistent forensic geotechnical analysis.

Computational demand

Computational demand is a significant concern in BNs with several variables, even though the number of probability relationships between variables is reduced by the chain rule (Section 2.2.6). Bensi et al. (2013) pointed out that for moderately sized BNs, the computational demand could be intractable, especially when exact algorithms such as variable elimination are used. Given that the computational demands may restrict the use of BNs in FGE, the following aspects require special attention:

- Reduce the number of variables and use only those that impact the outcomes. In other words, before constructing a BN, a sensitivity analysis should be performed in order to detect the relevant variables with the most significant influence on the results.

- Discretize relevant variables using a few intervals. As discussed at the beginning of this section, BNs require the discretization of continuous variables into a finite number of intervals. The number of intervals should be defined based on representativeness, loss of information criteria, and computational demand. The discretization should also be based on standard geotechnical knowledge, failure criteria, and deformation thresholds.

- Assign probability values to cause-effect relationships (i.e., eliciting CPT values). An efficient discretization process reduces the size of CPTs. Moreover, when parent and

child nodes (variables) are binary (i.e., two states), and child nodes only depend on one or two parents, the probability values in the CPT can be assigned by human experts. Eliciting probability values for CPTs or providing the shape of a probability distribution can be challenging for human experts when variables have several states and depend on many parent nodes. Morgan & Henrion (1990) present a comprehensive analysis of the cognitive difficulties of human experts in this regard.

- Computational demand may be a major concern if probability values that populate CPTs are estimated from complex numerical models. The examples of Sections 4.2 and 5.2 are good illustrations of computational demand issues. In the case of the forensic analysis of Section 4.3, 100,000 slope stability calculations using limit equilibrium equations were performed using D-Stability (Meij & Deltares, 2020). In order to obtain accurate estimates of the conditional probabilities, the 100,000 slope stability outcomes were considered more than sufficient. Convergence plots of **Figure 4-16** demonstrate this condition.

On the other hand, the forensic analysis described in Section 5.3 required a more sophisticated geotechnical model that included a complex finite element analysis. This sophisticated model was formulated based on failure and deformation criteria defined in the hypotheses. The number of finite element analyses was restricted to 8,000, given the limitations of the software and memory capacity. However, even with the restricted number of analyses, the number of outcomes was ample to complete the CPTs and use a BN for determining the most probable cause of failure. Both analyses, but especially the example of Section 5.3, required a compromise between the representativeness of the failure model and computational demand. Further research in this regard is required.

7. Conclusions and Suggestions for Future Work

7.1 Conclusions

7.1.1 Diagnosis of Standard Practices in Forensic Geotechnical Engineering

Deterministic analysis for hypotheses comparison is the standard practice in FGE. Deterministic models via back analysis are by far the most preferred tools used by geotechnical engineers. However, since analyzes are deterministic and the results are mainly based on expert judgment, some conclusions about the causes of geotechnical failures sometimes seem arbitrary. Two reasons may explain this situation. The first reason is associated with the uncertainty in soil/rock behavior and the uncertainty in failure scenarios. The second reason is related to the standard practices of FGE and how expert judgment is used.

In the case of soil/rock behavior and uncertainty in failure scenarios, geotechnical engineers recognize that those materials and the forces acting on them are highly uncertain. However, probabilistic tools to address uncertainty are rarely used in forensic geotechnical investigations. Consequently, failure scenarios (hypotheses) are usually characterized by deterministic models, which in turn lead to deterministic results. Deterministic analyses are contrary to the behavior of geotechnical models and may lead to erroneous or biased results.

Standard practices in FGE include comparing deterministic results from back-analysis against the collected evidence and supporting/disregarding hypotheses (failure scenarios) based on expert judgment. Since the results are deterministic and the expert judgment is difficult to trace, probabilistic comparison between hypotheses cannot be carried out. For example, deterministic analyses are unable to provide comparison such as (1) how many times the hypothesis H_i better supports the evidence than the hypothesis H_j? or (2) although

hypotheses H_i, H_j and H_k can explain the failure, which of them is the most likely explanation of the cause of failure? Current technical literature in FGE does not include probabilistic comparisons such as those mentioned above.

As a result, current standard practices of FGE require the support of probabilistic techniques that align with the actual behavior of geotechnical systems and the uncertainty associated with the process of hypotheses comparison. Bayesian statistics and Bayesian inference are promising tools that can support decisions in FGE. This doctoral thesis has aimed to explore these tools in order to establish whether they can provide traceable and reliable conclusions regarding the causes of geotechnical failures. In addition, the proposed Bayesian methodology is expected to improve transparency in decision-making processes in FGE.

7.1.2 Supporting Decisions in FGE. The Proposed Bayesian Methodology: Posterior Odds Ratio and Bayesian Networks

This doctoral thesis proposed a Bayesian methodology for supporting decisions about the causes of failures with geotechnical origins. The methodology applies two Bayesian tools: Bayesian inference via Posterior Odds Radio (POR) and Bayesian Networks (BN). Both tools are used to compare hypotheses and identify the most probable cause that led to geotechnical failures. Furthermore, the methodology facilitates the assessment of how the quantity of evidence influences the identification of failure causes.

The methodology includes three main stages: preliminary steps, construction of probabilistic failure models, and probabilistic hypotheses comparisons. The first stage aims to collect all available evidence required for the analysis and formulate credible hypotheses about the causes of failure. The purpose of collecting evidence is to validate or disregard hypotheses. Therefore, evidence should be systematically collected to avoid altering or losing information that could lead to erroneous results. In the case of hypotheses formulation, all hypotheses must be formulated based on the predictable behavior of materials and expected forces acting on the geotechnical structure. In geotechnical engineering, credible hypotheses are associated with changes in pore-water pressures, external/internal forces, and soil/rock behavior. Finally, hypotheses must fulfill two requirements: they should be verbalized so that cause-effect relationships between variables are identifiable, and they must be collectively

exhaustive, i.e., encompass the entire range of possible causes that could explain the geotechnical failure.

The second stage defines relevant random variables and develops probabilistic failure models. By default, all geotechnical variables are random. However, the randomness of some variables has limited influence on the geotechnical behavior. Therefore, random variables must be selected according to criteria such as hypotheses representativeness, service/strength limit state analyses, and performance of the geotechnical system. Regarding the probabilistic failure model, two aspects are relevant: (i) it must include all failure hypotheses and performance variables, and (ii) it must be simple but sufficiently accurate to provide reliable results.

The last stage is the core of the Bayesian methodology proposed in this thesis. It aims to compare probabilistically hypotheses about causes of geotechnical failures using two Bayesian tools: posterior odds ratio (POR) and Bayesian networks (BN). The POR tool compares the chance of hypothesis H_i being true to the chance of hypothesis H_j being true by means of Bayes' theorem (Bayes rule). The results obtained from the comparison are interpreted based on criteria defined in Jeffreys's (1961) scale and Kass & Raftery's (1995) modified scale. In the case of the BN tool, information about the probabilistic failure model is encoded in the BN structure. This encoding allows for more complex hypotheses comparison in the form of probability queries.

A benchmark exercise, a well-documented case from the technical literature, and a case study are presented to validate the applicability of the proposed Bayesian methodology. The benchmark exercise is a modification of the ECRT7 excavation (Schweiger, 2006), in which the POR and BN tools were used. The well-documented case also examines the causes of the Breitenhagen levee failure (Kool et al., 2019) using POR and BN tools. Finally, the case study applies both Bayesian tools to find the most probable cause that led to the failure of a deep excavation in Bogotá, Colombia (Unal, 2012). The following conclusions can be drawn from the results of these analyses:

- The POR tool can identify the most probable cause of a geotechnical failure even when the amount of evidence is scarce. After applying the stages proposed in the Bayesian methodology, the POR tool identified the most probable cause of failure in all case

studies. Moreover, it was able to estimate how many times better one hypothesis explained the geotechnical failure than the others.

- The amount of evidence used to compare hypotheses through the POR tool is essential to identify the cause of failure. Although including a few pieces of evidence may provide insight into the cause of failure, including all available evidence increases the probability that one hypothesis better explains the failure than others. In all case studies, the progressive increase in the amount of evidence also increased the probability of one of the hypotheses, especially when the stability condition and deformation measurements were included together.

- Similarly to the POR tool, the BN tool can also identify the most probable cause of a geotechnical failure. Moreover, the structure of the BN can encode additional information, such as human expert opinion and probability relationships between qualitative variables that cannot be included in the POR tool. These aspects of BNs allow for complex hypotheses comparisons and the evaluation of additional hypotheses not included in the early stages of forensic analyses. In addition, BNs can identify the most probable explanation (KMPE algorithm) among several probable explanations about the causes of a geotechnical failure.

- Large and complex geotechnical failures could be analyzed through BNs. Causality information between variables involved in a geotechnical failure can be translated into small pieces of causal relationships. Each piece may be constructed from fragments of semantic expression (idioms from the geotechnical jargon), reflecting cause-effect relationships between variables (nodes). Then, the pieces of causal relationships are depicted by simple direct acyclic graphs (DAG), which are joined together to form a large DAG. The probabilistic relationships between nodes in the DGA can be estimated by combining probabilistic failure models and human expert opinion.

- The failure case studies analyzed in this research went beyond the traditional tools applied in FGE. In addition to standard back analysis, probabilistic analyses via the proposed Bayesian methodology were also employed. The methodology recognizes the uncertainty associated with the engineering properties of soil/rock materials and

failure scenarios and expands the traditional back analysis to include the analysis of several hypotheses. This expansion allows for probability comparisons between hypotheses which can support decisions regarding the causes of geotechnical failures.

- The probabilistic comparison between hypotheses related to causes of geotechnical failures is the main contribution of this research to FGE knowledge. The results suggest that POR and BN tools can determine the most probable cause of failure. The practical implication of these findings is that Bayesian probabilistic techniques can support decisions regarding the causes of geotechnical failures. Moreover, the proposed methodology allows for tracking of the decision-making process, improves transparency and the reliability of conclusions. Therefore, the proposed methodology should reduce both bias and the perception of the arbitrariness related to conclusions in forensic geotechnical investigations.

7.2 Suggestions for Future Work

Based on the main results and limitations of this doctoral thesis, the following objectives for future work are suggested:

- To develop a methodology to assist in the construction of BNs in geotechnical engineering. The methodology may be valuable for translating complex geotechnical problems into reliable and helpful BNs. For example, in addition to applications of BNs in FGE, BNs can be beneficial in diagnosis, design optimization, and inclusion of expert judgment. Semantic expressions (geotechnical jargon), as used in this thesis, should be the starting point for this methodology, given their ability to indicate causal relationships between variables. Well-developed methodologies for constructing BNs in medicine and forensic science should be used as a baseline. For example, BNs for disease diagnosis or BNs used in criminalistic can be adapted to the characteristics of geotechnical engineering.

- Expand the applicability of the proposed Bayesian methodology by exploring more geotechnical failure cases. Future case studies should give priority to current challenging tasks such as hypotheses formulation, discretization of continuous

variables, complex probabilistic failure models, and CPT elicitation. For example, simplifying complex geotechnical models while maintaining accuracy and representativeness is an exciting field of research. In addition, well-known and well-documented geotechnical failures may be analyzed using the proposed Bayesian methodology in order to reinforce its applicability.

- Future work should involve the development of a probabilistic framework to include expert judgment for decision-making in geotechnical engineering. The Bayesian inference could be the probabilistic technique that supports the framework, given its ability to mimic human reasoning. The starting point for developing the framework may include well-established methodologies such as those used in risk analysis of nuclear power generation. Once the framework has been defined, Bayesian inference can be adapted to the characteristics of expert judgment in geotechnical engineering.

- Propose a multidisciplinary work between engineers and lawyers to improve the practice of FGE and the resolution of legal lawsuits. Multidisciplinary work should focus on the following aspects: (1) define a common language for engineers and lawyers regarding geotechnical failures, (2) improve oral skills of geotechnical engineers in order to explain technical aspects of failures in court, (3) train lawyers in engineering failures, and (4) train geotechnical engineers in litigation.

References

Allenby, G. M. (1990). Hypothesis Testing with Scanner Data: The Advantage of Bayesian Methods. *Journal of Marketing Research*, *27*(4), 379–389. https://doi.org/10.1177/002224379002700401

Alonso, E. E., Pinyol, N. M., & Puzrin, A. M. (2010). Geomechanics of failures. Advanced topics. In *Geomechanics of Failures. Advanced Topics*. https://doi.org/10.1007/978-90-481-3538-7

Alonso, E. E., Pinyol, N. P., & Fernández, P. (2016). *Caisson Failure Induced by Wave Action BT - Forensic Geotechnical Engineering* (V. V. S. Rao & G. L. Sivakumar Babu (Eds.); pp. 45–93). Springer India. https://doi.org/10.1007/978-81-322-2377-1_4

Antonucci, A. (2018). Reliable Discretisation of Deterministic Equations in Bayesian Networks. *The Thirty-Second International Florida Artificial Intelligence Research Society Conference (FLAIRS-32)*, 453–457.

ASCE. (2018). Guidelines for Failure Investigation. In R. S. Barrow, R. W. Anthony, K. J. Beasley, & S. M. Verhulst (Eds.), *Guidelines for Failure Investigation*. https://doi.org/10.1061/9780784415122

Babu, G. L. S. (2016). Briefing: Forensic geotechnical engineering. *Proceedings of the Institution of Civil Engineers - Forensic Engineering*, 1–4.

Babu, G. L., & Singh, V. P. (2016). *Back Analyses in Geotechnical Engineering* (V. V. S. Rao & G. L. Sivakumar Babu (Eds.); pp. 113–118). Springer India. https://doi.org/10.1007/978-81-322-2377-1_7

Babu, G., Raja, J., Munwar Basha, B., & Srivastava, A. (2016). Forensic Analysis of Failure of Retaining Wall. In V. V. S. Rao & G. L. Sivakumar Babu (Eds.), *Forensic Geotechnical Engineering* (pp. 451–465). Springer India. https://doi.org/10.1007/978-81-322-2377-1_30

Baecher, G. B. (2017). Bayesian Thinking in Geotechnics. *Geo-Risk 2017*, *June*, 1–18. https://doi.org/10.1061/9780784480694.001

Baecher, G. B., & Christian, J. T. (2003). *Reliability and statistics in geotechnical engineering* (Issue 1). John Wiley & Sons Ltd. https://doi.org/10.1198/tech.2005.s838

Bea, R. (2006). Reliability and Human Factors in Geotechnical Engineering. *Journal of Geotechnical and Geoenvironmental Engineering - J GEOTECH GEOENVIRON ENG*, *132*. https://doi.org/10.1061/(ASCE)1090-0241(2006)132:5(631)

Bell, G. R. (2000). Engineering Investigation of Structural Failures. In R. T. Ratay (Ed.),

Forensic Structural Engineering Handbook. McGraw-Hill.

Bensi, M., Kiureghian, A. Der, & Straub, D. (2013). Efficient Bayesian network modeling of systems. *Reliability Engineering and System Safety*, *112*, 200–213. https://doi.org/10.1016/j.ress.2012.11.017

Bensi, M. T. (2010). *A Bayesian Network Methodology for Infrastructure Seismic Risk Assessment and Decision Support*. University of California, Berkeley.

Bentley Systems. (2020). *Paxis 2D. Connect Edition* (20.04; p. 240). Bentley.

Berti, M., Martina, M. L. V., Franceschini, S., Pignone, S., Simoni, A., & Pizziolo, M. (2012). Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *Journal of Geophysical Research: Earth Surface*. https://doi.org/10.1029/2012JF002367

Biedermann, A., & Taroni, F. (2012). Bayesian networks for evaluating forensic DNA profiling evidence: A review and guide to literature. *Forensic Science International: Genetics*, *6*(2), 147–157. https://doi.org/10.1016/j.fsigen.2011.06.009

Biedermann, A., Taroni, F., Delemont, O., Semadeni, C., & Davison, A. C. (2005). The evaluation of evidence in the forensic investigation of fire incidents (Part I): An approach using Bayesian networks. *Forensic Science International*, *147*(1), 49–57. https://doi.org/10.1016/j.forsciint.2004.04.014

Bolstad, W. M. (2010). Understandeing Computational Bayesian Statistics. In P. Giudisi, G. Givens, & B. Mallik (Eds.), *Wiley Series in Computational Statistic*. John Wiley & Sons, Inc.

Boutang, J., & De Lara, M. (2015). The biased mind: How evolution shaped our psychology including anecdotes and tips for making sound decisions. In *The Biased Mind: How Evolution Shaped our Psychology Including Anecdotes and Tips for Making Sound Decisions*. https://doi.org/10.1007/978-3-319-16519-6

Bowen, J. (2018). Unfalsifiability. In *Bad Arguments* (pp. 403–406). https://doi.org/https://doi.org/10.1002/9781119165811.ch99

Brady, S. P. (2012). Role of the forensic process in investigating structural failure. *Journal of Performance of Constructed Facilities*, *26*(1), 2–6. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000274

Burland, J. (2012). Chapter 4 The geotechnical triangle. In *ICE manual of geotechnical engineering: Volume I* (pp. 17–26). Thomas Telford Ltd. https://doi.org/doi:10.1680/moge.57074.0017

Burland, J. B., Jamiolkowski, M., & Viggiani, C. (1998). Stabilising the leaning tower of Pisa. *Bulletin of Engineering Geology and the Environment*, *57*(1), 91–99. https://doi.org/10.1007/s100640050025

Calvello, M., Cuomo, S., & Ghasemi, P. (2017). The role of observations in the inverse analysis of landslide propagation. *Computers and Geotechnics*, *92*, 11–21.

Campos, L. M. De, Gamez, J. A., & Moral, S. (2001). Simplifying explanations in Bayesian belief networks. *International Journal of Uncertainty, Puzziness and Knowledge-Based Systems*, *9*(4), 461–489.

Caracol Radio. (2012). *Nuevo dolor de cabeza generan obras en la Carrera 11 con 98*.

https://caracol.com.co/radio/2012/01/20/bogota/1327066020_609716.html

Carper, K. L. (2000). *Forensic Engineering*. Taylor \& Francis. https://books.google.com.bo/books?id=gIu9BwAAQBAJ

Chen, S. H., & Pollino, C. A. (2012). Good practice in Bayesian network modelling. *Environmental Modelling & Software*, *37*, 134–145. https://doi.org/https://doi.org/10.1016/j.envsoft.2012.03.012

Chowdhury, R. N. (1987). Aspects of the Vajont slide. *Engineering Geology*, *24*(1), 533–540. https://doi.org/https://doi.org/10.1016/0013-7952(87)90085-8

Cleland, C. E. (2001). Historical science, experimental science, and the scientific method. *Geology*, *29*(11), 987–990. http://ecee.colorado.edu/ecen5009/Resources/Cleland01.pdf

Correa, M., Bielza, C., & Pamies-Teixeira, J. (2009). Comparison of Bayesian networks and artificial neural networks for quality detection in a machining process. *Expert Systems with Applications*, *36*(3 PART 2), 7270–7279. https://doi.org/10.1016/j.eswa.2008.09.024

Dahl, F. A., Grotle, M., Šaltyte Benth, J., & Natvig, B. (2008). Data splitting as a countermeasure against hypothesis fishing: With a case study of predictors for low back pain. *European Journal of Epidemiology*, *23*(4), 237–242. https://doi.org/10.1007/s10654-008-9230-x

Day, R. W. (2010). *Forensic Geotechnical and Foundation Engineering* (p. 528).

Drezner, Z., & Zerom, D. (2016). A Simple and Effective Discretization of a Continuous Random Variable. *Communication in Statistics- Simulation and Computation*, *45*, 3798–3810. https://doi.org/10.1080/03610918.2015.1071389

Druzdzel, M. J., & Gaag, L. C. van der. (2000). Building probabilistic networks: "Where do the numbers come from?" guest editors' introduction. *IEEE Transactions on Knowledge and Data Engineering*, *12*(4), 481–486. https://doi.org/10.1109/TKDE.2000.868901

Dykes, A. P., & Bromhead, E. N. (2018). The Vaiont landslide: re-assessment of the evidence leads to rejection of the consensus. *Landslides*, *15*(9), 1815–1832. https://doi.org/10.1007/s10346-018-0996-y

Ering, P., & Sivakumar Babu, G. L. (2017). A Bayesian framework for updating model parameters while considering spatial variability. *Georisk*, *11*(4), 285–298. https://doi.org/10.1080/17499518.2016.1255760

Feng, X. (2015). *Application of Bayesian approach in geotechnical engineering* [Universidad Politécnica de Madrid]. http://oa.upm.es/37270/

Fenton, N., & Neil, M. D. (2019). *Risk assessment and decision analysis with bayesian networks* (Second Edi). CRC Press. https://doi.org/10.1201/b13102

Friedman, N., & Goldszmidt, M. (1996). Discretizing Continuous Attributes While Learning Bayesian Networks. *Icml*, 157–165. https://doi.org/10.1001/archinte.159.12.1359

Friedman, N., Goldszmidt, M., Heckerman, D., & Russell, S. (1997). Challenge: What is the impact of Bayesian networks on learning? *IJCAI International Joint Conference on Artificial Intelligence*, *1*, 10–15.

Garbolino, P., & Taroni, F. (2002). Evaluation of scientific evidence using Bayesian networks. *Forensic Science International*, *125*(2–3), 149–155. https://doi.org/10.1016/S0379-0738(01)00642-9

Garcia-Feria, W. M., Colmenares Montañez, J. E., & Hernandez Perez, G. J. (2022). Testing the Causes of a Levee Failure Using Bayesian Networks. *Ingeniería*, *27*(2), e18538. https://doi.org/10.14483/23448393.18538

Gelman, A., Hill, J., & Yajima, M. (2012). Why We (Usually) Don't Have to Worry About Multiple Comparisons. *Journal of Research on Educational Effectiveness*, *5*(2), 189–211. https://doi.org/10.1080/19345747.2011.618213

Gelman, A., & Tuerlinckx, F. (2000). Type S error rates classical and Bayesian single and multiple compparison procedures. *Computational Statistics*, *15*(3), 373–390. https://doi.org/10.1007/s001800000040

Gens, A. (2010). Soil-environment interactions in geotechnical engineering. *Geotechnique*, *60*(1), 3–74. https://doi.org/10.1680/geot.9.P.109

Gilbert, R. B. (2016). Important Role of Uncertainty in Forensic Geotechnical Engineering. In V. V. S. Rao & G. L. Sivakumar Babu (Eds.), *Forensic Geotechnical Engineering* (pp. 493–503). Springer India. https://doi.org/10.1007/978-81-322-2377-1_32

Greenspan, H. F., O'kon, J. A., Beasley, K. J., & Ward, J. . (1989). *Guidelines for Failure Investigation* (ASCE (Ed.)).

Griffiths, D., & Fenton, G. A. (2007). *Probabilistic methods in geotechnical engineering*.

Grubert, P. (2013). *Saaledeich bei Breitenhagen Geotechnische Untersuchungen der Bruchstelle Empfehlungen zur Sanierung*.

Guerra, S. Z. (1992). Severe Soil Deformations, Leveling and Protection at the Metropolitan Cathedral in Mexico City. *APT Bulletin: The Journal of Preservation Technology*, *24*(1/2), 28–35. https://doi.org/10.2307/1504308

Hasan, S., & Najjar, S. (2013). Probabilistic Back Analysis of Failed Slopes using Bayesian Techniques. *Geo-Congress 2013*, *231 GSP*, 1013–1022. https://doi.org/10.1061/9780784412787.103

Holický, M., Marková, J., & Sýkora, M. (2013). Forensic assessment of a bridge downfall using Bayesian networks. *Engineering Failure Analysis*, *30*, 1–9. https://doi.org/10.1016/j.engfailanal.2012.12.014

Houck, M. M. (2006). Forensic Science: An Introduction to Scientific and Investigative Techniques. *Journal of Forensic Sciences*, *51*(1), 205. https://doi.org/https://doi.org/10.1111/j.1556-4029.2005.00042.x

Hu, J. L., Tang, X. W., & Qiu, J. N. (2016). Assessment of seismic liquefaction potential based on Bayesian network constructed from domain knowledge and history data. *Soil Dynamics and Earthquake Engineering*, *89*, 49–60. https://doi.org/10.1016/j.soildyn.2016.07.007

Hwang, R. N. (2016). Back Analyses in Forensic Geotechnical Engineering. In V. V. S. Rao & G. L. Sivakumar Babu (Eds.), *Forensic Geotechnical Engineering* (pp. 131–143). Springer India. https://doi.org/10.1007/978-81-322-2377-1_9

Iai, S. (2016). Backwards Problem in Geotechnical Earthquake Engineering. In V. V. S. Rao & G. L. Sivakumar Babu (Eds.), *Forensic Geotechnical Engineering* (pp. 187–196). Springer India. https://doi.org/10.1007/978-81-322-2377-1_13

Iwasaki, Y. (2016). *A Case Study of Observational Method for a Failed Geotechnical Excavation in Singapore BT - Forensic Geotechnical Engineering* (V. V. S. Rao & G. L. Sivakumar Babu (Eds.); pp. 535–548). Springer India. https://doi.org/10.1007/978-81-322-2377-1_35

Jeffreys, H. (1961). *Theory of Probability* (Third Edit). Clarendon Press.

Jensen, F. V., & Nielsen, T. D. (2007). Bayesian Networks and Decision Graphs. In *Information Science and Statistics*. Springer, New York, NY. https://doi.org/https://doi.org/10.1007/978-0-387-68282-2

Jessep, R. A., de Mello, L. G., & Rao, V. V. S. (2016). *Technical Shortcomings Causing Geotechnical Failures: Report of Task Force 10, TC 302 BT - Forensic Geotechnical Engineering* (V. V. S. Rao & G. L. Sivakumar Babu (Eds.); pp. 267–295). Springer India. https://doi.org/10.1007/978-81-322-2377-1_19

Johnson, D. H. (1999). The Insignificance of Statistical Significance Testing. *The Journal of Wildlife Management*, *63*(3), 763–772. https://doi.org/10.2307/3802789

Kadane, J. B., & Schum, D. A. (1998). A Probabilistic Analysis of the Sacco and Vanzetti Evidence. In *Technometrics* (Vol. 40, Issue 2). https://doi.org/10.1080/00401706.1998.10485222

Kardon, J. B. (2003). *GUIDELINES FOR FORENSIC ENGINEERING PRACTICE* (J. B. Kardon (Ed.); Second edi, Vol. 7, Issue 2). ASCE. https://doi.org/10.1080/10903120390936923

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/http://dx.doi.org/10.1080/01621459.1995.10476572

Kiureghian, A., Bensi, M., & Straub, D. (2009). Bayesian network methodology for post-earthquake infrastructure risk management. In *Frontier Technologies for Infrastructures Engineering* (pp. 201–214).

Kjærulff, U. B., & Madsen, A. L. (2013). *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis* (M. Jordan, R. Nowak, & B. Scholkopf (Eds.); Second Edi). Springer. https://doi.org/10.1007/978-1-4614-5104-4

Koller, D., & Friedman, N. (2009). Probabilistic Graphical Models: Principles and Techniques. In *Adaptive Computation and Machine Learning*. The MIT Press.

Kool, J. J., Kanning, W., Heyer, T., Jommi, C., & Jonkman, S. N. (2019). Forensic Analysis of Levee Failures : The Breitenhagen Case. *International Journal of Geoengineering Case Histories*, *5*(2), 70–92. https://doi.org/10.4417/IJGCH-05-02-02

Kool, J. J., Kanning, W., Jommi, C., & Jonkman, S. N. (2020). A Bayesian hindcasting method of levee failures applied to the Breitenhagen slope failure. *Georisk*, *0*(0), 1–18. https://doi.org/10.1080/17499518.2020.1815213

Kruschke, J. K. (2010). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences*, *14*(7), 293–300. https://doi.org/https://doi.org/10.1016/j.tics.2010.05.001

Kruschke, J. K. (2015). *Doing Bayesian Data Analysis* (Second edi). Academic Press is an imprint of Elsevier.

Kwan, M., Chow, K.-P., Frank, L., & Lai, P. (2008). Reasoning About Evidence Using Bayesian Networks. In R. I. & S. S. (Eds.), *Advances in Digital Forensics IV. DigitalForensics 2008. IFIP — The International Federation for Information Processing, vol 285* (pp. 275–289). Springer. https://doi.org/https://doi.org/10.1007/978-0-387-84927-0_22

Lacasse, S. (2016). Forensic Geotechnical Engineering Theory and Practice. In V. V. S. Rao & G. L. Sivakumar Babu (Eds.), *Forensic Geotechnical Engineering* (pp. 17–37). Springer India. https://doi.org/10.1007/978-81-322-2377-1_2

Leonards, G. A. (1982). Investigation of Failures. 16th Terzaghi Lecture. *Journal of the Geotechnical Engineering Division. ASCE*, *108*(2), 187–246. https://doi.org/10.1061/AJGEB6.0001241

Marcot, B. G., Steventon, J. D., Sutherland, G. D., & McCann, R. K. (2006). Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. *Canadian Journal of Forest Research*, *36*(12), 3063–3074. https://doi.org/10.1139/x06-135

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*(3), 679–690. https://doi.org/10.3758/s13428-010-0049-5

McGrayne, S. B. (2011). *The theory that would not die : how Bayes' rule cracked the enigma code, hunted down Russian submarines, and emerged triumphant from two centuries of controversy*. Yale University Press.

Meij, R. van der, & Deltares. (2020). *D-Stability. Slope stability software for soft soil engineering* (2020.03.3; p. 167). Deltares.

Melchers, R. E., & Beck, A. T. (2018). *Structural reliability analysis and prediction* (Third Edit). JohnWiley & Sons Ltd.

Mkrtchyan, L., Podofillini, L., & Dang, V. N. (2016). Methods for building Conditional Probability Tables of Bayesian Belief Networks from limited judgment: An evaluation for Human Reliability Application. *Reliability Engineering & System Safety*, *151*, 93–112. https://doi.org/https://doi.org/10.1016/j.ress.2016.01.004

Mohan, V. K. D., Vardon, P. J., Hicks, M. A., & Gelder, P. H. A. J. M. van. (2019). Uncertainty Tracking and Geotechnical Reliability Updating Using Bayesian Networks Varenya. In J. Ching, D.-Q. Li, & J. Zhang (Eds.), *Proceedings ofthe 7th InternationalSymposiumonGeotechnical SafetyandRisk(ISGSR)* (Issue December, pp. 978–981). Research Publishing. https://doi.org/10.3850/978-981-11-2725-0

Morales-Nápoles, O., Delgado-Hernández, D. J., De-León-Escobedo, D., & Arteaga-Arcos, J. C. (2014). A continuous Bayesian network for earth dams' risk assessment: Methodology and quantification. In *Structure and Infrastructure Engineering* (Vol. 10, Issue 5, pp. 589–603). Taylor & Francis. https://doi.org/10.1080/15732479.2012.757789

Morgan, M. G., & Henrion, M. (1990). Monte Carlo and Other Sampling Methods. In *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Poilcy Analysis*. http://ebooks.cambridge.org/ref/id/CBO9780511840609%5Cnhttp://books.google.co

m/books?hl=en&lr=&id=ajd1V305PgQC&oi=fnd&pg=PR11&dq=A+Guide+to+Dealing+with+Uncertainty+in+Quantitative+Risk+and+Poilcy+Analysis&ots=U5HUq7Jo8s&sig=CqVVzpABYsV156jGILhbm6R1nGc

Müller, L. (1968). New Considerations on the Vaiont Slide. *Rock Mechanics & Engineering Geology*, *6*(1–2), 1–91.

Nadim, F., & Liu, Z. Q. (2013). Quantitative risk assessment for earthquake-triggered landslides using Bayesian network. *18th International Conference on Soil Mechanics and Geotechnical Engineering: Challenges and Innovations in Geotechnics, ICSMGE 2013*, *3*(3), 2221–2224.

Neil, M., Fenton, N., Lagnado, D., & Gill, R. D. (2019). Modelling competing legal arguments using Bayesian model comparison and averaging. *Artificial Intelligence and Law*, *27*(4), 403–430. https://doi.org/10.1007/s10506-019-09250-3

Neil, M., Fenton, N., & Nielsen, L. (2000). Building large-scale Bayesian networks. *Knowledge Engineering Review*, *15*(3), 257–284. https://doi.org/10.1017/S0269888900003039

Neil, M., Tailor, M., & Marquez, D. (2007). Inference in hybrid Bayesian networks using dynamic discretization. *Statistics and Computing*, *17*(3), 219–233. https://doi.org/10.1007/s11222-007-9018-y

Noon, R. (2001). *Forensic engineering investigation*. CRC Press.

Noon, R. (2009). *Scientific method: applications in failure investigation and forensic science*. CRC Press.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc.

Peck, R. B. (1969). Advantages and Limitations of the Observational Method in Applied Soil Mechanics. *Géotechnique*, *19*(2), 171–187. https://doi.org/10.1680/geot.1969.19.2.171

Phoon, K.-K., & Kulhawy, F. H. (1999a). Characterization of geotechnical variability. *Canadian Geotechnical Journal*, *36*(4), 612–624. https://doi.org/10.1139/t99-038

Phoon, K.-K., & Kulhawy, F. H. (1999b). Evaluation of geotechnical property variability. *Canadian Geotechnical Journal*, *36*(4), 625–639. https://doi.org/10.1139/t99-039

Phoon, K. K., Sivakumar Babu, G. L., & Uzielli, M. (2016). Role of Reliability in Forensic Geotechnical Engineering. In V. V. S. Rao & G. L. Sivakumar Babu (Eds.), *Forensic Geotechnical Engineering* (pp. 467–491). Springer India. https://doi.org/10.1007/978-81-322-2377-1_31

Popescu, M. E., & Schaefer, V. R. (2016). *Back Analysis of Slope Failures to Design Landslide Stabilizing Piles* (V. V. S. Rao & G. L. Sivakumar Babu (Eds.); pp. 119–130). Springer India. https://doi.org/10.1007/978-81-322-2377-1_8

Popper, K. (2002). *The Logic of Scientific Discovery* (2nd Editio). Routledge. https://doi.org/https://doi.org/10.4324/9780203994627

Poulos, H. G. (2016). A Framework for Forensic Foundation Engineering. In V. V. S. Rao & G. L. Sivakumar Babu (Eds.), *Forensic Geotechnical Engineering* (pp. 1–15). Springer India. https://doi.org/10.1007/978-81-322-2377-1_1

Puzrin, A. M., Alonso, E. E., & Pinyol, N. M. (2010). *Geomechanics of Failures*. Springer Science+Business Media B.V.

Radio Santafe. (2012). *Distrito anuncia que abrir la carrera 11 con 98 depende de pruebas de carga*. https://www.radiosantafe.com/2012/08/14/distrito-anuncia-que-abrir-la-carrera-11-con-98-depende-de-pruebas-de-carga/

Rao, V. V. S. (2016). Guidelines for Forensic Investigation of Geotechnical Failures. In V. V. S. Rao & G. L. Sivakumar Babu (Eds.), *Forensic Geotechnical Engineering* (pp. 39–44). Springer India. https://doi.org/10.1007/978-81-322-2377-1_3

Rao, V. V. S., & Babu, G. L. . (2009). Administrative report: TC 40 - Forensic geotechnical engineering. *Proceedings of the 17th International Conference on Soil Mechanics and Geotechnical Engineering: The Academia and Practice of Geotechnical Engineering*, *5*, 3801–3804. https://doi.org/10.3233/978-1-60750-031-5-3801

Rao, V. V. S., & Babu, G. L. S. (2016). *Forensic Geotechnical Engineering* (First Edit). Springer India 2016. https://doi.org/10.1007/978-81-322-2377-1

Ratay, R. T. (Ed.). (2000). *Forensic Structural Engineering Handbook*. McGraw-Hill.

Robertson, P. K., & Campanella, R. G. (1985). Liquefaction Potential of Sands Using the CPT. *Journal of Geotechnical Engineering*, *111*(3), 384–403. https://doi.org/10.1061/(ASCE)0733-9410(1985)111:3(384)

Rocscience Inc. (2006). *Slide* (5.0).

Rohmer, J. (2020). Uncertainties in conditional probability tables of discrete Bayesian Belief Networks: A comprehensive review. *Engineering Applications of Artificial Intelligence*, *88*, 103384. https://doi.org/https://doi.org/10.1016/j.engappai.2019.103384

Sakurai, S., Akutagawa, S., Takeuchi, K., Shinji, M., & Shimizu, N. (2003). Back analysis for tunnel engineering as a modern observational method. *Tunnelling and Underground Space Technology*, *18*(2–3), 185–196. https://doi.org/10.1016/S0886-7798(03)00026-9

Schweiger, H. F. (2006). ERTC7 benchmark exercise. In H. F. Schweiger (Ed.), *Numerical Methods in Geotechnical Engineering* (pp. 3–8). CRC Press.

Scutari, M., & Denis, J.-B. (2015). Bayesian networks with examples in R. In *Texts in Statistical Science*. CRC Press.

Smith, M. (2006). Dam Risk Analysis Using Bayesian Networks. *2006 ECI Conference on Geohazards*.

Sowers, G. F. (1993). Human factors in civil and geotechnical engineering failures. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, *30*(5), 308. https://doi.org/10.1016/0148-9062(93)92542-x

Špačková, O., & Straub, D. (2011). Probabilistic risk assessment of excavation performance in tunnel projects using Bayesian networks: a case study. *Proceedings of the 3rd International Symposium on Geotechnical Safety and Risk*, 651–660. http://vzb.baw.de/e-medien/geotechnical-safety-and-risk-2011/PDF/6 Practical Applications and Case Studies/6_04.pdf

Spross, J. (2016). *Toward a reliability framework for the observational method* (Issue September). https://doi.org/10.1063/1.2719228

Spross, J., & Johansson, F. (2017). When is the observational method in geotechnical engineering favourable? *Structural Safety*, *66*. https://doi.org/10.1016/j.strusafe.2017.01.006

Spross, J., Johansson, F., Uotinen, L. K. T., & Rafi, J. Y. (2016). Using Observational Method to Manage Safety Aspects of Remedial Grouting of Concrete Dam Foundations. *Geotechnical and Geological Engineering*, *34*(5), 1613–1630. https://doi.org/10.1007/s10706-016-0069-8

Stone, J. V. (2013). *Bayes' Rule. A Tutorial Introduction to Bayesian Analysis* (First Edit). Sebtel Press. https://doi.org/10.1002/0470055480.ch6

Straub, D. (2005). Natural hazards risk assessment using Bayesian networks. *9th International Conference on Structural Safety and Reliability ICOSSAR 05 Rome Italy*, *March*, 2509–2516.

Straub, D., & Grêt-Regamey, A. (2006). A Bayesian probabilistic framework for avalanche modelling based on observations. *Cold Regions Science and Technology*, *46*(3), 192–203. https://doi.org/10.1016/j.coldregions.2006.08.024

Szucs, D., & Ioannidis, J. P. A. (2017). When null hypothesis significance testing is unsuitable for research: A reassessment. *Frontiers in Human Neuroscience*, *11*(August). https://doi.org/10.3389/fnhum.2017.00390

Taroni, F., Biedermann, A., Bozza, S., Garbolino, P., & Aitken, C. (2014). Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science. In *Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science* (Second edi). John Wiley & Sons, Ltd Registered. https://doi.org/10.1002/9781118914762

Taroni, F., Biedermann, A., Garbolino, P., & Aitken, C. G. G. (2004). A general approach to Bayesian networks for the interpretation of evidence. *Forensic Science International*, *139*(1), 5–16. https://doi.org/10.1016/j.forsciint.2003.08.004

Tendeiro, J. N., & Kiers, H. A. L. (2019). A review of issues about null hypothesis Bayesian testing. In *Psychological Methods* (Vol. 24, pp. 774–795). American Psychological Association. https://doi.org/10.1037/met0000221

Terwel, K., Schuurman, M., & Loeve, A. (2018). Improving reliability in forensic engineering: The Delft approach. *Proceedings of the Institution of Civil Engineers: Forensic Engineering*, *171*(3), 99–106. https://doi.org/10.1680/jfoen.18.00006

Unal. (2012). *Estudio Para Determinar Las Posibles Causas Que Afectaron Las Redes De Acueducto Y Alcantarillado En El Sector Comprendido Entre Las Calles 97 Y 100 Y Las Carreras 11 Y 11 A Contrato Interadministrativo Unal-Eaab 1-02-26200-0841-2011*.

Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling*, *203*(3–4), 312–318. https://doi.org/10.1016/j.ecolmodel.2006.11.033

Wagenmakers, E.-J., Lee, M., Lodewyckx, T., & Iverson, G. J. (2008). Bayesian Versus Frequentist Inference. In H. Hoijtink, I. Klugkist, & P. A. Boelen (Eds.), *Bayesian*

Evaluation of Informative Hypotheses (pp. 181–207). Springer New York. https://doi.org/10.1007/978-0-387-09612-4_9

Wang, Y., Cao, Z., & Li, D. (2016). Bayesian perspective on geotechnical variability and site characterization. *Engineering Geology*, *203*, 117–125. https://doi.org/10.1016/j.enggeo.2015.08.017

Wu, T. H. (2011). 2008 Peck Lecture: The Observational Method: Case History and Models. *Journal of Geotechnical and Geoenvironmental Engineering*, *137*(10), 862–873. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000509

Xu, Y., & Zhang, L. (2016). Diagnosis of Geotechnical Failure Causes Using Bayesian Networks. In V. V. S. Rao & G. L. Sivakumar Babu (Eds.), *Forensic Geotechnical Engineering* (pp. 103–112). Springer India. https://doi.org/10.1007/978-81-322-2377-1_6

Yildirim, I. (2012). *Bayesian Inference : Gibbs Sampling*. *14627*, 1–6.

Zhang, L. L., Zhang, J., Zhang, L. M., & Tang, W. H. (2010). Back analysis of slope failure with Markov chain Monte Carlo simulation. *Computers and Geotechnics*, *37*(7–8), 905–912.

A. Annex: Scripts and K MPE Results

A1. ERTC7 Benchmark Exercise. Script of the K MPE Algorithm

```
################################################################################
# 2022-12-30
# By: wmgarciaf@unal.edu.co
#
# R SCRIPT para análisis Bayesiano
# Red Bayesiana para análisis ERTC7

# H.F. Schweiger (2006)
# Results from the ERTC7 benchmark exercise
# Numerical Methods in Geotechnical Engineering – Schweiger (ed.)
# © 2006 Taylor & Francis Group, London, ISBN 0-415-40822-9
################################################################################

#-------- PRELIMINARES ----------
# se leen los datos y se depuran para simplificar a 22 tipos de sup. de falla
setwd("C:/Users/mgfer/OneDrive - Universidad Nacional de Colombia/D_drive/Simulations/Excavation")
# 1. Leer el archivo con los resultados de las 100mil simulaciones de D-stability
df = read.csv(file="results_ERTC7_for_BN.csv", header = T)
# Convertir todas las columnas de "numeric" or "character" a "factor"
for (i in 1: ncol(df)) {
  df[, i] = as.factor(df[, i])
}

# -------- GRAFICOS PRELIMINARES -----------
windows(20, 20)
par(mfrow= c(3,2))
barplot(table(df$EL), horiz=F)
barplot(table(df$WT), horiz=F)
barplot(table(df$Sch), horiz=F)
barplot(table(df$SC), horiz=F)
barplot(table(df$SS), horiz=F)

# -------- RED BAYESIANA -----------
library(bnlearn)
model1 = paste0("[EL]","[WT]","[Sch]", "[SC|EL:WT:Sch]","[SS|EL:WT:Sch]")
dag1   = model2network(model1)
windows(20, 20)
graphviz.plot(dag1, layout = "dot")
ERTC7_BN = bn.fit(x = dag1, data = df)

# -------- ABDUCTIVE QUERIES  -----------
# Hypothesis H0 - Random event, act of god
cpquery(ERTC7_BN, event= (EL=="3.5")&(WT=="4")&((Sch=="50")|(Sch=="40")),
      evidence = (SC=="Fail")&((SS=="L1")|(SS=="L2")|(SS=="L3")) , n=10E7)

# Hypothesis H1 - Elevation of the water table
cpquery(ERTC7_BN, event= (EL=="3.5")&((WT=="2")|(WT=="1")|(WT=="0"))&((Sch=="50")|(Sch=="40")),
      evidence = (SC=="Fail")&((SS=="L1")|(SS=="L2")|(SS=="L3")) , n=10E6)

# Hypothesis H2 - Increase of the Surcharge
cpquery(ERTC7_BN, event= (EL=="3.5")&(WT=="4")&((Sch=="60")|(Sch=="70")),
      evidence = (SC=="Fail")&((SS=="L1")|(SS=="L2")|(SS=="L3")) , n=10E7)

# Hypothesis H3 - Short embedded lengh of the wall
cpquery(ERTC7_BN, event= ((EL=="3")|(EL=="2.5")|(EL=="2")|(EL=="1.5"))&(WT=="4")&((Sch=="50")|(Sch=="40")),
      evidence = (SC=="Fail")&((SS=="L1")|(SS=="L2")|(SS=="L3")) , n=10E7)
```

A1. ERTC7 Benchmark Exercise. Script of the K MPE Algorithm

```
# -------- K-MPE -----------
KMPE_df = expand.grid(EL  = unique(df$EL),
                      WT  = unique(df$WT),
                      Sch = unique(df$Sch))

KMPE_df$query = NA
progress_bar = txtProgressBar(min=0, max=length(KMPE_df$EL), style = 1, char="=")
for(i in 1:length(KMPE_df$EL)){
  KMPE_df[i,4]= cpquery(ERTC7_BN, event= (EL==as.character(KMPE_df[i,1]))&
                        (WT ==as.character(KMPE_df[i,2]))&
                        (Sch==as.character(KMPE_df[i,3])),
                      evidence = (SC=="Fail")& ((SS=="L1")|(SS=="L2")|(SS=="L3")), n=1E6)
  setTxtProgressBar(progress_bar, value = i)
}

KMPE_df = KMPE_df[order(-KMPE_df$query),]
KMPE_df$query = formatC(KMPE_df$query,format="e")
#write.csv(x = KMPE_df, file = "KMPE_ERTC7.csv")

#_______
KMPE_df2 = expand.grid(EL  = unique(df$EL),
                       Sch = unique(df$Sch))

KMPE_df2$query = NA
progress_bar = txtProgressBar(min=0, max=length(KMPE_df2$EL), style = 1, char="=")
for(i in 1:length(KMPE_df2$EL)){
  KMPE_df2[i,3]= cpquery(ERTC7_BN, event= (EL==as.character(KMPE_df2[i,1]))&
                         (Sch==as.character(KMPE_df2[i,2])),
                         evidence = (SC=="Fail")& ((SS=="L1")|(SS=="L2")|(SS=="L3"))&(WT=="4"), n=1E6)
  setTxtProgressBar(progress_bar, value = i)
}

KMPE_df2 = KMPE_df2[order(-KMPE_df2$query),]
KMPE_df2$query = formatC(KMPE_df2$query,format="e")
```

A2. K Most Probable Explanations (K MPE) of the ERTC7 Benchmark

K	*EL*	*WT*	*Sch*	*Prob.*
1	1.5	0	70	4.11E-02
2	1.5	0	40	4.10E-02
3	1.5	0	60	4.06E-02
4	1.5	0	50	4.06E-02
5	2	0	50	3.34E-02
6	2	0	70	3.29E-02
7	2	0	60	3.13E-02
8	2	0	40	3.12E-02
9	1.5	1	50	3.12E-02
10	1.5	1	70	3.08E-02
11	1.5	1	40	3.07E-02
12	1.5	1	60	3.05E-02
13	2.5	0	70	2.31E-02
14	2.5	0	40	2.21E-02
15	2	1	70	2.19E-02
16	2.5	0	60	2.18E-02
17	2.5	0	50	2.12E-02
18	2	1	50	2.09E-02
19	1.5	2	40	2.01E-02
20	2	1	40	1.99E-02
21	2	1	60	1.93E-02
22	1.5	2	60	1.87E-02
23	1.5	2	50	1.79E-02
24	1.5	2	70	1.79E-02
25	3	0	70	1.48E-02
26	3	0	40	1.42E-02
27	3	0	60	1.39E-02
28	3	0	50	1.25E-02
29	2.5	1	60	1.19E-02
30	2.5	1	70	1.15E-02
31	1.5	3	70	1.09E-02
32	2.5	1	50	1.08E-02
33	2.5	1	40	1.07E-02
34	2	2	40	1.05E-02
35	2	2	70	1.03E-02
36	2	2	50	1.01E-02
37	1.5	3	40	9.98E-03
38	1.5	3	50	9.91E-03
39	2	2	60	9.73E-03
40	1.5	3	60	9.31E-03
41	3	1	40	7.73E-03
42	2	3	40	7.14E-03
43	2.5	2	60	6.89E-03
44	3	1	50	6.72E-03
45	2.5	2	70	6.69E-03

K	*EL*	*WT*	*Sch*	*Prob.*
46	3	1	60	6.56E-03
47	2.5	2	50	6.35E-03
48	2.5	2	40	6.33E-03
49	1.5	4	50	6.30E-03
50	2	3	60	6.26E-03
51	3	1	70	5.83E-03
52	1.5	4	40	5.77E-03
53	2	3	70	5.49E-03
54	1.5	4	70	5.48E-03
55	2	3	50	5.32E-03
56	1.5	4	60	5.11E-03
57	3	2	70	3.83E-03
58	2.5	3	70	3.79E-03
59	2.5	3	50	3.74E-03
60	2.5	3	40	3.59E-03
61	3	2	50	3.46E-03
62	3	2	60	3.46E-03
63	2	4	70	3.42E-03
64	2	4	60	3.41E-03
65	2	4	40	3.27E-03
66	3	2	40	3.17E-03
67	2	4	50	2.84E-03
68	2.5	3	60	2.59E-03
69	3	3	70	2.11E-03
70	2.5	4	50	1.97E-03
71	2.5	4	60	1.89E-03
72	2.5	4	70	1.72E-03
73	3	3	50	1.64E-03
74	3	3	60	1.61E-03
75	3	3	40	1.47E-03
76	2.5	4	40	1.14E-03
77	3	4	50	9.61E-04
78	3	4	60	7.25E-04
79	3	4	40	6.64E-04
80	3	4	70	5.50E-04
81	3.5	2	40	4.94E-04
82	3.5	2	60	4.73E-04
83	3.5	2	70	2.65E-04
84	3.5	2	50	1.78E-04
85	3.5	3	70	1.58E-04
86	3.5	3	60	1.38E-04
87	3.5	3	50	4.73E-05
88	3.5	4	60	2.94E-05
89	3.5	4	50	2.44E-05
90	3.5	4	40	1.22E-05

A3. Python Script. Breitenhagen Levee Failure. Mohr-Coulomb

```
# -*- coding: utf-8 -*-
"""
Created on Wed Sep 16 11:53:36 2020

@author: wmgarciaf

This code is the second version of python execution od D-stability console
The idea is to use a better code with user-defined functions
Input variables have prior uniform distributions
This .py file extracts coordinates X and Z and the Radius of slip circle
"""
# Call the modules used in this code
import pandas as pd
import numpy as np
import random
#import matplotlib.pyplot as plt
import zipfile  # Import library to use zipfiles in Python
import subprocess # to run the D-stability console
import re  #Import Regex library to read float numbers
#import time
from tqdm import tqdm
#import gc

# Función para crear el data frame
def data_frame(n_sim):
    c_values     = np.arange(0.0, 16.0, 1.0);
    c_values     = ["%.1f" % member for member in c_values] #Format to c_data values as float
    fi_values    = np.arange(15.0 ,35.0, 1.0)
    fi_values    = ["%.1f" % member for member in fi_values]
    gamma_values = np.arange(18.0 ,18.1, 1.0)
    gamma_values = ["%.1f" % member for member in gamma_values]
    ST_values    = ['High', 'Medium', 'Low']
    CL_values    = ['Yes', 'No']
    PC_values    = ['High', 'Low']

    c_sample     = random.choices(c_values,     weights=None, k= n_sim)
    fi_sample    = random.choices(fi_values,    weights=None, k= n_sim)
    gamma_sample = random.choices(gamma_values, weights=None, k= n_sim)
    ST_sample    = random.choices(ST_values,    weights=None, k= n_sim)
    CL_sample    = random.choices(CL_values,    weights=None, k= n_sim)
    PC_sample    = random.choices(PC_values,    weights=None, k= n_sim)

    df = pd.DataFrame(list(zip(c_sample, fi_sample, gamma_sample,
                               ST_sample, CL_sample, PC_sample)),
                      columns =['c', 'fi', 'gamma', 'ST', 'CL', 'PC'])
    df['FoS']= np.repeat('NA', len(c_sample))  # add FoS column to dataframe an fill it with NA's
    df['X_center'] = np.repeat('NA', len(c_sample))  # add X_center column to dataframe an fill it with NA's
    df['Z_center'] = np.repeat('NA', len(c_sample))  # add Z_center column to dataframe an fill it with NA's
    df['R_center'] = np.repeat('NA', len(c_sample))  # add R_center column to dataframe an fill it with NA's
    return(df)

# Function to replace values in *soils.json file (c', fi', gamma)
def soils (iter):  # definition and name of the function
    f = open('soils.json', 'rt')   # open soils.json as read
    dat = f.read()  # put the information in the variable dat
    dat = dat.splitlines()   # split lines to convert in a list
    dat[7] = """      "VolumetricWeightAbovePhreaticLevel": """ + df.iloc[iter,2]+"," # replace line 7 for a value of c'
    dat[8] = """      "VolumetricWeightBelowPhreaticLevel": """ + df.iloc[iter,2]+","
    dat[11]= """      "Cohesion": """ + df.iloc[iter,0] + ","
    dat[17]= """      "FrictionAngle": """ + df.iloc[iter,1] + ","
    dat = '\n'.join(''.join(sub) for sub in dat)  # convert dat again to a string separated by new lines
    f.close() # close f
    f = open('soils.json', "wt") # open soils.json as write
    f.write(dat)  # rewrite f(soils.json) with values in dat
    f.close() # close f (soils.json)

# Function to replace values in waternets.json file (saturated condition, conductive layer, pond connection)
def waternets (iter):
    g = open('waternets.json', 'rt')
    data = g.read()
    data = data.splitlines()
    # Saturated levee condition
    if df.iloc[iter,3] == "High":  # Saturation condition at column 3 in df
```

A3. Python Script. Breitenhagen Levee Failure. Mohr-Coulomb

```
        data[19] = """          "X": 34.4,"""
        data[20] = """          "Z": 54.7"""
    elif df.iloc[iter,3] == "Medium":
        data[19] = """          "X": 39.2,"""
        data[20] = """          "Z": 54.7"""
    else:
        data[19] = """          "X": 38.0,"""
        data[20] = """          "Z": 51.5"""
    # Precence of conductive layer
    if df.iloc[iter,4] == "Yes": # conductive layer condition at column 4
        data[65] = """      "TopHeadLineId": "20","""
        data[66] = """      "BottomHeadLineId": "19","""
    else:
        data[65] = """      "TopHeadLineId": null,"""
        data[66] = """      "BottomHeadLineId": null,"""
    # Pond connection
    if df.iloc[iter,5] == "High": # Pond connection condition at column 5 in df
        data[82] = """      "TopHeadLineId": "21","""
        data[83] = """      "BottomHeadLineId": "21","""
    else:
        data[82] = """      "TopHeadLineId": null,"""
        data[83] = """      "BottomHeadLineId": null,"""
    data = '\n'.join(''.join(sub) for sub in data)  # convert dat again to a string separated by new lines
    g.close()
    g = open('waternets.json', "wt")
    g.write(data)
    g.close()

# Function to built a *stix file with the new soils.json and waternets.json files
def new_stix ():
    # Lines to delete 'soils.json' and 'waternets/waternets.json'
    zin  = zipfile.ZipFile ('Levee_Test_2_seed.stix', 'r') # variable containg with the original .stix file
    zout = zipfile.ZipFile ('Levee_Test_2_new.stix', 'w') # variable with the new .stix file
    for item in zin.infolist():  # for loop to read files
        buffer = zin.read(item.filename) #Read each file in zin
        if (item.filename[-100:] != 'soils.json' and
            item.filename[-100:] != 'waternets/waternets.json'): #Conditional to exclude soils.json
            zout.writestr(item, buffer) # Include files in zout
    zout.close()
    zin.close()
    # Lines to append 'soils.json' and 'waternets/waternets.json'
    z = zipfile.ZipFile('Levee_Test_2_new.stix', 'a') # variable to append into Test.stix
    z.write('soils.json')
    z.write('waternets.json', 'waternets/waternets.json') #file to append
    z.close()

# Function to extrac values of FoS, x_entry and x_exit
def extract (iter):
    zip_file = zipfile.ZipFile('Levee_Test_2_new.stix')
    results = zip_file.read('results/bishopbruteforce/bishopbruteforceresult.json')
    text_results = results.decode('utf-8')
    df.iloc[iter,6] = float(re.findall('-?\d+\.\d+', text_results.splitlines()[9])[0]) # get the Factor of Safety
    df.iloc[iter,7] = float(re.findall('-?\d+\.\d+', text_results.splitlines()[3])[0]) # get X_center
    df.iloc[iter,8] = float(re.findall('-?\d+\.\d+', text_results.splitlines()[4])[0]) # get Z_center
    df.iloc[iter,9] = float(re.findall('-?\d+\.\d+', text_results.splitlines()[6])[0]) # get R_center

#
n = 100000 # number of simulations
df = data_frame(n)

for ii in tqdm(range(0,n)): #tqdm add a progress bar
    soils(ii)
    waternets(ii)
    new_stix()
    subprocess.run(["D:/Software/Deltares/bin/D-GEO_Console.exe",
                    "D:/Simulations/Levee_V1/Levee_Test_2_new.stix"])
   # time.sleep(0.01)
    extract(ii)
    #gc.collect()

#df.to_csv("D:/Simulations/Levee_V1/100mil_MORH.csv")
```

A4. Python Script. Breitenhagen Levee Failure. SHANSEP

```
# -*- coding: utf-8 -*-
"""
Created on Wed Sep 16 11:53:36 2020

@author: wmgarciaf

This code uses the SHANSEP method in the D-stability console
The idea is to use a better code with user-defined functions
Input variables have prior uniform distributions
This .py file extracts coordinates X and Z and the Radius of slip circle
"""
# Call the modules used in this code
import pandas as pd
import numpy as np
import random
#import matplotlib.pyplot as plt
import zipfile  # Import library to use zipfiles in Python
import subprocess # to run the D-stability console
import re  #Import Regex library to read float numbers
#import time
from tqdm import tqdm
#import gc

# Función para crear el data frame
def data_frame(n_sim):
   S_values    = np.arange(0.23, 0.505, 0.02);
   S_values    = ["%.2f" % member for member in S_values] #Format to c_data values as float
   m_values    = np.arange(0.50 ,1.01, 0.04)
   m_values    = ["%.2f" % member for member in m_values]
   POP_values = np.arange(0.0 ,151, 10.0)
   POP_values = ["%.1f" % member for member in POP_values]
   ST_values   = ['High', 'Medium', 'Low']
   CL_values   = ['Yes', 'No']
   PC_values   = ['High', 'Low']
   S_sample    = random.choices(S_values,    weights=None, k= n_sim)
   m_sample    = random.choices(m_values,    weights=None, k= n_sim)
   POP_sample  = random.choices(POP_values, weights=None, k= n_sim)
   ST_sample   = random.choices(ST_values,   weights=None, k= n_sim)
   CL_sample   = random.choices(CL_values,   weights=None, k= n_sim)
   PC_sample   = random.choices(PC_values,   weights=None, k= n_sim)

   #plt.hist(PC_sample, bins=100)

   df = pd.DataFrame(list(zip(S_sample, m_sample, POP_sample,
                       ST_sample, CL_sample, PC_sample)),
                columns =['S', 'm', 'POP', 'ST', 'CL', 'PC'])
   df['FoS']= np.repeat('NA', len(S_sample))  # add FoS column to dataframe an fill it with NA's
   df['X_center'] = np.repeat('NA', len(S_sample))  # add X_center column to dataframe an fill it with NA's
   df['Z_center'] = np.repeat('NA', len(S_sample))  # add Z_center column to dataframe an fill it with NA's
   df['R_center'] = np.repeat('NA', len(S_sample))  # add R_center column to dataframe an fill it with NA's
   return(df)

# Function to replace values in *soils.json file (S, m)
def soils (iter):  # definition and name of the function
   f = open('soils.json', 'rt')   # open soils.json as read
   dat = f.read()  # put the information in the variable dat
   dat = dat.splitlines()   # split lines to convert in a list

   dat[30]= """      "ShearStrengthRatio": """ + df.iloc[iter,0] + ","
   dat[36]= """      "StrengthIncreaseExponent": """ + df.iloc[iter,1] + ","

   dat = '\n'.join(''.join(sub) for sub in dat)  # convert dat again to a string separated by new lines
   f.close() # close f
   f = open('soils.json', "wt") # open soils.json as write
   f.write(dat)  # rewrite f(soils.json) with values in dat
   f.close() # close f (soils.json)

# Function to replace values in *states.json file (POP)
def states(iter):
  h = open('states.json', 'rt')
  da = h.read()
  da = da.splitlines()
```

A4. Python Script. Breitenhagen Levee Failure. SHANSEP

```
    da[14]= """        "Pop": """ + df.iloc[iter,2] + ","
    da = '\n'.join(''.join(sub) for sub in da)  # convert dat again to a string separated by new lines
    h.close() # close f
    h = open('states.json', "wt") # open soils.json as write
    h.write(da)  # rewrite f(soils.json) with values in dat
    h.close() # close f (soils.json)

# Function to replace values in waternets.json file (saturated condition, conductive layer, pond connection)
def waternets (iter):
    g = open('waternets.json', 'rt')
    data = g.read()
    data = data.splitlines()
    # Saturated levee condition
    if df.iloc[iter,3] == "High":  # Saturation condition at column 3 in df
        data[19] = """          "X": 34.4,"""
        data[20] = """          "Z": 54.7"""
    elif df.iloc[iter,3] == "Medium":
        data[19] = """          "X": 39.2,"""
        data[20] = """          "Z": 54.7"""
    else:
        data[19] = """          "X": 38.0,"""
        data[20] = """          "Z": 51.5"""
    # Precence of conductive layer
    if df.iloc[iter,4] == "Yes": # conductive layer condition at column 4
        data[65] = """      "TopHeadLineId": "20","""
        data[66] = """      "BottomHeadLineId": "19","""
    else:
        data[65] = """      "TopHeadLineId": null,"""
        data[66] = """      "BottomHeadLineId": null,"""
    # Pond connection
    if df.iloc[iter,5] == "High": # Pond connection condition at column 5 in df
        data[82] = """      "TopHeadLineId": "21","""
        data[83] = """      "BottomHeadLineId": "21","""
    else:
        data[82] = """      "TopHeadLineId": null,"""
        data[83] = """      "BottomHeadLineId": null,"""
    data = '\n'.join(''.join(sub) for sub in data)  # convert dat again to a string separated by new lines
    g.close()
    g = open('waternets.json', "wt")
    g.write(data)
    g.close()

# Function to built a *stix file with the new soils.json, waternets.json and states.json files
def new_stix ():
    # Lines to delete 'soils.json' and 'waternets/waternets.json'
    zin  = zipfile.ZipFile ('Levee_Test_3_seed.stix', 'r') # variable containg with the original .stix file
    zout = zipfile.ZipFile ('Levee_Test_3_new.stix', 'w') # variable with the new .stix file
    for item in zin.infolist():  # for loop to read files
        buffer = zin.read(item.filename) #Read each file in zin
        if (item.filename[-100:] != 'soils.json' and
            item.filename[-100:] != 'waternets/waternets.json' and
            item.filename[-100:] != 'states/states.json'): #Conditional to exclude soils.json, waternets and states.json
            zout.writestr(item, buffer) # Include files in zout
    zout.close()
    zin.close()
    # Lines to append 'soils.json' ,'waternets/waternets.json' and 'states.json'
    z = zipfile.ZipFile('Levee_Test_3_new.stix', 'a') # variable to append into Test.stix
    z.write('soils.json')
    z.write('waternets.json', 'waternets/waternets.json') #file to append
    z.write('states.json', 'states/states.json')
    z.close()

# Function to extrac values of FoS, x_entry and x_exit
def extract (iter):
    zip_file = zipfile.ZipFile('Levee_Test_3_new.stix')
    results = zip_file.read('results/bishopbruteforce/bishopbruteforceresult.json')
    text_results = results.decode('utf-8')
    df.iloc[iter,6] = float(re.findall('-?\d+\.\d+', text_results.splitlines()[9])[0]) # get the Factor of Safety
    df.iloc[iter,7] = float(re.findall('-?\d+\.\d+', text_results.splitlines()[3])[0]) # get X_center
    df.iloc[iter,8] = float(re.findall('-?\d+\.\d+', text_results.splitlines()[4])[0]) # get Z_center
    df.iloc[iter,9] = float(re.findall('-?\d+\.\d+', text_results.splitlines()[6])[0]) # get R_center

#
```

A4. Python Script. Breitenhagen Levee Failure. SHANSEP

```
n = 100000# number of simulations
df = data_frame(n)

for ii in tqdm(range(0,n)): #tqdm add a progress bar
   soils(ii)
   waternets(ii)
   states(ii)
   new_stix()
   subprocess.run(["D:/Software/Deltares/bin/D-GEO_Console.exe",
            "D:/Simulations/Levee_V2/Levee_Test_3_new.stix"], check= False)
  # time.sleep(0.01)
   extract(ii)
   #gc.collect()
#df.to_csv("D:/Simulations/Levee_V2/100mil_SHANSEP.csv")
```

A5. K Most Probable Explanations (K MPE) of the Levee Failure

MOHR - COULOMB						
K	ST	CL	PC	c	phi	Prob
1	Medium	Yes	Low	4	17	4.93E-03
2	High	No	Low	1	24	4.79E-03
3	Medium	Yes	Low	3	18	4.77E-03
4	Medium	No	Low	1	18	4.76E-03
5	High	No	Low	1	29	4.73E-03
6	Medium	Yes	Low	3	32	4.69E-03
7	High	Yes	Low	2	32	4.68E-03
8	High	Yes	High	4	23	4.66E-03
9	High	Yes	Low	6	23	4.65E-03
10	High	No	Low	3	15	4.62E-03
11	High	Yes	High	6	31	4.62E-03
12	High	Yes	Low	4	21	4.59E-03
13	Medium	Yes	Low	3	16	4.58E-03
14	High	Yes	Low	6	21	4.57E-03
15	Medium	No	Low	1	16	4.57E-03
16	Medium	Yes	Low	3	19	4.55E-03
17	High	Yes	High	4	17	4.55E-03
18	Medium	No	Low	2	18	4.55E-03
19	High	Yes	Low	6	20	4.54E-03
20	High	No	Low	3	26	4.54E-03
21	High	Yes	Low	4	22	4.52E-03
22	High	No	Low	4	15	4.51E-03
23	High	Yes	High	4	18	4.51E-03
24	High	Yes	High	5	24	4.50E-03
25	High	Yes	High	3	18	4.49E-03
26	High	No	Low	1	20	4.49E-03
27	High	No	Low	1	34	4.48E-03
28	High	Yes	Low	5	23	4.47E-03
29	High	Yes	High	1	15	4.46E-03
30	High	Yes	Low	6	19	4.46E-03
31	High	Yes	Low	5	18	4.46E-03
32	High	Yes	High	5	18	4.45E-03
33	High	No	Low	2	21	4.45E-03
34	High	Yes	Low	2	29	4.45E-03
35	High	No	Low	1	30	4.45E-03
36	High	Yes	High	5	25	4.44E-03
37	High	No	Low	2	34	4.44E-03

SHANSEP							
K	ST	CL	PC	S	m	POP	Prob
1	High	Yes	High	0.31	0.78	10	1.51E-03
2	High	Yes	Low	0.33	0.74	20	1.49E-03
3	High	No	Low	0.27	0.54	20	1.46E-03
4	High	Yes	Low	0.39	0.54	0	1.44E-03
5	High	No	Low	0.37	0.94	0	1.40E-03
6	High	Yes	High	0.29	0.82	10	1.37E-03
7	High	No	Low	0.33	0.94	0	1.36E-03
8	Medium	Yes	Low	0.27	0.74	10	1.35E-03
9	High	No	Low	0.43	0.9	0	1.34E-03
10	High	Yes	High	0.27	0.86	0	1.34E-03
11	High	No	Low	0.27	0.94	0	1.34E-03
12	High	Yes	Low	0.23	0.7	50	1.33E-03
13	High	Yes	Low	0.49	0.54	0	1.32E-03
14	High	No	Low	0.33	0.9	0	1.32E-03
15	High	Yes	High	0.37	0.86	10	1.31E-03
16	High	Yes	High	0.49	0.62	0	1.31E-03
17	High	Yes	Low	0.49	0.7	0	1.31E-03
18	High	Yes	Low	0.45	0.7	0	1.31E-03
19	Medium	Yes	Low	0.27	0.54	20	1.30E-03
20	High	Yes	Low	0.25	0.58	80	1.30E-03
21	High	No	Low	0.31	0.86	10	1.29E-03
22	High	Yes	High	0.29	0.54	0	1.29E-03
23	High	Yes	High	0.45	0.58	0	1.28E-03
24	High	Yes	High	0.35	0.9	0	1.28E-03
25	High	No	Low	0.43	0.78	0	1.28E-03
26	High	Yes	Low	0.35	0.94	10	1.28E-03
27	High	Yes	High	0.37	0.78	10	1.27E-03
28	High	Yes	Low	0.41	0.74	0	1.26E-03
29	High	Yes	Low	0.25	0.94	0	1.26E-03
30	High	Yes	High	0.23	0.9	0	1.26E-03
31	High	Yes	High	0.25	0.82	0	1.26E-03
32	High	Yes	High	0.25	0.78	10	1.25E-03
33	High	No	Low	0.29	0.58	0	1.25E-03
34	High	Yes	High	0.39	0.74	0	1.25E-03
35	High	No	Low	0.41	0.58	0	1.25E-03
36	High	Yes	Low	0.23	0.62	70	1.25E-03
37	High	Yes	High	0.25	0.7	20	1.25E-03

A5. K Most Probable Explanations (K MPE) of the Levee Failure

MOHR - COULOMB						
K	ST	CL	PC	c	phi	Prob
38	High	Yes	High	2	30	4.44E-03
39	High	No	Low	2	30	4.43E-03
40	High	No	Low	3	32	4.43E-03
41	High	No	Low	1	16	4.42E-03
42	High	Yes	Low	6	25	4.42E-03
43	Medium	Yes	Low	2	21	4.42E-03
44	High	Yes	Low	6	24	4.42E-03
45	Medium	Yes	Low	4	16	4.41E-03
46	Medium	Yes	Low	4	23	4.41E-03
47	High	No	Low	1	27	4.40E-03
48	High	Yes	Low	1	15	4.39E-03
49	Medium	Yes	Low	2	24	4.39E-03
50	High	Yes	High	1	17	4.39E-03
51	High	No	Low	4	17	4.38E-03
52	Medium	Yes	Low	3	21	4.38E-03
53	High	No	Low	2	26	4.37E-03
54	High	No	Low	3	33	4.36E-03
55	High	No	Low	4	19	4.36E-03
56	Medium	No	Low	1	17	4.35E-03
57	High	Yes	High	2	32	4.35E-03
58	High	No	Low	1	26	4.34E-03
59	High	No	Low	2	27	4.34E-03
60	High	No	Low	3	25	4.33E-03
61	Medium	Yes	Low	2	16	4.33E-03
62	High	Yes	Low	5	22	4.33E-03
63	High	No	Low	1	32	4.32E-03
64	High	Yes	Low	6	34	4.31E-03
65	High	Yes	High	6	25	4.31E-03
66	Medium	No	Low	1	19	4.31E-03
67	Medium	No	Low	2	21	4.31E-03
68	High	No	Low	1	23	4.31E-03
69	High	No	Low	2	33	4.30E-03
70	High	No	Low	1	19	4.30E-03
71	Medium	Yes	Low	3	30	4.29E-03
72	High	Yes	Low	4	16	4.29E-03
73	High	No	Low	2	31	4.28E-03

SHANSEP							
K	ST	CL	PC	S	m	POP	Prob
38	High	No	Low	0.45	0.5	10	1.25E-03
39	High	Yes	Low	0.27	0.7	20	1.25E-03
40	Medium	Yes	Low	0.31	0.78	10	1.24E-03
41	Medium	Yes	Low	0.31	0.5	20	1.24E-03
42	High	Yes	Low	0.23	0.9	30	1.23E-03
43	High	Yes	Low	0.23	0.74	30	1.23E-03
44	High	No	Low	0.23	0.5	40	1.22E-03
45	High	Yes	Low	0.23	0.98	20	1.22E-03
46	High	Yes	High	0.31	0.78	0	1.22E-03
47	Medium	Yes	Low	0.31	0.9	10	1.22E-03
48	Medium	Yes	Low	0.31	0.58	10	1.22E-03
49	High	Yes	Low	0.41	0.5	0	1.22E-03
50	High	No	Low	0.31	0.54	0	1.22E-03
51	High	No	Low	0.33	0.54	20	1.22E-03
52	High	Yes	Low	0.41	0.94	0	1.22E-03
53	High	Yes	High	0.41	0.82	10	1.21E-03
54	High	No	Low	0.49	0.58	0	1.21E-03
55	Medium	Yes	Low	0.23	0.9	10	1.21E-03
56	High	No	Low	0.41	0.78	0	1.21E-03
57	High	Yes	Low	0.45	0.86	0	1.21E-03
58	High	No	Low	0.25	0.5	0	1.21E-03
59	High	Yes	High	0.39	0.94	0	1.21E-03
60	High	Yes	High	0.41	0.78	10	1.20E-03
61	High	No	Low	0.43	0.7	0	1.19E-03
62	Medium	Yes	Low	0.35	0.74	10	1.19E-03
63	High	Yes	High	0.45	0.86	10	1.19E-03
64	High	No	Low	0.31	0.82	10	1.19E-03
65	Low	No	Low	0.23	0.66	10	1.19E-03
66	High	Yes	High	0.47	0.54	0	1.19E-03
67	High	Yes	Low	0.43	0.7	0	1.19E-03
68	High	No	Low	0.47	0.7	0	1.19E-03
69	High	Yes	High	0.25	0.78	20	1.19E-03
70	High	Yes	Low	0.23	0.66	60	1.18E-03
71	High	No	Low	0.35	0.82	10	1.18E-03
72	High	No	Low	0.31	0.58	10	1.18E-03
73	High	Yes	High	0.43	0.7	0	1.18E-03

A5. K Most Probable Explanations (K MPE) of the Levee Failure

MOHR - COULOMB						
K	ST	CL	PC	c	phi	Prob
74	High	Yes	High	6	32	4.28E-03
75	High	No	Low	1	31	4.28E-03
76	Medium	Yes	Low	3	29	4.27E-03
77	High	Yes	Low	6	30	4.27E-03
78	Low	No	Low	1	15	4.27E-03
79	High	Yes	High	4	15	4.27E-03
80	High	No	Low	2	20	4.27E-03
81	High	No	Low	3	18	4.27E-03
82	High	No	Low	5	15	4.26E-03
83	High	No	Low	2	24	4.25E-03
84	Medium	Yes	Low	3	26	4.25E-03
85	High	No	Low	1	28	4.25E-03
86	High	Yes	High	1	16	4.25E-03
87	High	No	Low	1	33	4.24E-03
88	High	Yes	High	4	22	4.24E-03
89	High	No	Low	3	24	4.24E-03
90	High	Yes	Low	4	23	4.24E-03
91	High	No	Low	1	25	4.23E-03
92	High	Yes	Low	5	24	4.23E-03
93	High	No	Low	2	23	4.23E-03
94	High	Yes	Low	4	17	4.23E-03
95	High	Yes	Low	5	26	4.22E-03
96	Medium	No	Low	1	15	4.22E-03
97	High	Yes	High	5	17	4.22E-03
98	Medium	Yes	Low	5	16	4.21E-03
99	Medium	Yes	Low	2	19	4.21E-03
100	Medium	Yes	Low	2	25	4.21E-03
101	Medium	Yes	Low	3	24	4.21E-03
102	High	Yes	Low	5	20	4.20E-03
103	High	Yes	Low	4	18	4.20E-03
104	Medium	Yes	Low	4	18	4.20E-03
105	High	Yes	Low	6	18	4.20E-03
106	High	Yes	High	6	24	4.19E-03
107	High	No	Low	3	23	4.19E-03
108	Medium	Yes	Low	3	28	4.19E-03
109	Medium	Yes	Low	2	23	4.19E-03
110	Medium	Yes	Low	3	22	4.19E-03

SHANSEP							
K	ST	CL	PC	S	m	POP	Prob
74	High	Yes	High	0.29	0.78	0	1.18E-03
75	High	Yes	High	0.45	0.9	10	1.18E-03
76	Medium	Yes	Low	0.31	0.54	20	1.18E-03
77	High	No	Low	0.33	0.86	10	1.18E-03
78	High	Yes	Low	0.23	0.58	90	1.18E-03
79	High	No	Low	0.23	0.94	0	1.18E-03
80	High	Yes	High	0.29	0.5	0	1.17E-03
81	High	Yes	High	0.45	0.82	10	1.17E-03
82	High	Yes	Low	0.25	0.66	30	1.17E-03
83	High	Yes	Low	0.43	0.82	0	1.16E-03
84	High	Yes	Low	0.49	0.82	10	1.16E-03
85	High	No	Low	0.49	0.62	0	1.16E-03
86	High	No	Low	0.29	0.62	10	1.16E-03
87	High	Yes	Low	0.25	0.62	60	1.16E-03
88	High	Yes	High	0.39	0.82	0	1.16E-03
89	High	No	Low	0.33	0.54	0	1.16E-03
90	High	No	Low	0.31	0.98	0	1.16E-03
91	High	Yes	Low	0.33	0.54	0	1.16E-03
92	High	Yes	Low	0.47	0.78	0	1.16E-03
93	High	Yes	High	0.47	0.86	0	1.15E-03
94	High	Yes	High	0.43	0.98	0	1.15E-03
95	High	No	Low	0.31	0.66	20	1.15E-03
96	High	No	Low	0.33	0.5	0	1.15E-03
97	High	Yes	High	0.39	0.66	0	1.15E-03
98	High	Yes	Low	0.35	0.9	10	1.15E-03
99	High	No	Low	0.25	0.54	0	1.15E-03
100	High	Yes	High	0.31	0.94	10	1.15E-03
101	High	Yes	High	0.47	0.5	0	1.15E-03
102	High	Yes	High	0.37	0.78	0	1.15E-03
103	High	Yes	Low	0.33	0.74	0	1.15E-03
104	Medium	Yes	Low	0.23	0.5	20	1.15E-03
105	High	Yes	Low	0.39	0.7	0	1.15E-03
106	High	Yes	Low	0.27	0.78	20	1.14E-03
107	High	No	Low	0.25	0.54	30	1.13E-03
108	High	Yes	High	0.33	0.82	10	1.13E-03
109	High	Yes	High	0.43	0.54	0	1.13E-03
110	High	No	Low	0.37	0.82	10	1.13E-03

A5. K Most Probable Explanations (K MPE) of the Levee Failure

MOHR - COULOMB						
K	ST	CL	PC	c	phi	Prob
111	Medium	Yes	Low	2	22	4.19E-03
112	High	Yes	Low	6	26	4.18E-03
113	High	Yes	Low	5	15	4.18E-03
114	High	Yes	Low	1	16	4.18E-03
115	Medium	Yes	Low	2	18	4.18E-03
116	Medium	Yes	Low	2	17	4.18E-03
117	High	Yes	Low	6	17	4.18E-03
118	High	Yes	Low	6	33	4.18E-03
119	High	Yes	High	4	16	4.17E-03
120	Medium	Yes	Low	3	20	4.17E-03
121	High	Yes	High	5	28	4.17E-03
122	High	Yes	Low	6	32	4.16E-03
123	High	Yes	Low	5	16	4.16E-03
124	Medium	Yes	Low	2	15	4.16E-03
125	High	No	Low	3	22	4.15E-03
126	High	No	Low	2	29	4.15E-03
127	High	No	Low	2	18	4.15E-03
128	High	Yes	Low	5	30	4.15E-03
129	High	Yes	High	6	28	4.15E-03
130	High	Yes	High	6	33	4.15E-03
131	High	No	Low	2	15	4.14E-03
132	High	Yes	Low	5	17	4.14E-03
133	High	Yes	Low	4	24	4.14E-03
134	High	Yes	Low	5	19	4.14E-03
135	Medium	Yes	Low	3	23	4.13E-03
136	High	No	Low	3	28	4.13E-03
137	High	No	Low	5	16	4.13E-03
138	High	Yes	Low	6	15	4.13E-03
139	High	No	Low	1	22	4.12E-03
140	High	Yes	High	6	26	4.11E-03
141	High	Yes	High	4	21	4.11E-03
142	High	No	Low	4	16	4.11E-03
143	Medium	Yes	Low	3	27	4.11E-03
144	High	Yes	High	5	22	4.11E-03
145	High	No	Low	3	31	4.11E-03
146	High	No	Low	4	22	4.11E-03

SHANSEP							
K	ST	CL	PC	S	m	POP	Prob
111	High	No	Low	0.47	0.66	0	1.13E-03
112	High	Yes	High	0.25	0.7	0	1.13E-03
113	High	Yes	High	0.33	0.9	0	1.13E-03
114	High	Yes	Low	0.49	0.78	0	1.13E-03
115	High	Yes	Low	0.25	0.74	30	1.13E-03
116	High	No	Low	0.49	0.86	0	1.12E-03
117	High	Yes	Low	0.23	0.62	60	1.12E-03
118	High	Yes	High	0.47	0.78	10	1.12E-03
119	High	No	Low	0.37	0.78	10	1.12E-03
120	High	Yes	High	0.31	0.7	20	1.12E-03
121	High	Yes	High	0.33	0.74	20	1.12E-03
122	High	Yes	High	0.33	0.54	0	1.12E-03
123	High	Yes	Low	0.33	0.7	30	1.12E-03
124	High	Yes	Low	0.25	0.54	0	1.12E-03
125	High	No	Low	0.27	0.7	0	1.12E-03
126	High	No	Low	0.23	0.78	10	1.12E-03
127	High	Yes	High	0.49	0.58	0	1.12E-03
128	High	Yes	Low	0.37	0.74	20	1.12E-03
129	High	Yes	Low	0.27	0.66	40	1.12E-03
130	High	Yes	Low	0.35	0.74	20	1.12E-03
131	Medium	Yes	Low	0.33	0.62	10	1.12E-03
132	High	No	Low	0.39	0.94	0	1.12E-03
133	Medium	Yes	Low	0.23	0.7	10	1.12E-03
134	Medium	Yes	Low	0.23	0.66	10	1.12E-03
135	High	Yes	Low	0.49	0.74	0	1.12E-03
136	High	Yes	High	0.37	0.5	0	1.12E-03
137	High	Yes	High	0.49	0.86	0	1.12E-03
138	High	No	Low	0.25	0.62	20	1.12E-03
139	High	Yes	Low	0.43	0.82	10	1.12E-03
140	High	No	Low	0.29	0.86	0	1.12E-03
141	Medium	Yes	Low	0.29	0.58	10	1.12E-03
142	High	Yes	Low	0.47	0.62	0	1.12E-03
143	High	Yes	Low	0.31	0.82	0	1.12E-03
144	High	Yes	High	0.23	0.82	10	1.12E-03
145	High	Yes	Low	0.43	0.5	0	1.11E-03
146	High	Yes	Low	0.23	0.78	20	1.11E-03

A5. K Most Probable Explanations (K MPE) of the Levee Failure

MOHR - COULOMB						
K	ST	CL	PC	c	phi	Prob
147	High	No	Low	4	23	4.09E-03
148	High	Yes	Low	3	19	4.09E-03
149	High	Yes	High	3	15	4.08E-03
150	High	No	Low	3	20	4.08E-03
151	High	No	Low	1	15	4.07E-03
152	Medium	Yes	Low	4	19	4.07E-03
153	High	No	Low	3	19	4.07E-03
154	High	Yes	Low	2	30	4.07E-03
155	High	Yes	High	6	22	4.07E-03
156	High	Yes	High	2	29	4.07E-03
157	High	Yes	Low	6	16	4.05E-03
158	High	Yes	High	3	16	4.05E-03
159	Medium	No	Low	2	15	4.04E-03
160	High	Yes	High	4	24	4.04E-03
161	High	Yes	Low	1	17	4.04E-03
162	High	Yes	High	5	29	4.04E-03
163	Medium	No	Low	3	17	4.03E-03
164	High	Yes	Low	3	16	4.03E-03
165	High	Yes	Low	6	22	4.03E-03
166	High	Yes	Low	6	31	4.03E-03
167	High	Yes	Low	5	21	4.03E-03
168	High	No	Low	1	17	4.03E-03
169	Medium	Yes	Low	3	25	4.03E-03
170	High	No	Low	2	17	4.02E-03
171	High	Yes	High	5	20	4.02E-03
172	High	Yes	High	6	23	4.02E-03
173	High	No	Low	4	21	4.02E-03
174	High	No	Low	5	18	4.02E-03
175	High	Yes	High	5	19	4.02E-03
176	High	No	Low	2	19	4.01E-03
177	Medium	Yes	Low	4	20	3.99E-03
178	Medium	Yes	Low	3	31	3.98E-03
179	Low	No	Low	1	17	3.98E-03
180	High	No	Low	1	21	3.97E-03
181	High	Yes	Low	2	27	3.97E-03
182	High	Yes	High	5	26	3.97E-03
183	High	Yes	High	4	20	3.96E-03

SHANSEP							
K	ST	CL	PC	S	m	POP	Prob
147	High	Yes	Low	0.43	0.78	0	1.11E-03
148	High	Yes	High	0.23	0.94	10	1.11E-03
149	High	No	Low	0.37	0.74	10	1.11E-03
150	High	Yes	High	0.31	0.62	40	1.11E-03
151	High	No	Low	0.25	0.54	20	1.11E-03
152	High	Yes	High	0.47	0.74	0	1.11E-03
153	High	No	Low	0.23	0.54	20	1.10E-03
154	High	Yes	Low	0.41	0.82	10	1.09E-03
155	High	Yes	Low	0.29	0.62	50	1.09E-03
156	High	No	Low	0.43	0.94	0	1.09E-03
157	High	Yes	High	0.43	0.78	0	1.09E-03
158	High	Yes	High	0.31	0.86	10	1.09E-03
159	High	No	Low	0.49	0.78	0	1.09E-03
160	High	Yes	Low	0.41	0.66	0	1.09E-03
161	High	Yes	High	0.49	0.78	10	1.09E-03
162	High	Yes	High	0.27	0.66	30	1.09E-03
163	High	Yes	Low	0.45	0.78	10	1.09E-03
164	High	No	Low	0.23	0.66	20	1.09E-03
165	High	No	Low	0.35	0.66	10	1.09E-03
166	High	No	Low	0.33	0.78	0	1.09E-03
167	High	Yes	Low	0.39	0.7	20	1.09E-03
168	High	No	Low	0.35	0.58	10	1.09E-03
169	High	Yes	Low	0.41	0.82	0	1.09E-03
170	High	Yes	High	0.43	0.82	0	1.09E-03
171	High	Yes	High	0.41	0.5	0	1.09E-03
172	High	Yes	Low	0.33	0.98	10	1.09E-03
173	High	No	Low	0.39	0.54	0	1.09E-03
174	High	No	Low	0.25	0.54	40	1.09E-03
175	High	No	Low	0.35	0.54	20	1.09E-03
176	High	Yes	High	0.23	0.58	60	1.08E-03
177	High	Yes	Low	0.35	0.82	0	1.08E-03
178	High	No	Low	0.35	0.5	20	1.08E-03
179	High	No	Low	0.37	0.5	0	1.08E-03
180	High	Yes	High	0.35	0.86	10	1.08E-03
181	High	Yes	High	0.33	0.98	0	1.08E-03
182	High	Yes	High	0.37	0.86	0	1.08E-03
183	High	Yes	Low	0.31	0.78	0	1.08E-03

A5. K Most Probable Explanations (K MPE) of the Levee Failure

MOHR - COULOMB						
K	ST	CL	PC	c	phi	Prob
184	High	Yes	High	5	21	3.95E-03
185	High	No	Low	3	17	3.95E-03
186	High	Yes	High	3	19	3.95E-03
187	Medium	Yes	Low	4	15	3.95E-03
188	High	Yes	High	6	34	3.95E-03
189	High	Yes	High	5	30	3.93E-03
190	High	Yes	High	3	17	3.93E-03
191	High	Yes	Low	4	19	3.93E-03
192	High	No	Low	3	30	3.92E-03
193	High	No	Low	2	32	3.92E-03
194	High	Yes	High	6	29	3.92E-03
195	High	Yes	High	6	27	3.91E-03
196	High	Yes	High	6	30	3.91E-03
197	High	Yes	Low	5	28	3.90E-03
198	Low	No	Low	1	16	3.89E-03
199	High	Yes	Low	4	15	3.89E-03
200	High	No	Low	3	21	3.89E-03
201	High	No	Low	2	22	3.88E-03
202	High	Yes	Low	3	15	3.88E-03
203	High	No	Low	2	28	3.88E-03
204	Medium	No	Low	2	16	3.88E-03
205	High	No	Low	4	24	3.86E-03
206	High	Yes	High	2	28	3.86E-03
207	High	No	Low	3	29	3.85E-03
208	Medium	Yes	Low	3	15	3.85E-03
209	High	No	Low	3	27	3.84E-03
210	Medium	Yes	Low	4	21	3.84E-03
211	High	No	Low	3	16	3.84E-03
212	High	No	Low	1	18	3.84E-03
213	High	No	Low	2	25	3.83E-03
214	High	Yes	Low	5	29	3.83E-03
215	High	Yes	Low	3	17	3.82E-03
216	Medium	Yes	Low	5	15	3.82E-03
217	High	Yes	Low	6	29	3.81E-03
218	High	Yes	Low	6	27	3.81E-03
219	High	Yes	Low	2	31	3.80E-03

SHANSEP							
K	ST	CL	PC	S	m	POP	Prob
184	Medium	Yes	Low	0.39	0.62	10	1.08E-03
185	High	Yes	High	0.23	0.58	70	1.08E-03
186	High	Yes	Low	0.23	0.78	0	1.08E-03
187	High	Yes	Low	0.35	0.94	0	1.08E-03
188	High	Yes	Low	0.29	0.94	10	1.08E-03
189	High	Yes	Low	0.47	0.7	0	1.07E-03
190	High	Yes	High	0.45	0.54	0	1.07E-03
191	High	No	Low	0.27	0.58	20	1.07E-03
192	High	Yes	Low	0.33	0.7	20	1.07E-03
193	High	No	Low	0.41	0.5	10	1.07E-03
194	High	Yes	High	0.43	0.78	10	1.06E-03
195	High	No	Low	0.23	0.74	10	1.06E-03
196	High	No	Low	0.33	0.9	10	1.06E-03
197	High	No	Low	0.25	0.74	10	1.06E-03
198	High	Yes	High	0.35	0.94	10	1.06E-03
199	High	No	Low	0.23	0.78	0	1.06E-03
200	High	No	Low	0.25	0.78	0	1.06E-03
201	High	No	Low	0.29	0.5	0	1.06E-03
202	High	No	Low	0.33	0.58	10	1.06E-03
203	High	Yes	High	0.47	0.82	10	1.06E-03
204	High	Yes	High	0.23	0.78	0	1.06E-03
205	High	Yes	Low	0.31	0.78	20	1.06E-03
206	High	Yes	High	0.33	0.78	10	1.06E-03
207	High	No	Low	0.33	0.74	10	1.06E-03
208	High	Yes	High	0.39	0.98	0	1.06E-03
209	High	Yes	High	0.41	0.98	0	1.06E-03
210	High	Yes	Low	0.37	0.98	10	1.06E-03
211	High	Yes	High	0.35	0.62	0	1.06E-03
212	High	No	Low	0.41	0.7	0	1.06E-03
213	Medium	Yes	Low	0.29	0.5	10	1.06E-03
214	Low	No	Low	0.23	0.58	10	1.06E-03
215	High	Yes	High	0.29	0.9	10	1.06E-03
216	High	No	Low	0.43	0.58	10	1.06E-03
217	High	No	Low	0.27	0.5	0	1.06E-03
218	Medium	Yes	Low	0.33	0.86	10	1.06E-03
219	High	No	Low	0.39	0.7	0	1.06E-03

A5. K Most Probable Explanations (K MPE) of the Levee Failure

MOHR - COULOMB						
K	**ST**	**CL**	**PC**	**c**	**phi**	**Prob**
220	High	Yes	Low	5	25	3.80E-03
221	Medium	No	Low	2	17	3.79E-03
222	Medium	No	Low	2	20	3.78E-03
223	High	No	Low	5	17	3.78E-03
224	Medium	Yes	Low	3	17	3.77E-03
225	High	No	Low	4	20	3.77E-03
226	High	Yes	Low	4	20	3.76E-03
227	Medium	Yes	Low	2	20	3.75E-03
228	Medium	No	Low	3	18	3.75E-03
229	High	Yes	High	2	27	3.72E-03
230	Medium	No	Low	2	19	3.70E-03
231	High	No	Low	2	16	3.69E-03
232	High	Yes	Low	3	18	3.68E-03
233	High	Yes	Low	5	27	3.64E-03
234	Medium	Yes	Low	4	22	3.64E-03
235	High	Yes	High	4	19	3.64E-03
236	High	No	Low	4	18	3.64E-03
237	High	Yes	High	2	31	3.62E-03
238	High	Yes	Low	6	28	3.58E-03
239	High	Yes	High	5	23	3.57E-03
240	High	Yes	High	5	27	3.40E-03
241	High	Yes	Low	2	28	3.40E-03

SHANSEP							
K	**ST**	**CL**	**PC**	**S**	**m**	**POP**	**Prob**
220	Medium	Yes	Low	0.37	0.74	10	1.05E-03
221	High	Yes	High	0.23	0.82	0	1.05E-03
222	High	Yes	High	0.29	0.82	0	1.05E-03
223	High	Yes	Low	0.35	0.54	0	1.05E-03
224	High	Yes	High	0.35	0.78	10	1.05E-03
225	High	Yes	Low	0.29	0.82	0	1.05E-03
226	High	Yes	Low	0.25	0.9	10	1.05E-03
227	High	Yes	High	0.29	0.62	40	1.05E-03
228	High	No	Low	0.35	0.9	10	1.05E-03
229	High	No	Low	0.35	0.9	0	1.05E-03
230	High	Yes	High	0.29	0.58	0	1.05E-03
231	High	Yes	Low	0.29	0.9	20	1.05E-03
232	High	Yes	High	0.37	0.58	0	1.05E-03
233	High	No	Low	0.35	0.62	10	1.05E-03
234	High	Yes	High	0.47	0.58	0	1.05E-03
235	High	No	Low	0.25	0.62	0	1.05E-03
236	High	Yes	Low	0.23	0.9	20	1.05E-03
237	Medium	Yes	Low	0.37	0.7	10	1.04E-03
238	High	No	Low	0.33	0.58	20	1.04E-03
239	High	Yes	High	0.39	0.54	0	1.04E-03
240	High	No	Low	0.49	0.82	0	1.04E-03
241	High	Yes	Low	0.45	0.78	0	1.04E-03

A6. Python Script. The Green Office Excavation Failure. Slope Stability Model

```
# -*- coding: utf-8 -*-
"""
Created on Wed Jun 22 17:53:42 2022

@author: mgfer
"""

import numpy as np
import random
from scipy.stats import truncnorm
#import matplotlib.pyplot as plt
import pandas as pd
import os        # To rename files
import shutil    # To copy files
import subprocess # to run the Slide console
from tqdm import tqdm

# Function to create random values and storage them in a dataframe
def data_frame(n_sim):
   random.seed(2022)
   # Parameters to create a truncate normal distribution for c' and phi'
   # Soil1: Soft Clay
   c1_mean = 5.; COV1 = 0.1; c1_std = c1_mean*COV1; c1_clip_a = c1_mean-3.*c1_std ; c1_clip_b = c1_mean+3.*c1_std
   c1_a, c1_b = (c1_clip_a - c1_mean) / c1_std, (c1_clip_b - c1_mean) / c1_std

   fi1_mean = 24.; COV2=0.1; fi1_std = fi1_mean*COV2; fi1_clip_a = fi1_mean-3.*fi1_std; fi1_clip_b = fi1_mean+3.*fi1_std
   fi1_a, fi1_b = (fi1_clip_a - fi1_mean) / fi1_std, (fi1_clip_b - fi1_mean) / fi1_std

   lamda1_mean = 0.188; COV3 = 0.1; lamda1_std = lamda1_mean*COV3; lamda1_clip_a = lamda1_mean-3.*lamda1_std ; lamda1_clip_b =
lamda1_mean+3.*lamda1_std
   lamda1_a, lamda1_b = (lamda1_clip_a - lamda1_mean) / lamda1_std, (lamda1_clip_b - lamda1_mean) / lamda1_std

   kappa1_mean = 0.047; COV4 = 0.1; kappa1_std = kappa1_mean*COV4; kappa1_clip_a = kappa1_mean-3.*kappa1_std ; kappa1_clip_b =
kappa1_mean+3.*kappa1_std
   kappa1_a, kappa1_b = (kappa1_clip_a - kappa1_mean) / kappa1_std, (kappa1_clip_b - kappa1_mean) / kappa1_std

   einit1_mean = 2.; COV5 = 0.07; einit1_std = einit1_mean*COV5; einit1_clip_a = einit1_mean-3.*einit1_std ; einit1_clip_b =
einit1_mean+3.*einit1_std
   einit1_a, einit1_b = (einit1_clip_a - einit1_mean) / einit1_std, (einit1_clip_b - einit1_mean) / einit1_std

   # Soil2: Very Soft Clay
   c2_mean = 6.3; COV1 = 0.1; c2_std = c2_mean*COV1; c2_clip_a = c2_mean-3.*c2_std ; c2_clip_b = c2_mean+3.*c2_std
   c2_a, c2_b = (c2_clip_a - c2_mean) / c2_std, (c2_clip_b - c2_mean) / c2_std

   fi2_mean = 19.3; COV2=0.1; fi2_std = fi2_mean*COV2; fi2_clip_a = fi2_mean-3.*fi2_std; fi2_clip_b = fi2_mean+3.*fi2_std
   fi2_a, fi2_b = (fi2_clip_a - fi2_mean) / fi2_std, (fi2_clip_b - fi2_mean) / fi2_std

   lamda2_mean = 0.137; COV3 = 0.1; lamda2_std = lamda2_mean*COV3; lamda2_clip_a = lamda2_mean-3.*lamda2_std ; lamda2_clip_b =
lamda2_mean+3.*lamda2_std
   lamda2_a, lamda2_b = (lamda2_clip_a - lamda2_mean) / lamda2_std, (lamda2_clip_b - lamda2_mean) / lamda2_std

   kappa2_mean = 0.0457; COV4 = 0.1; kappa2_std = kappa2_mean*COV4; kappa2_clip_a = kappa2_mean-3.*kappa2_std ; kappa2_clip_b =
kappa2_mean+3.*kappa2_std
   kappa2_a, kappa2_b = (kappa2_clip_a - kappa2_mean) / kappa2_std, (kappa2_clip_b - kappa2_mean) / kappa2_std

   einit2_mean = 2.8; COV6 = 0.07; einit2_std = einit2_mean*COV6; einit2_clip_a = einit2_mean-3.*einit2_std ; einit2_clip_b =
einit2_mean+3.*einit2_std
   einit2_a, einit2_b = (einit2_clip_a - einit2_mean) / einit2_std, (einit2_clip_b - einit2_mean) / einit2_std

   # Generate n_sim values following a truncate normal distribution
   c1_sample     = truncnorm.rvs(c1_a, c1_b, loc=c1_mean, scale=c1_std, size=n_sim, random_state=2022)
   fi1_sample    = truncnorm.rvs(fi1_a, fi1_b, loc=fi1_mean, scale=fi1_std, size=n_sim, random_state=2022)
   lamda1_sample  = truncnorm.rvs(lamda1_a, lamda1_b, loc=lamda1_mean, scale=lamda1_std, size=n_sim, random_state=2022)
   kappa1_sample  = truncnorm.rvs(kappa1_a, kappa1_b, loc=kappa1_mean, scale=kappa1_std, size=n_sim, random_state=2022)
   einit1_sample   = truncnorm.rvs(einit1_a, einit1_b, loc=einit1_mean, scale=einit1_std, size=n_sim, random_state=2022)

   c2_sample     = truncnorm.rvs(c2_a, c2_b, loc=c2_mean, scale=c2_std, size=n_sim, random_state=2022) # random_state=seed
   fi2_sample    = truncnorm.rvs(fi2_a, fi2_b, loc=fi2_mean, scale=fi2_std, size=n_sim, random_state=2022)
   lamda2_sample  = truncnorm.rvs(lamda2_a, lamda2_b, loc=lamda2_mean, scale=lamda2_std, size=n_sim, random_state=2022)
   kappa2_sample  = truncnorm.rvs(kappa2_a, kappa2_b, loc=kappa2_mean, scale=kappa2_std, size=n_sim, random_state=2022)
   einit2_sample   = truncnorm.rvs(einit2_a, einit2_b, loc=einit2_mean, scale=einit2_std, size=n_sim, random_state=2022)

   df = pd.DataFrame(list(zip(c1_sample, fi1_sample, lamda1_sample, kappa1_sample, einit1_sample,
```

A6. Python Script. The Green Office Excavation Failure. Slope Stability Model

```
                    c2_sample, fi2_sample, lamda2_sample, kappa2_sample, einit2_sample)), # Create a data frame using the values
                columns= ['c1', 'phi1', 'lamda1', 'kappa1', 'eini1', 'c2', 'phi2', 'lamda2', 'kappa2', 'eini2' ])

    var_names = ['xc','yc','r', 'x1', 'y1', 'x2', 'y2', 'fs'] # variable names to read
    method_names = ['or_fell', 'bish_simp', 'jan_simp', 'ja_corr' , 'spenc']  # methods names
    col_names =  [y+'-'+x for x in method_names for y in var_names ]  # combine var_names and method names https://blog.finxter.com/how-to-
write-a-nested-for-loop-in-one-line-python/
    for t in col_names:   # Create columns in the
      df[t]       = np.repeat('NA', len(c1_sample))
    return(df)

# This function use the original "Modelo_slide_M1.sli" and replace the c and phi values from the dataframe df
def create_sli(it):

    filename = "modelo_slide_m1.sli" # Create file name based on excel codification File "List_of_files.xlsx"
    dir_ini = "C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/Modelos_slide/source_files/" # Directory where the initial file is located
    dir_fin = "C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/Modelos_slide/simulations/" #directory were the copy file will be located

    original = dir_ini + filename     # complete path of the original file
    target   = dir_fin + filename     # complete path of the copy file with a value indicating the iteration number
    shutil.copyfile(original, target)       # command to copy the file
    base = os.path.splitext(target)[0] # split the file name and its extension
    os.rename(target, base + ".txt") # change extension of the file
    work_file = dir_fin  + "modelo_slide_m1.txt"

    g = open(work_file, 'r') # open the work_file
    lines = g.readlines()  # read each line and convert the to a list

    # Replace line 34 and 35 with the c- and phi- values for soft and very soft soils
    lines[34] = "  soil1 = type: 0 water: 1 wtable: 0 c: "+str(float(df.iloc[it,0]))+" phi: "+str(float(df.iloc[it,1]))+" uw: 14.8 hutype: 0\n"
    lines[35] = "  soil2 = type: 0 water: 1 wtable: 0 c: "+str(float(df.iloc[it,5]))+" phi: "+str(float(df.iloc[it,6]))+" uw: 13.4 hutype: 0\n"
    lines = ''.join(lines)    # change list to string

    g = open(work_file, "wt") #open work_file in write mode
    g.write(lines)           #write the work_file with the variable lines
    g.close()               #close the file
    os.rename(work_file, dir_fin + "Modelo_slide_m1.sli") # change extension of the file

# This function change the extension of the output file *.s01 to *.txt in oder to read results
def create_txt():
    dir_fin ="C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/Modelos_slide/simulations/" #directory were the copy file will be located
    for i in os.listdir(dir_fin):
      files = os.path.join(dir_fin,i)
      split= os.path.splitext(files)
      if split[1]=='.s01':
        os.rename(files,split[0]+'.txt')

#This function read the values of the output file
def read_value(iii):
    dir_fin = "C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/Modelos_slide/simulations/" #directory were the copy file will be located
    g = open(dir_fin  + "modelo_slide_m1.txt")
    line = g.readlines()
    id1 = line.index("* Global Minimum FS (xc,yc,r,x1,y1,x2,y2,fs,name)\n")  # Find the line in the *txt file with the value "* Global Minimum FS
(xc,yc,r,x1,y1,x2,y2,fs,name)\n"
    fellenius_line   = line[id1+1].split()[0:8] #select the 7 first data
    bishop_sim_line  = line[id1+2].split()[0:8]
    janbu_sim_line   = line[id1+3].split()[0:8]
    janbu_cor_line   = line[id1+4].split()[0:8]
    spencer_line     = line[id1+5].split()[0:8]
    all_values       = [fellenius_line, bishop_sim_line, janbu_sim_line, janbu_cor_line, spencer_line]
    #all_values       = np.concatenate(all_values) # convertlist to a single vector
    #print(values_stability)
    return(all_values)

#_____________________________________________________________________
#  COMMNANDS TO RUN SLIDE CONSOLE AND READ OUTPUT VALUES
#_____________________________________________________________________
```

A6. Python Script. The Green Office Excavation Failure. Slope Stability Model

```
# Loop to delete all previous files
#path = "C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/Modelos_slide/simulations/"
#for f in os.listdir(path):
#    os.remove(os.path.join(path, f))

n_sim  = 1000   # Number of simulations
df     = data_frame(n_sim)

################################################################################
#
#
################################################################################
def cre_txt():
   dir_fin = "C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/Modelos_slide/simulations/" #directory were the copy file will be located
   file_name = "modelo_slide_m1.s01"
   file = dir_fin + file_name
   pre, ext = os.path.splitext(file)
   os.rename(file,pre +'.txt')

# Loop to create the *.sli files

for index, row in tqdm(df.iterrows()):

   create_sli(index)

# Run the Slide Console

   filename = "modelo_slide_m1.sli"

   subprocess.run(["C:/Program Files (x86)/Rocscience/Slide 5.0/aslidew.exe",
                   "C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/Modelos_slide/simulations/"+filename ], check= False)
   # Create *.txt files
   cre_txt()

   for t in range(5):
      for s in range(8):
         df.iloc[index,8*t+s+10 ]  = read_value(index)[t][s]

   path = "C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/Modelos_slide/simulations"
   for f in os.listdir(path):
      os.remove(os.path.join(path, f))

df.to_csv("C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/Modelos_slide/"+'results_M1.csv')
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

```
# PLAXIS PRELIMINARIES ###
import subprocess  # Use the subprocess module
import os        # To use the cmd? (verify)
from plxscripting.easy import *  # Import PLAXIS library

import math
import numpy as np
import random
from scipy.stats import truncnorm
#import matplotlib.pyplot as plt
import pandas as pd
from tqdm import tqdm

inputport = 10000
plaxispw= r'**********'  # Password
plaxis_path = r'C:\Program Files\Bentley\Geotechnical\***** #path of executable
plaxis_input = 'PLAXIS2DxInput.exe--AppServerPort=21403 --NO_CONTROLLERS'

# Connect to PLAXIS application
s_i, g_i = new_server('localhost', 10000, password = plaxispw) # Conecta al servidor de entrada
s_o, g_o = new_server('localhost', 10001, password = plaxispw) # Conecta al servidor de salida
#s_i.new()
```

In [2]:

```
df = pd.read_csv ("C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/input_Pijao_plaxis1K.csv")
point_names = ["T1','T2','T3','T4','T5','T6','T7','T8','T9','T10',
'W1','W2','W3','W4','W5','W6','W7','W8','W9','W10','B1','B2','B3','B4','B5','B6','B7','B8','B9','B10']
step_names  = ['Phase3','Phase4', 'Phase5' ]
col_names = [y+'-'+x for x in step_names for y in point_names ]
for t in col_names:   # Create columns in df
   df[t] = np.repeat('NA', len(df.index))
#print(df.head())
```

In [3]:

```
# STRATIGRAPHY AND GEOMETRY (SOIL MODE)
def soil_stratigraphy(s_i, g_i):
   #inputport = 10000
   #plaxispw= plaxispw #Password
   #plaxis_path = r'C:\Program Files\Bentley\Geotechnical\PLAXIS 2D CONNECT Edition V20' #path of executable
   #plaxis_input = 'PLAXIS2DxInput.exe--AppServerPort=21403 --NO_CONTROLLERS'

   # Connect to PLAXIS application
   #s_i, g_i = new_server('localhost', 10000, password = plaxispw) # Conecta al servidor de entrada
   #s_o, g_o = new_server('localhost', 10001, password = plaxispw) # Conecta al servidor de salida


   s_i.new()
   g_i.SoilContour.initializerectangular(-120, 0 ,120 ,50) #define project limits
   g_i.borehole(0)                  # add a borehole at coodinate x=0
   g_i.borehole_1.Head = 50         # Set phreatic level
   g_i.soillayer(0)                 # Create a soil layer of thick=0. Default name Soillayer_1
   g_i.soillayer(1)                 # Create a soil layer of thick=0. Default name Soillayer_2
   g_i.soillayer(2)                 # Create a soil layer of thick=0. Default name Soillayer_3
   g_i.soillayer(3)                 # Create a soil layer of thick=0. Default name Soillayer_4
   g_i.Soillayer_1.Zones[0].Top = 50      # Modify top position of Soillayer_1 to y = 30
   g_i.Soillayer_1.Zones[0].Bottom = 44   # Modify bottom position of Soillayer_1 to y = 44
   g_i.Soillayer_2.Zones[0].Bottom = 15
   g_i.Soillayer_3.Zones[0].Bottom = 10
   g_i.Soillayer_4.Zones[0].Bottom = 0
```

In [4]:

```
# CREATE AND ASIGN SOIL MATERIALS (SOIL MODE)
def soil_material(s_i, g_i, c1, phi1, lamda1, kappa1, einit1,
            c2, phi2, lamda2, kappa2, einit2):
   soil1 = g_i.soilmat()                  # Default name : Soil_1
   soil2 = g_i.soilmat()                  # Default name : Soil_2
   soil3 = g_i.soilmat()                  # Default name : Soil_3
   soil4 = g_i.soilmat()                  # Default name : Soil_4

   ### Soil 1 Properties ###
   soil1.setproperties(
   # General tab
   "MaterialName", "Soft Clay",              # Name of the soil material.
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

```
    "SoilModel", 5,                    # Material model. 1 Linear Elastic, 2 Mohr-Coulomb, 3 Hardening Soil... see
file:///C:/Program%20Files/Bentley/Geotechnical/PLAXIS%202D%20CONNECT%20Edition%20V20/Manuals/English/input_objects/objects_Soi
lMat.html
    "DrainageType", 1,                 # Drainage type. 0 Drained, 1 Undrained (A) 3 Undrained (C)...
    "Colour", 11129544,                # Colour of the soil material in OLE colour notation
    "MaterialNumber", 0,               # Identification number of the soil material
    "UserDefinedIndex", 0,             # Index number of the soil model from the DLL file that is currently used
    "Comments", "Soil for Green Office Building",   # Comments on the soil material
    "gammaUnsat", 13.8,
    "gammaSat", 14.8,
    "DilatancyCutOff", False,
    "ninit",einit1/(1+einit1),                 # porosity n=0.667 to get a initial void ratio e_init = 2.0.
    # Parameters tab
    "lambdaModified", lamda1,        # λ* = Cc/(2.3(1+e))
    "kappaModified", kappa1,                # (λ* / κ* ) = 3.0 (reference: PLaxis Material Models Manual. p. 116)
    "cref", c1,                      # Reference value of the cohesion
    "phi", phi1,                     # Friction angle of the soil material
    "psi", 0.0,                      # Dilatancy angle of the soil material
    "DefaultValuesAdvanced", True,
    #"nu", 0.3,                      # Equivalent to nurf
    "UndrainedBehaviour", 0,                # Undrained behaviour 0 Standard Undrained behaviour of the soil material is standard, 1 Manual
    # Groundwater tab
    #"DataSetFlow", 0,
    #"ModelFlow", 0,
    "SoilTypeFlow", 4,
    #"UseDefaultsFlow", 1,
    #"FlowDataModel" , 0 ,              # 000 Standard, 100 Hypres,  200 USDA,  300 Staring, 400 User define
    #"DefaultValuesFlow",True,          # False Disables the use of default values for the flow parameters
    "perm_primary_horizontal_axis",0.5996,       # Permeability of the soil along the primary horizontal axis
    "perm_vertical_axis", 0.5996,  #

    # Interfaces tab
    "InterfaceStiffness", 0,           # 0 standard 1 manual
    "InterfaceStrength", 0,            # 1 Rigid Rinter = 1.0, 0 Manual
    "Rinter", 0.67,                  # Strength reduction factor for interfaces
    "ConsiderGapClosure", True,
    # Initial tab
    "K0Determination", 1,
    "OCR", 3.0,
    "POP", 0.0)

    ### Soil 2 Properties ###
    soil2.setproperties(
    # General tab
    "MaterialName", "Very Soft Clay",          # Name of the soil material.
    "SoilModel", 5,                  # Material model. 1 Linear Elastic, 2 Mohr-Coulomb, 3 Hardening Soil... see
file:///C:/Program%20Files/Bentley/Geotechnical/PLAXIS%202D%20CONNECT%20Edition%20V20/Manuals/English/input_objects/objects_Soi
lMat.html
    "DrainageType", 1,                 # Drainage type. 0 Drained, 1 Undrained (A) 3 Undrained (C)...
    "Colour", 8433599,                 # Colour of the soil material in OLE colour notation
    "MaterialNumber", 0,               # Identification number of the soil material
    "UserDefinedIndex", 0,             # Index number of the soil model from the DLL file that is currently used
    "Comments", "Soil for Green Office Building",   # Comments on the soil material
    "gammaUnsat", 12.4,
    "gammaSat", 13.4,
    "DilatancyCutOff", False,
    "ninit", einit2/(1+einit2),                # porosity n=0.0.736 to get a initial void ratio e_init = 2.8.
    # Parameters tab
    "lambdaModified", lamda2,        # λ* = Cc/(2.3(1+e))
    "kappaModified", kappa2,                # (λ* / κ* ) = 3.0 (reference: PLaxis Material Models Manual. p. 116)
    "cref", c2,                      # Reference value of the cohesion
    "phi", phi2,                     # Friction angle of the soil material
    "psi", 0.0,                      # Dilatancy angle of the soil material
    "DefaultValuesAdvanced", True,
    "nu", 0.3,                       # Equivalent to nurf
    "UndrainedBehaviour", 0,                # Undrained behaviour 0 Standard Undrained behaviour of the soil material is standard, 1 Manual
    # Groundwater tab
    "FlowDataModel" , 000 ,            # 000 Standard, 100 Hypres,  200 USDA,  300 Staring, 400 User define
    "DefaultValuesFlow", False,          # False Disables the use of default values for the flow parameters
    "perm_primary_horizontal_axis", 0.5996,       # Permeability of the soil along the primary horizontal axis
    "perm_vertical_axis", 0.5996,  #
    # Interfaces tab
    "InterfaceStiffness", 0,           # 0 standard 1 manual
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

```
    "InterfaceStrength", 0,                 # 1 Rigid Rinter = 1.0, 0 Manual
    "Rinter", 0.67,                         # Strength reduction factor for interfaces
    "ConsiderGapClosure", True,
    # Initial tab
    "K0Determination", 1,
    "OCR", 1.0,
    "POP", 0.0)

    ### Soil 3 Properties ###
    soil3.setproperties(
    # General tab
    "MaterialName", "Clayed sand",          # Name of the soil material.
    "SoilModel", 2,                         # Material model. 1 Linear Elastic, 2 Mohr-Coulomb, 3 Hardening Soil... see
file:///C:/Program%20Files/Bentley/Geotechnical/PLAXIS%202D%20CONNECT%20Edition%20V20/Manuals/English/input_objects/objects_Soi
lMat.html
    "DrainageType", 0,                      # Drainage type. 0 Drained, 1 Undrained (A) 3 Undrained (C)...
    "Colour", 9160924,                      # Colour of the soil material in OLE colour notation
    "MaterialNumber", 0,                    # Identification number of the soil material
    "UserDefinedIndex", 0,                  # Index number of the soil model from the DLL file that is currently used
    "Comments", "Soil for the Benchmark problem",   # Comments on the soil material
    "gammaUnsat", 16.8,
    "gammaSat", 17.8,
    "DilatancyCutOff", False ,
    # Parameters tab
    "Eref", 1146,                           # Reference Young's modulus
    #"E50ref", 2E4,                         # Secant stiffness in standard drained triaxial test
    #"EoedRef", 2E4,                        # Tangent stiffness for primary oedometer loading
    #"EurRef", 6E4,                         # Unloading/reloading stiffness at engineering strains
    #"powerm", 0.5,                         # Power for stress-level dependency of stiffness
    "cref", 0,                              # Reference value of the cohesion
    "phi", 34.8,                            # Friction angle of the soil material
    "psi", 0.0,                             # Dilatancy angle of the soil material
    "nu", 0.3,                              # Equivalent to nurf
    "DefaultValuesAdvanced", True,
    #"K0nc", 0.5,                           # Coefficient of lateral earth pressure for a normally consolidated stress state
    "UndrainedBehaviour", 0,                # Undrained behaviour 0 Standard Undrained behaviour of the soil material is standard, 1 Manual
    # Groundwater tab
    "FlowDataModel" , 000 ,                 # 000 Standard, 100 Hypres,  200 USDA,  300 Staring, 400 User define
    "DefaultValuesFlow", False,             # False Disables the use of default values for the flow parameters
    "perm_primary_horizontal_axis", 0.5996,     # Permeability of the soil along the primary horizontal axis
    "perm_vertical_axis", 0.5996,  #
    # Interfaces tab
    "InterfaceStiffness", 0,                # 0 standard 1 manual
    "InterfaceStrength", 0,                 # 1 Rigid Rinter = 1.0, 0 Manual
    "Rinter", 0.67,                         # Strength reduction factor for interfaces
    "ConsiderGapClosure", True,
    # Initial tab
    "K0Determination", 1,
    "OCR", 1.0,
    "POP", 0.0)

  ### Soil 4 Properties ###
    soil4.setproperties(
    # General tab
    "MaterialName", "Colluvion",            # Name of the soil material.
    "SoilModel", 2,                         # Material model. 1 Linear Elastic, 2 Mohr-Coulomb, 3 Hardening Soil... see
file:///C:/Program%20Files/Bentley/Geotechnical/PLAXIS%202D%20CONNECT%20Edition%20V20/Manuals/English/input_objects/objects_Soi
lMat.html
    "DrainageType", 0,                      # Drainage type. 0 Drained, 1 Undrained (A) 3 Undrained (C)...
    "Colour", 8962011,                      # Colour of the soil material in OLE colour notation
    "MaterialNumber", 0,                    # Identification number of the soil material
    "UserDefinedIndex", 0,                  # Index number of the soil model from the DLL file that is currently used
    "Comments", "Soil for the Benchmark problem",   # Comments on the soil material
    "gammaUnsat", 17.5,
    "gammaSat", 18.5,
    "DilatancyCutOff", False ,
    # Parameters tab
    "Eref", 1736,                           # Reference Young's modulus
    #"E50ref", 2E4,                         # Secant stiffness in standard drained triaxial test
    #"EoedRef", 2E4,                        # Tangent stiffness for primary oedometer loading
    #"EurRef", 6E4,                         # Unloading/reloading stiffness at engineering strains
    #"powerm", 0.5,                         # Power for stress-level dependency of stiffness
    "cref", 0,                              # Reference value of the cohesion
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

```
    "phi", 42.4,                    # Friction angle of the soil material
    "psi", 0.0,                     # Dilatancy angle of the soil material
    "nu", 0.3,                      # Equivalent to nurf
    "DefaultValuesAdvanced", True,
    #"K0nc", 0.5,                  # Coefficient of lateral earth pressure for a normally consolidated stress state
    "UndrainedBehaviour", 0,        # Undrained behaviour 0 Standard Undrained behaviour of the soil material is standard, 1 Manual
    # Groundwater tab
    "FlowDataModel" , 000 ,         # 000 Standard, 100 Hypres,  200 USDA,  300 Staring, 400 User define
    "DefaultValuesFlow", False,     # False Disables the use of default values for the flow parameters
    "perm_primary_horizontal_axis", 0.5996,  # Permeability of the soil along the primary horizontal axis
    "perm_vertical_axis", 0.5996,  #
    # Interfaces tab
    "InterfaceStiffness", 0,        # 0 standard 1 manual
    "InterfaceStrength", 0,         # 1 Rigid Rinter = 1.0, 0 Manual
    "Rinter", 0.67,                 # Strength reduction factor for interfaces
    "ConsiderGapClosure", True,
    # Initial tab
    "K0Determination", 1,
    "OCR", 1.0,
    "POP", 0.0)

    # ASSIGN MATERIALS TO LAYERS
    g_i.Soils[0].Material = soil1   # Assign soil 1 to polygon Soils[0]
    g_i.Soils[1].Material = soil2   # Assign soil 2 to polygon Soils[1]
    g_i.Soils[2].Material = soil3   # Assign soil 3 to polygon Soils[2]
    g_i.Soils[3].Material = soil4   # Assign soil 4 to polygon Soils[3]
```

In [5]:

```
# DEFINE AND ASSIGN STRUCTURAL ELEMENTS (STRUCTURES MODE)
def structural_elements(s_i, g_i):      # h_wall = wall total lenght
    g_i.gotostructures()            # Move to Structures tab

    # WALLS
    # Define wall lines
    left_wall_line  = g_i.line(-35.50, 50, -35.50, 33) [-1] # Left wall. the value [-1] creates the object in plaxis
    right_wall_line = g_i.line(+36.50, 50, +36.50, 33) [-1] # Right wall

    # Define wall properties and assign to plates
    # According to the implementation of the "platemat" command in PLAXIS, for the elastic part,
    # you need to specify, at least, the following quantities:
    #"EA" (...) "EA2" (...) "Gref" (...) "nu" (...)"d" (...)
    # It is better to calculate d, E and G from equations.

    # Wall Material
    wall_mat = g_i.platemat()               # Create a plate (wall) material
    wall_mat.setproperties(                 # Set the properties of the wall material
    # Material set
    "MaterialName", "Walls",
    "Colour", 16711680,
    "Comments", "Structural Wall Properties",
    "Elasticity", 0,
    "IsIsotropic", True,
    #Properties
    "EA", 3.0E7*(0.4*1.0),                  # E=3.0E7 ; A=0.4*1.0
    "EA2",3.0E7*(0.4*1.0),
    "EI", 3.0E7*1.0*math.pow(0.4, 3)/12,    # I=1.0*d^3/12 = 5.33E-3,
    "nu", 0.2,
    "w", 24*1.0*0.4,                        # weight per length. 24*0.4
    "d", math.sqrt(12 * (3.0E7*1.0*math.pow(0.4, 3)/12) / (3.0E7*0.4*1.0)), # d = math.sqrt(12 * EI / EA)
    "Gref", 3.0E7 / (2 * (1 + 0.2)),        # Gref = E / (2 * (1 + nu))
    "PreventPunching", True)

    # Create plates (walls)
    Left_Wall = g_i.plate(left_wall_line,  "Material", wall_mat)    # Create a plate in left_wall_line and assign the wall_material
    Right_Wall= g_i.plate(right_wall_line, "Material", wall_mat)

    # Create Interfaces
    g_i.neginterface(left_wall_line, right_wall_line )      # negative interface
    g_i.posinterface(left_wall_line, right_wall_line )      # positive interface

    # PILES
    # Define pile lines
    pile_line_7 = g_i.line(-30.50, 39,  -30.50, 10) [-1]  # create a line for pile in axis 7-7. [-1] draw the line in plaxis
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

```
pile_line_6 = g_i.line(-19.75, 39,   -19.75, 10) [-1]
pile_line_5 = g_i.line(-09.75, 39,   -09.75, 10) [-1]
pile_line_4 = g_i.line(+00.25, 39,   +00.25, 10) [-1]
pile_line_3 = g_i.line(+10.25, 39,   +10.25, 10) [-1]
pile_line_2 = g_i.line(+20.25, 39,   +20.25, 10) [-1]
pile_line_1 = g_i.line(+31.75, 39,   +31.75, 10) [-1]

# Pile material
pile_mat = g_i.embeddedbeammat()
pile_mat.setproperties(
# Material Set
"MaterialName", "Pile",
"Colour", 0,
"Comments", "Structural Pile Properties",
"Elasticity", 0,
# Properties
"E", 3.0E7,
"w", 24.0,
"BeamType", 0,
"PredefinedBeamType", 0,
"Diameter", 0.8,
"A", math.pi*math.pow(0.8, 2)/4,
"I", 1/4*math.pi*math.pow(0.8, 4),
"Lspacing", 10,
"SkinResistance", 0,
"Tstart", 1,
"Tend", 100,
"LateralResistance", 0,
"Fmax", 100,
"DefaultStiffnessFactors", True
)
# Create embeddedbeamrow (Piles)
Piles = g_i.embeddedbeamrow(pile_line_7, pile_line_6, pile_line_5,
                    pile_line_4, pile_line_3, pile_line_2,
                    pile_line_1, "Material", pile_mat)
# COLUMNS
# Define column lines
# Assign a variable to each axis lines. [-3:] means that includes all 3 lines created

C7_7 = g_i.line((-30.50, 50), (-30.50, 46), (-30.50, 42.3), (-30.50, 39) ) [-3:] # Column lines for axis 7-7
C6_6 = g_i.line((-19.75, 50), (-19.75, 46), (-19.75, 42.3), (-19.75, 39) ) [-3:]
C5_5 = g_i.line((-09.75, 50), (-09.75, 46), (-09.75, 42.3), (-09.75, 39) ) [-3:]
C4_4 = g_i.line((+00.25, 50), (+00.25, 46), (+00.25, 42.3), (+00.25, 39) ) [-3:]
C3_3 = g_i.line((+10.25, 50), (+10.25, 46), (+10.25, 42.3), (+10.25, 39) ) [-3:]
C2_2 = g_i.line((+20.25, 50), (+20.25, 46), (+20.25, 42.3), (+20.25, 39) ) [-3:]
C1_1 = g_i.line((+31.75, 50), (+31.75, 46), (+31.75, 42.3), (+31.75, 39) ) [-3:]

# Define Column Material
column_mat = g_i.anchormat()
column_mat.setproperties(
 # Material Set
"MaterialName", "Column",
"Colour", 0,
"Comments", "Structural Column Properties",
"Elasticity", 0,
"EA", 3.0E7*math.pi*math.pow(0.8, 2) / 4,
"Lspacing", 10
)
# Assign nod2node anchor (columns) to each axis lines
Columns7_7 = g_i.n2nanchor(C7_7,"Material", column_mat)
Columns6_6 = g_i.n2nanchor(C6_6,"Material", column_mat)
Columns5_5 = g_i.n2nanchor(C5_5,"Material", column_mat)
Columns4_4 = g_i.n2nanchor(C4_4,"Material", column_mat)
Columns3_3 = g_i.n2nanchor(C3_3,"Material", column_mat)
Columns2_2 = g_i.n2nanchor(C2_2,"Material", column_mat)
Columns1_1 = g_i.n2nanchor(C1_1,"Material", column_mat)

# FOUNDATIONS SLAB
foundation_slab_line  = g_i.line(-35.50, 39, 36.50, 39) [-1] # Left wall. the value [-1] creates the object in plaxis

# Foundation Slab Material
foundation_mat = g_i.platemat()                    # Create a plate (wall) material
foundation_mat.setproperties(                      # Set the properties of the wall material
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

```
# Material set
"MaterialName", "Foundation",
"Colour", 16681509,
"Comments", "Foundation Slab Properties",
"Elasticity", 0,
"IsIsotropic", True,
#Properties
"EA", 3.0E7*(1.0*1.638),                    # E=3.0E7 ; A: deduce from slab geometry
"EA2", 3.0E7*(1.0*1.638),
"EI", 3.0E7*1.0*math.pow(1.638, 3)/12,                    # I: deduce from slab geometry
"nu", 0.2,
"w", 8.0,                    # weight per length. 24*0.4
"d", math.sqrt(12 * (3.0E7*1.0*math.pow(1.638, 3)/12) / (3.0E7*(1.0*1.638))),
"Gref", 3.0E7 / (2 * (1 + 0.2)),                    # Gref = E / (2 * (1 + nu))
"PreventPunching", False
)

# Create Foundation slab
Foundation_Slab= g_i.plate(foundation_slab_line,  "Material", foundation_mat)

# BASEMENT SLABS
basement_0_line = g_i.line(-35.50, 50, 36.50, 50) [-1]
basement_1_line = g_i.line(-35.50, 46, 36.50, 46) [-1]
basement_2_line = g_i.line(-35.50, 42.3, 36.50, 42.3) [-1]

# Basement Slab Material
basement_mat = g_i.platemat()
basement_mat.setproperties(
# Material set
"MaterialName", "Basement Slab",
"Colour", 16681509,
"Comments", "Basement Slab Properties",
"Elasticity", 0,
"IsIsotropic", True,
#Properties
"EA", 3.0E7*(1.0*1.219),                    # E=3.0E7 ; A: deduce from slab geometry
"EA2", 3.0E7*(1.0*1.219),
"EI", 3.0E7*1.0*math.pow(1.219, 3)/12,                    # I: deduce from slab geometry
"nu", 0.2,
"w", 6.0,                    # weight per length. 24*0.4
"d", math.sqrt(12 * (3.0E7*1.0*math.pow(1.219, 3)/12) / (3.0E7*(1.0*1.219))),
"Gref", 3.0E7 / (2 * (1 + 0.2)),                    # Gref = E / (2 * (1 + nu))
"PreventPunching", False
)

# Pavement Material
pav_mat = g_i.platemat()                    # Create a plate (wall) material
pav_mat.setproperties(                    # Set the properties of the wall material
# Material set
"MaterialName", "Pavement",
"Colour", 16711680,
"Comments", "Pavement Properties",
"Elasticity", 0,
"IsIsotropic", True,
#Properties
"EA", 3.0E7*(0.4*1.0),                    # E=3.0E7 ; A=0.4*1.0
"EA2",3.0E7*(0.4*1.0),
"EI", 3.0E7*1.0*math.pow(0.4, 3)/12,                    # I=1.0*d^3/12 = 5.33E-3,
"nu", 0.2,
"w", 24*1.0*0.4,                    # weight per length. 24*0.4
"d", math.sqrt(12 * (3.0E7*1.0*math.pow(0.4, 3)/12) / (3.0E7*0.4*1.0)), # d = math.sqrt(12 * EI / EA)
"Gref", 3.0E7 / (2 * (1 + 0.2)),                    # Gref = E / (2 * (1 + nu))
"PreventPunching", True)

# Create Basement slab
Foundation_Slab= g_i.plate(basement_0_line, basement_1_line, basement_2_line,  "Material", basement_mat)

# Pavament and Other Buildings
pavement1_line = g_i.line(+37.00, 50, +69.0, 50) [-1] # Pavement 1
building_line  = g_i.line(+69.50, 50, +100.00, 50) [-1] # Adjacent building

Pavement_Building = g_i.plate(pavement1_line, building_line,  "Material", pav_mat)
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

In [6]:

```
# DEFINE AND ASSIGN POINT AND LINE LOADS (STRUCTURES MODE)
def line_point_loads(s_i, g_i):
    # Adjacent Building
    building_line  = g_i.line(+69.50, 50, +100.00, 50) [-1] # Adjacent building
    building_load =  g_i.lineload(building_line, "qy_start", -20)

    # Traffic Point loads
    points_traffic_loads = g_i.point((43,50), (45.8,50), (47,50), (49.5,50), (55,50), (57.5,50), (59,50), (61.5,50)) [-8:]
    traffic_loads = g_i.pointload(points_traffic_loads, "Fy", -90)

    # Pijao building load
    pijao_line1 = g_i.line((-35.5, 50), (36.5, 50))[-1]
    pijao_load1 = g_i.lineload(pijao_line1, "qy_start", -21)  # LineLoad_3
```

In [7]:

```
# GENERATE THE MESH
def meshing_model():
    g_i.gotomesh()                  # Move to Mesh tab
    g_i.mesh(0.12, True)            # Generate a mesh. 0.06: medium coarse. True: Enhance mesh refinements
    #g_i.viewmesh()                 # To view the generated mesh
```

In [ ]:

In [8]:

```
# CONSTRUCTION STAGE
def construction_stages(s_i, g_i):
    g_i.gotostages()
    # Initial phase
    g_i.InitialPhase.Identification = "InitialPhase"    # Name of the phase
    g_i.InitialPhase.DeformCalcType = "K0 procedure"    # K0 procedure to generate initial stresses
    g_i.InitialPhase.PorePresCalcType = "Phreatic"      # Phreatic is selected as the initial pore presure calculation
    g_i.InitialPhase.Deform.IgnoreSuction=True          # Ignore suction

   # PHASE 1: Adjacent building and traffic
    # The features not defaind here remain as default (Calculation Type=Plastic, Pore pressure=Phreatic, ...)
    g_i.phase(g_i.InitialPhase)                         # Create a new phase from InitialPhase. Default name Phase_1
    g_i.setcurrentphase(g_i.Phase_1)                    # Make Phase_1 current
    g_i.Phase_1.Identification = "Building+traffic"     # Name the phase
    g_i.activate((g_i.Plate_7, g_i.Plate_8), (g_i.Phase_1))
    g_i.LineLoad_1.activate(g_i.Phase_1)                # Line load adjacent building
    g_i.PointLoads.activate(g_i.Phase_1)


    # PHASE 2: Activation of walls
    # The features not defaind here remain as default (Calculation Type=Plastic, Pore pressure=Phreatic, ...)
    g_i.phase(g_i.Phase_1)                              # Create a new phase from InitialPhase. Default name Phase_1
    g_i.setcurrentphase(g_i.Phase_2)                    # Make Phase_1 current
    g_i.Phase_2.Identification = "Walls+Foundation"     # Name the phase
    g_i.Phase_2.Deform.ResetDisplacementsToZero = True

    g_i.activate((g_i.Plate_1,g_i.Plate_2, g_i.Plate_4), (g_i.Phase_2))         # Chech the Plate_number in the graphical model

    g_i.Interfaces.activate(g_i.Phase_2)                # Activate all interfaces at phase 1
    g_i.NodeToNodeAnchors.activate(g_i.Phase_2)
    g_i.EmbeddedBeamRows.activate(g_i.Phase_2)

    # PHASE 3 First excavation
    g_i.phase(g_i.Phase_2)                              # Create a new phase from InitialPhase. Default name Phase_2
    g_i.Phase_3.Identification = "First Exavation"      # Name the phase
    g_i.setcurrentphase(g_i.Phase_3)                    # Make Phase_2 current
    g_i.BoreholePolygon_1_1.deactivate(g_i.Phase_3)

    WT_1 = g_i.waterlevel((-120,50), (-35.5,46), (36.5,46), (120,50))
    g_i.set(g_i.Water.GlobalWaterLevel, (g_i.Phase_3), WT_1)
    g_i.Phase_3.PorePresCalcType = "Steady state groundwater flow"
    g_i.Plate_5.activate(g_i.Phase_3)
    #g_i.PointLoads.activate(g_i.Phase_3)

    # PHASE 4 Second excavation
    g_i.phase(g_i.Phase_3)                              # Create a new phase from InitialPhase. Default name Phase_2
    g_i.Phase_4.Identification = "Second Exavation"     # Name the phase
    g_i.setcurrentphase(g_i.Phase_4)                    # Make Phase_2 current
    g_i.BoreholePolygon_1_4.deactivate(g_i.Phase_4)
    g_i.BoreholePolygon_2_1.deactivate(g_i.Phase_4)
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

```
  WT_2 = g_i.waterlevel((-120,50), (-35.5,42.3), (36.5,42.3), (120,50))
  g_i.set(g_i.Water.GlobalWaterLevel, (g_i.Phase_4), WT_2)
  g_i.Phase_4.PorePresCalcType = "Steady state groundwater flow"
  g_i.Plate_6.activate(g_i.Phase_4)
  #g_i.borehole_1.setproperties("Head ",-30)

  # Phase 5 Third excavation
  g_i.phase(g_i.Phase_4)                          # Create a new phase from InitialPhase. Default name Phase_2
  g_i.Phase_5.Identification = "Third Excavation"          # Name the phase
  g_i.setcurrentphase(g_i.Phase_5)                # Make Phase_2 current
  g_i.BoreholePolygon_2_2.deactivate(g_i.Phase_5)

  WT_3 = g_i.waterlevel((-120,50), (-35.5,39), (36.5,39), (120,50))
  g_i.set(g_i.Water.GlobalWaterLevel, (g_i.Phase_5), WT_3)
  g_i.Phase_5.PorePresCalcType = "Steady state groundwater flow"
  g_i.Plate_3.activate(g_i.Phase_5)
  g_i.activate(g_i.LineLoad_2, g_i.Phase_5)
```

In [9]:

```
# RUN THE MODEL
for index, row in tqdm(df.iterrows()):
  soil_stratigraphy(s_i, g_i)
  soil_material(s_i, g_i, df.iloc[index,0], df.iloc[index,1], df.iloc[index,2], df.iloc[index,3], df.iloc[index,4],
          df.iloc[index,5], df.iloc[index,6], df.iloc[index,7], df.iloc[index,8], df.iloc[index,9])

  structural_elements(s_i, g_i)
  line_point_loads(s_i, g_i)
  meshing_model()
  construction_stages(s_i, g_i)

 # RUN THE MODEL
  g_i.calculate()
  g_i.view(g_i.Phase_1)
  # Read deformation results and put them in the dataframe

  #PHASE 3 FIRST EXCAVATION
  #Street and near building
  df.iloc[index,10]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (37, 50))
  df.iloc[index,11]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (41, 50))
  df.iloc[index,12]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (45, 50))
  df.iloc[index,13]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (49, 50))
  df.iloc[index,14]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (53, 50))
  df.iloc[index,15]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (57, 50))
  df.iloc[index,16]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (61, 50))
  df.iloc[index,17]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (65, 50))
  df.iloc[index,18]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (69, 50))
  df.iloc[index,19]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (73, 50))

  # Soil adjacent to wall
  df.iloc[index,20]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 50))
  df.iloc[index,21]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 47))
  df.iloc[index,22]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 44))
  df.iloc[index,23]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 41))
  df.iloc[index,24]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 38))
  df.iloc[index,25]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 35))
  df.iloc[index,26]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 32))
  df.iloc[index,27]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 29))
  df.iloc[index,28]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 26))
  df.iloc[index,29]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Ux, (37, 23))

  # Soil at the bottom of excavation
  df.iloc[index,30]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (0.0, 46))
  df.iloc[index,31]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (3.5, 46))
  df.iloc[index,32]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (7.0, 46))
  df.iloc[index,33]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (10.5, 46))
  df.iloc[index,34]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (14.0, 46))
  df.iloc[index,35]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (17.5, 46))
  df.iloc[index,36]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (21.0, 46))
  df.iloc[index,37]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (24.5, 46))
  df.iloc[index,38]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (28.0, 46))
  df.iloc[index,39]  =  g_o.getsingleresult(g_o.Phase_3, g_o.ResultTypes.Soil.Uy, (31.5, 46))

  #PHASE 4 SECOND EXCAVATION
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

```
#Street and near building
df.iloc[index,40]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (37, 50))
df.iloc[index,41]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (41, 50))
df.iloc[index,42]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (45, 50))
df.iloc[index,43]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (49, 50))
df.iloc[index,44]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (53, 50))
df.iloc[index,45]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (57, 50))
df.iloc[index,46]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (61, 50))
df.iloc[index,47]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (65, 50))
df.iloc[index,48]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (69, 50))
df.iloc[index,49]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (73, 50))

# Soil adjacent to wall
df.iloc[index,50]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 50))
df.iloc[index,51]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 47))
df.iloc[index,52]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 44))
df.iloc[index,53]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 41))
df.iloc[index,54]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 38))
df.iloc[index,55]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 35))
df.iloc[index,56]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 32))
df.iloc[index,57]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 29))
df.iloc[index,58]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 26))
df.iloc[index,59]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Ux, (37, 23))

# Soil at the bottom of excavation
df.iloc[index,60]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (0.0, 42.3))
df.iloc[index,61]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (3.5, 42.3))
df.iloc[index,62]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (7.0, 42.3))
df.iloc[index,63]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (10.5, 42.3))
df.iloc[index,64]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (14.0, 42.3))
df.iloc[index,65]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (17.5, 42.3))
df.iloc[index,66]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (21.0, 42.3))
df.iloc[index,67]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (24.5, 42.3))
df.iloc[index,68]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (28.0, 42.3))
df.iloc[index,69]  =  g_o.getsingleresult(g_o.Phase_4, g_o.ResultTypes.Soil.Uy, (31.5, 42.3))

#PHASE 5 THIRD EXCAVATION
#Street and near building
df.iloc[index,70]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (37, 50))
df.iloc[index,71]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (41, 50))
df.iloc[index,72]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (45, 50))
df.iloc[index,73]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (49, 50))
df.iloc[index,74]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (53, 50))
df.iloc[index,75]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (57, 50))
df.iloc[index,76]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (61, 50))
df.iloc[index,77]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (65, 50))
df.iloc[index,78]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (69, 50))
df.iloc[index,79]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (73, 50))

# Soil adjacent to wall
df.iloc[index,80]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 50))
df.iloc[index,81]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 47))
df.iloc[index,82]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 44))
df.iloc[index,83]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 41))
df.iloc[index,84]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 38))
df.iloc[index,85]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 35))
df.iloc[index,86]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 32))
df.iloc[index,87]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 29))
df.iloc[index,88]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 26))
df.iloc[index,89]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Ux, (37, 23))

# Soil at the bottom of excavation
df.iloc[index,90]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (0.0, 39))
df.iloc[index,91]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (3.5, 39))
df.iloc[index,92]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (7.0, 39))
df.iloc[index,93]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (10.5, 39))
df.iloc[index,94]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (14.0, 39))
df.iloc[index,95]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (17.5, 39))
df.iloc[index,96]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (21.0, 39))
df.iloc[index,97]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (24.5, 39))
df.iloc[index,98]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (28.0, 39))
df.iloc[index,99]  =  g_o.getsingleresult(g_o.Phase_5, g_o.ResultTypes.Soil.Uy, (31.5, 39))
```

A7. Python Script. The Green Office Excavation Failure. Plaxis Model

```
    #g_i.Phase_1.ShouldCalculate=True
    #g_i.Phase_2.ShouldCalculate=True
    #g_i.Phase_3.ShouldCalculate=True
    #g_i.Phase_4.ShouldCalculate=True
    #g_i.Phase_5.ShouldCalculate=True

    g_o.update()
#save_path = r'C:\Users\mgfer\OneDrive - Universidad Nacional de
Colombia\D_drive\Simulations\Excavation\pruebas\Model_ERTC7_plaxis_V1.p2dx'
    #g_i.save(save_path)
#g_i.view(g_i.Phase_4)
```

A8. R Script. The Green Office Excavation Failure. Bayesian Network and K MPE

```
#_____________________________________________________________________
# 2020-01-05
# By: wmgarciaf@unal.edu.co
#
# R SCRIPT para an?lisis Bayesiano
# Red Bayesiana para an?lisis modelo Pijao
#_____________________________________________________________________

#-------- PRELIMINARES ----------
setwd("C:/Users/mgfer/OneDrive - Universidad Nacional de Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1")
df = read.csv(file="Condensed_Results_BN_xxx.csv", header = T)
# Convertir todas las columnas de "numeric" or "character" a "factor"
for (i in 1: ncol(df)) {
  df[, i] = as.factor(df[, i])
}
# -------- GRAFICOS PRELIMINARES ----------
windows(20, 20)
par(mfrow= c(3,3))
barplot(table(df$LT))
barplot(table(df$WT))
barplot(table(df$FoS))
barplot(table(df$SC))
barplot(table(df$Ro))
barplot(table(df$i))
barplot(table(df$h))


# -------- CONSTRUCCION RED BAYESIANA ----------
library(bnlearn)
model1 = paste0("[Ar]","[Dm]","[LT]", "[Ra|Ar]", "[WT|Dm:Ar]", "[In|Dm]", "[Pz|WT]",
            "[FoS|LT:WT]", "[SC|FoS]", "[Ro|SC]", "[i|SC]", "[h|SC]", "[Eo|SC]")
dag1   = model2network(model1)
windows(20, 20)
graphviz.plot(dag1)

# LABELS
In_label = c("Damage", "No_Damage")
Ar_label = c("True", "False")
Dm_label = c("True", "False")
Ra_label = c("Records", "No_Records")
LT_label = c("LT1", "LT2")
FoS_label= c("<1.0", ">1.0")
WT_label = c("0.0m", "-3.0m")
Pz_label = c("0.0m_lecture", "-3.0m_lecture")
SC_label = c("Stable", "Unstable")
Ro_label = c("<0.3", ">0.3")
i_label  = c("<0.2", ">0.2")
h_label  = c("<0.3", ">0.3")
Eo_label = c("Observe_failure", "Do_not_observe_failure")

# CPTs
In_prob = array(c(0.99, 0.01, 0.05, 0.95), dim = c(2,2),
            dimnames = list(In=In_label, Dm=Dm_label))
Ar_prob = array(c(0.5, 0.5), dim = 2, dimnames = list(Ar=Ar_label))
Dm_prob = array(c(0.5, 0.5), dim = 2, dimnames = list(Dm=Dm_label))
Ra_prob = array(c(0.99, 0.01, 0.05, 0.95), dim = c(2,2),
            dimnames = list(Ra=Ra_label, Ar=Ar_label))
LT_prob = array(c(0.5, 0.5), dim = 2, dimnames = list(LT=LT_label))
WT_prob = array(c(0.99, 0.01, 0.95, 0.05, 0.15, 0.85,0.01, 0.99), dim = c(2,2,2),
            dimnames = list(WT=WT_label, Ar=Ar_label, Dm=Dm_label))
Pz_prob = array(c(0.99, 0.01, 0.05, 0.95), dim = c(2,2),
            dimnames = list(Pz=Pz_label, WT=WT_label))
FoS_prob= array(c(0.999, 0.001, 0.486, 0.514, 0.999, 0.001,0.094, 0.906), dim = c(2,2,2),
            dimnames = list(FoS=FoS_label, LT=LT_label, WT=WT_label))
SC_prob = array(c(0.0, 1.0, 1.0, 0.0), dim = c(2,2),
            dimnames = list(SC=SC_label, FoS=FoS_label))
Ro_prob = array(c(0.64, 0.36, 0.04, 0.96), dim = c(2,2),
            dimnames = list(Ro=Ro_label, SC=SC_label))
i_prob  = array(c(0.999, 0.001, 0.797, 0.203), dim = c(2,2),
            dimnames = list(i=i_label, SC=SC_label))
h_prob  = array(c(0.999, 0.001, 0.225, 0.775), dim = c(2,2),
            dimnames = list(h=h_label, SC=SC_label))
Eo_prob = array(c(0.2, 0.8, 1.0, 0.0), dim = c(2,2),
```

A8. R Script. The Green Office Excavation Failure. Bayesian Network and K MPE

```
            dimnames = list(Eo=Eo_label, SC=SC_label))

cpt = list(In=In_prob, Ar=Ar_prob, Dm=Dm_prob, Ra=Ra_prob, LT=LT_prob, WT=WT_prob,
        Pz=Pz_prob, FoS=FoS_prob, SC=SC_prob, Ro=Ro_prob, i=i_prob, h=h_prob,
        Eo=Eo_prob)

BN_pijao = custom.fit(x = dag1, dist = cpt)

# -------- PREDICTIVE QUERIES  -----------
library(gRain)
junction_BN_Pijao = compile(as.grain(BN_pijao))
querygrain(junction_BN_Pijao, nodes = c("FoS", "WT"))

#Table 5.24
# Evidence e1
e1 = setEvidence(junction_BN_Pijao, nodes = "Eo", states = "Observe_failure")
LTxWT_cpt_1 = querygrain(object = e1, nodes = c("LT","WT" ), type = "joint" )

#Evidence e2
e2 = setEvidence(junction_BN_Pijao, nodes = c("Eo",  "Ro"), states = c("Observe_failure", ">0.3"))
LTxWT_cpt_2 = querygrain(object = e2, nodes = c("LT","WT" ), type = "joint" )

#Evidence e2
e3 = setEvidence(junction_BN_Pijao, nodes = c("Eo",  "Ro", "i", "h"),
             states = c("Observe_failure", ">0.3", ">0.2", "<0.3"))
LTxWT_cpt_3 = querygrain(object = e3, nodes = c("LT","WT" ), type = "joint" )

#Examples for section 5.3
# Case No.1 Evidence> Pz=0.0 m lecture,In=No damage
jPz = setEvidence(junction_BN_Pijao, nodes = "Pz", states = "0.0m_lecture")
querygrain(jPz, nodes = c("SC", "WT", "LT"))

# Case # 4. Actual evidence
jAll = setEvidence(junction_BN_Pijao, nodes =c("Ra", "Pz", "In", "Ro", "h", "i") ,
        states = c("Records", "-3.0m_lecture", "Damage", ">0.3",
                   ">0.3", ">0.2"))
LTxWT_cpt_All = querygrain(object = jAll, nodes = c("LT","WT" ), type = "joint" )


# -------- K-MPE. MOST PROBABLE EXPLANATION -----------
# Due to the fact that the BN and the CPTs were created using  a
# combination of expert opinion and numerical models, it is necessary
# to create a simulation from the BN. This simulation is created using
# cpdist function


df_sim = cpdist(fitted = BN_pijao, nodes = c("Dm", "Ar", "WT", "LT", "FoS",
                              "SC", "Eo", "Ro", "i", "h", "Pz"),
            evidence = ((Ro==">0.3")&(i=="<0.2")&(h==">0.3")&
                      (Pz=="-3.0m_lecture")&(In=="Damage")&
                      (Ra=="Records")), n=10^6)

# Identify unique rows in df_sim an count them
library(dplyr)
K_MPE = df_sim%>%group_by_all%>%count
K_MPE = K_MPE[order(K_MPE$n, decreasing = T),]
write.csv(K_MPE, file = "C:/Users/mgfer/OneDrive - Universidad Nacional de
Colombia/D_drive/Simulations/Pijao_V1/Modelos_V1/K_MPE_Pijao.csv")
```