Reliability Benchmarking of Eurocode 7 Design Examples

CIE5050-09: Additional Thesis Project 2021

Muhammad Rayyan

TU Delft

Reliability Benchmarking of Eurocode 7 Design Examples

by

Muhammad Rayyan

5032334

Supervisor:	Dr. ir. A.P. (Bram) van den Eijnden
	Dr. ir. Timo Schweckendiek
Institution:	Delft University of Technology (TU Delft)
Track:	Geo-Engineering
Project Duration:	September, 2021 - December, 2021

Cover Image: Tangguh LNG project, Indonesia (2019)

TU Delft

Summary

The application of reliability analysis in geotechnical engineering is relatively new compared to the other sections of civil engineering such as structural engineering and hydraulic engineering. However, due to its increases use in recent years, reliability analysis is planned to be included extensively in the upcoming Eurocode 7 (EN 1997). This research aims to compare the accuracy and efficiency between the applications of 22 selected reliability methods in 9 selected geotechnical engineering problems with various number of independent variables and modes of failure. The accuracy of the reliability methods are determined based on the Probability of Failure (P_f) errors, while the efficiency is based on the number of realizations (N) each method needs. The Monte Carlo Simulation is found to be the most accurate method despite its shortcomings in efficiency (ranked as the least efficient). Moreover, the FOSM method is found to be the most efficient despite its serious shortcoming in accuracy where it is also ranked as the most inaccurate. However, putting both accuracy and efficiency into account, the AK-MCS 0 order is proven to be the best method when applied to the discussed geotechnical engineering problems. The research also points out the necessity to perform multiple reliability methods for each geotechnical engineering problem.

Preface

This report is the final product of an additional thesis project in Geo-Engineering track at Delft University of Technology (TU Delft) in the Netherlands. Since I am enjoying Python programming in geotechnical engineering very much, I am very grateful when the offer for this topic came up. Moreover, the opportunity also allowed me to further study the reliability application in geotechnical engineering problems, which is relatively new in practice. I hope this additional thesis project would give a contribution for those who are seeking to learn or further understand reliability analysis in the field of geotechnical engineering.

Furthermore, I would like to express my gratitude to everyone who has helped in overcoming daily problems regarding the project, especially to Mr. Bram van den Eijnden and Mr. Timo Schweckendiek who directly guided me throughout the entire process, I really learnt a lot from you.

Last but not least, I would like to extend my gratitude for my wife Rosi and my family for all of the love and support, and my employer PT PP (Persero) Tbk. for all of the support and opportunity.

Muhammad Rayyan
Delft, December 2021

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
ADS	Adaptive Directional Sampling
AIS	Adaptive Importance Sampling
AK- MCS	Active Learning Kriging-based Monte Carlo Simulation
AK-MCS 0	AK-MCS 0 order (simple kriging)
AK-MCS 1	AK-MCS 1 order (universal kriging)
COV	Coefficient of Variation (variant coefficient)
DIY	Do-it-yourself (manual)
DS	Directional Sampling
FORM	First Order Reliability Method
FOSM	First Order Second Moment
FS	Factor of Safety
FS_d	Factor of Safety of the deep failure mode (slope stability problem)
FS_s	Factor of Safety of the shallow failure mode (slope stability problem)
GEOSNet	Geotechnical Safety Network
IS	Importance Sampling
LHS	Latin Hypercube Sampling
MCS	Monte Carlo Simulation
NI	Numerical Integration
NB	Numerical Bisection
OT	OpenTURNS
PTK	Probabilistic Toolkit
SORM	Second Order Reliability Method
SS	Subset Simulation

Symbols

Symbol	Definition	Unit
β	reliability index	[-]
β_{deep}	reliability index of the deep failure (slope stability problem)	[-]
$\beta_{shallow}$	reliability index of the shallow failure (slope stability problem)	[-]
β_{sys}	reliability index of the problem (slope stability problem)	[-]
γ	moist unit weight of the surficial soil (GEOSNet EX1)	[kN/m^3]
$gamma_{sat}$	saturated unit weight of the surficial soil (GEOSNet EX1)	[kN/m^3]
Φ	the standard normal probability density function	[-]
ϕ	effective stress internal friction angle (GEOSNet EX1)	[deg.]

Symbol	Definition	Unit
σ_p	the standard deviation of P_f	[-]
θ	slope inclination (GEOSNet EX1)	[deg.]
$c_{d1}(X)$	deep failure coefficient of layer 1 (slope stability problem)	[-]
$c_{d2}(X)$	deep failure coefficient of layer 2 (slope stability problem)	[-]
$c_{s1}(X)$	shallow failure coefficient of layer 1 (slope stability problem)	[-]
e	void ratio of soil (GEOSNet EX1)	[-]
$f_x(X)$	Joint probability density function of X	[-]
f_{var}	variance factor per variable (Importance Sampling)	[-]
$g(u)$	performance function / limit state function of u	[-]
$g(X)$	performance function / limit state function of X	[-]
G_s	specific gravity of soil (GEOSNet EX1)	[-]
$g_d(X)$	limit state function of the deep failure mode (slope stability problem)	[-]
$g_s(X)$	limit state function of the shallow failure mode (slope stability problem)	[-]
$g_{sys}(X)$	limit state function (slope stability problem)	[-]
H	depth of soil above bedrock (GEOSNet EX1)	[m]
h	height of groundwater above bedrock (GEOSNet EX1)	[m]
I	Indicator function, 0 for failure and 1 for non-failure	[-]
$L(u)$	Linear line as the approximation of g(u)	[-]
N	number of realizations	[-]
$N_{failure}$	number of failure realizations	[-]
P_f	probability of failure	[-] or [%]
$P^N_{f,LHS}$	Latin Hypercube's probability of failure	[-] or [%]
$s_{u,1}$	undrained shear strength of layer 1 soil (slope stability problem)	[kN/m^2]
$s_{u,2}$	undrained shear strength of layer 2 soil (slope stability problem)	[kN/m^2]
s_{var}	variance shift per variable (Importance Sampling)	[-]
u	random standardized variables vector	[-]
u	each realization in standard normal space (Importance Sampling)	[-]
u_{imp}	Importance Sampling realization (Importance Sampling)	[-]
X	random variables vector	[-]

List of Figures

List of Tables

1 Introduction

1.1. Background

The European standard (i.e. code of practice) for design of geotechnical structures, Eurocode 7 (EN 1997), is currently undergoing a revision. Since the Eurocodes are reliability-based, the second generation of Eurocodes (including EN 1997) will include significantly more explicit reliability elements compared to the first generation. This additional graduation topic has been defined to support the work of the task group who is responsible in facilitating the use of reliability-based methods in geotechnical engineering for the Eurocodes revision.

1.2. Problem Analysis

Reliability analysis is needed in geotechnical engineering to deal with the uncertainties it possesses. This is especially true since geotechnical engineers deal with little-known data regarding natural soil properties (among others). In the past, deterministic analysis with uniformly-assigned material properties are normally applied to solve numerous geotechnical engineering problems. However, the soil properties is functions of a statistical distributions, rather than some fixed values which are constant throughout the layers. The deterministic approach also leads to a single factor of safety, ignoring the uncertainty and variability within the soil properties. Therefore, reliability analysis gives a more meaningful definition of stability: i.e. reliability, the probability that failure will not occur (Hicks et al. 2002 [15]).

Geotechnical engineering problems normally deal with many uncertainties and multiple modes of failure. However, reliability analysis application in geotechnical engineering is relatively new compared to other fields of engineering despite the methods have been developing significantly in the last few decades. Moreover, the uses of these methods in geotechnical engineering are increasing remarkably in recent years. Thus comparison regarding their applications are relatively limited. This research aims to compare the accuracy and efficiency between 22 selected reliability methods in their applications to 9 selected geotechnical engineering problems.

1.3. Research Objective

The main objectives of this research are:

- To carry out (full probabilistic) reliability analysis on various geotechnical engineering problems with numerous modes of failure, material distributions, and number of independent variables.
- To benchmark the performance of different reliability methods.
- Recommend the most suitable reliability method for given type of problems.

1.4. Research Method

Multiple problems in Geotechnical Engineering such as slope stability, foundation stability, and consolidation problem are simulated in this research by using several reliability methods. The reliability methods used in this reports and their abbreviations throughout the report are:

- Monte Carlo Simulation (MCS)
- First Order Reliability Method (FORM)
- Second Order Reliability Method (SORM)
- First Order Second Moment (FOSM)
- Subset simulation (SS)
- Directional Sampling (DS)
- Adaptive Directional Sampling (ADS)
- Important Sampling (IS)
- Adaptive Importance Sampling (AIS)
- Latin Hypercube Sampling (LHS)
- Active Learning Kriging-based Monte Carlo Simulation (AK-MCS)
- Numerical Integration (NI)
- Numerical Bisection (NB)

The above-mentioned reliability methods are carried by Deltares' Probabilistic Toolkit software (PTK), Python open source's OpenTurns (OT), and Python-based script (DIY).

1.4.1. Deltares' Probabilistic Toolkit

The Probabilistic Toolkit (PTK) is a software developed by Deltares to analyze the effects of uncertainty to any model. These models range from python scripts to geotechnical and hydrodinamical Deltares and non-Deltares applicaation. One of the main feature of PTK is its ability to compute the probability of undesired events (probability of failure), thus providing the reliability of a model. Therefore, PTK provides some of the reliability methods performed in this research. Reliability methods performed by PTK in this research are MCS (PTK), FORM (PTK), DS (PTK), IS (PTK), LHS (PTK), NI (PTK), and NB (PTK). More details regarding PTK can be found in [7].

1.4.2. OpenTURNS

OpenTURNS (OT) is an open source initiative for the treatment of uncertainties, risks, and statistics. It is funded by Airbus Group, EDF Research and Development, IMACS, ONERA, and Phimeca Engineering. OpenTURNS provides open source C++/Python scipts related to data analysis, probabilistic and reliability modeling, meta modeling, calibration, and functional modeling. This research utilize its reliability modeling features. Reliability methods performed by OT in this research are MCS (OT), FORM (OT), SORM (OT), SS (OT), DS (OT), IS (OT), LHS (OT), and ADS (OT). More details regarding OT can be found in [1].

1.4.3. Do-it-yourself Script (DIY)

Some of the reliability methods are performed by python-based scripts which were based on numerous references mentioned in section 2. Reliability methods performed by python-based DIY scripts in this research are MCS (DIY), FORM (DIY), FOSM (DIY), SS (DIY), DS (DIY), IS (DIY), AIS (DIY), and AK-MCS with 0 and 1 order (AK-MCS 0 and AK-MCS 1).

2

Theoretical Background

Reliability analysis is an analysis in calculating how much the probability of an unwanted event to occur. An unwanted event in this research is the failure of a geotechnical structure. Failure is evaluated through a limit state function (or also known as a performance function), normally written as $g(X)$, where X is the random variables vector which has their own distributions. The limit state function is defined in such a way that a failure occurs when $g(X)$ < 0. The failure probability (P_f) is determined by calculating the cumulative distribution function of $g(X)$ < 0. Mathematically, probability of failure can be written as equation (2.1) where $f_X(X)$ is the joint probability density function of $g(X)$.

$$P_f = \int_{g(X)<0} f_X(X)\, dX \tag{2.1}$$

Moreover, since P_f is normally a really small number, it is also a very common way to express the reliability by a reliability index (β). The reliability index is defined as the inverse of the cumulative standard normal distribution function of P_f, as can be written in equation (2.2).

$$\beta = -\Phi^{-1}(P_f) \tag{2.2}$$

This research implements several reliability methods to determine P_f from the limit state function of several cases, and compare the results. The theoretical background of each reliability method is discussed in the following section.

2.1. Reliability methods levels

Reliability methods could be divided into 4 levels, however this report only involving Level II and Level III methods. The brief description of the levels are as follows (Schweckendiek 2006 [25]).

2.1.1. Level I Methods

Level I method is also known as semi-probabilistic method. It is usually applied in design codes for the verification of structures, and requires previous knowledge about the basic random variables. This research does not involve Level 1 method.

2.1.2. Level II Methods (fully probabilistic with approximations

Level II methods take all the probabilistic properties of the random variables into account. However, they include approximations that at the same time could also be severe limitations for their use in specific problems. The examples of level II reliability methods are:

- First Order Reliability Method (FORM)
- Second Order Reliability Method (SORM)
- First Order Second Moment (FOSM)

2.1.3. Level III Methods (fully probabilistic)

Level III methods are characterized as fully probabilistic and exact methods (no simplifying assumptions are implied). The accuracy of these methods can usually be controlled by parameters like the variance of the resulting failure probability or step sizes which also have an impact on the calculation time. The example of level III reliability methods are:

- Monte Carlo Simulation (MCS)
- Subset Simulation (SS)
- Directional Sampling (DS)
- Adaptive Directional Sampling (ADS)
- Important Sampling (IS)
- Adaptive Importance Sampling (AIS)
- Latin Hypercube Sampling (LHS)
- Active Learning Kriging-based Monte Carlo Simulation (AK-MCS)
- Numerical Integration (NI)
- Numerical Bisection (NB)

2.1.4. Level IV Methods

Level IV methods include more aspects into consideration (e.g. economical aspect). This report does not include level IV methods into application.

2.2. Monte Carlo Simulation (MCS)

A normal Monte Carlo Simulation (MCS), also known as ”Crude Monte Carlo”, is a sampling-based simulation that obtain results by using repeated random realizations on the limit state function. The underlying concept is to use randomness to solve problems that might be deterministic in principle [27]. Each realization will give a failure ($g(X)$ < 0) or a success ($g(X) \geq$ 0). Thus the P_f is calculated by a straightforward method as shown in equation (2.3).

$$P_f = \frac{N_{failure}}{N} \tag{2.3}$$

where $N_{failure}$ is the total number of failure realizations and N is the total number of realizations. The MCS method is often considered as the most robust method since it put ”all possibilities” into account, however, this leads to a time-consuming simulation and not very efficient to be applied in computer. Therefore, it is considered as inefficient and expensive. Other reliability methods are developed mainly to tackle this problem.

2.3. First Order Reliability Method (FORM).

The FORM method searches the design point directly by estimating the shortest distance (β) between the origin to the standardized limit state function $g(u)$ in the Standard Normal Space. A straightforward depiction of FORM can be seen in Figure 2.1. FORM tries to recreate the limit state function $g(u)$ with a linear line $L(u)$ (thus it is called ”first-order”) by using Taylor series approximation. Furthermore, the closest distance between $L(u)$ and the origin is then measured and is defined as β.

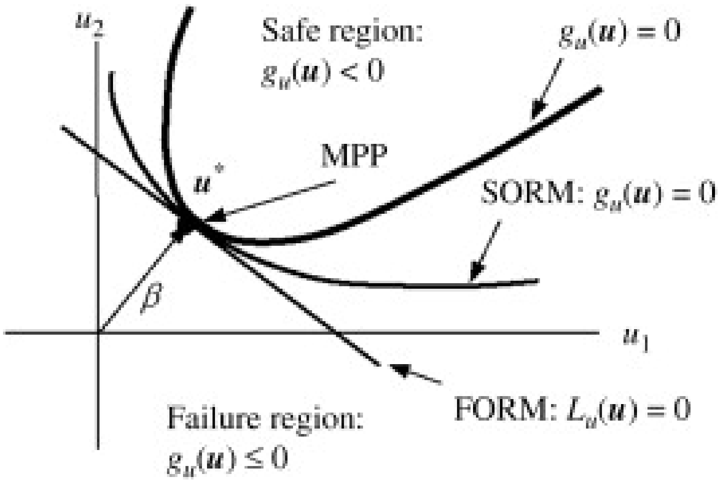


Figure 2.1: FORM and SORM reliability method (Chang 2015 [5]).

For the OpenTurn's FORM approach, the original Hasofer-Lind method is implemented in determining the shortest distance (β). Another FORM method implemented in this research is the Rackwitz-Fiesler iterative algorithm (FORM (DIY)) (Rackwitz et al. 1978 [24]). Although Figure 2.1 only shows a problem with 2 independent variables, FORM procedure and algorithm can also be implemented to a limit state function with more than 2 variables. Moreover, it is very difficult (if not impossible) to depict a FORM problem with more than 2 independent variables. However, the basic idea and mathematical implementation remains the same. Moreover, in this research, the starting point of the FORM method is set from the independent variables' mean values (or the origin of the standard normal space)

The main advantages of FORM is that it takes much less computational effort compared to a Monte Carlo Simulation. However, the main disadvantage is in its inaccuracy in estimating a really complex limit state function, especially when it has more than 1 design point (or more than 1 mode of failure).

2.4. Second Order Reliability Method (SORM).

The main idea is similar to FORM, however, SORM recreate the limit state function $g(u)$ with a curved line (thus it is called "second-order") by using a further truncation of the Taylor series approximation. A straight-forward scheme of SORM can also be observed in Figure 2.1. The main advantage is the limit state function can be more accurately recreated since it is imitated by a curved line, as opposed to FORM's linear approach. Moreover, OpenTurns' SORM that is performed in this research is using Breitung method (OpenTURNS 2021 [20]).

2.5. First Order Second Moment (FOSM).

Similar to FORM method, SORM is based on the first order Taylor series approximation of a performance function that is linearized at the mean values of the random variables by only using the second moment statistics (mean and variance) of the random variables. Therefore it is called first order "second moment". However, the main disadvantages of FOSM are it does not consider the information regarding the distribution of the independent variables and gives significant error in the truncation if the limit state function is non-linear. Despite many improvement of FOSM in recent years, a more detailed explanation of FOSM can be seen in the original FOSM formulation proposed by Cornell (Cornell 1967 [6]).

2.6. Subset Simulation (SS)

The basic idea of Subset Simulation method is to express the failure probability as a product of larger conditional failure probabilities by introducing intermediate failure events. With a proper choice of the conditional events, the conditional failure probabilities can be made sufficiently large so that they can be estimated by means of simulation with a small number of samples (Au 2001 [4]).

In simpler terms, subset simulation divide the parameter space within the problem domain into a several required simulation steps, and perform a Metropolis-based [2] Markov Chain Monte Carlo for each step to determine the domain of the next step. The probability of failure can be defined as [4]:

$$P_f = P(F_i)\, \Pi_{i=1}^{m-1} P(F_{i+1}|F_i) \tag{2.4}$$

Where m is the total number of required step until the failure event of interest has been reached. To compute P_f on equation 2.4, one needs to compute the probability of $P(F_1)$, $\{P(F_{i+1}|F_i) : i = 1, ..., m-1\}$.. $P(F_1)$ can be readily estimated by MCS as [4]:

$$P(F_1) \approx \widetilde{P_1} = \frac{1}{N} \sum_{k=1}^{N} I_{F_1}(\theta_k) \tag{2.5}$$

Where $\{\theta_k : k = 1, ..., N\}$ are independent and identically distributed samples according to PDF. Moreover, by applying Metropolis-based Markov Chain Monte Carlo, $P(F_{i+1}|F_i)$ can be defined as [4]:

$$P(F_{i+1}|F_i) \approx \widetilde{P_{i+1}} = \frac{1}{N} \sum_{k=1}^{N} I_{Fi+1}(\theta_k^{(i)}) \tag{2.6}$$

Finally, combining Equation 2.4, 2.5, and 2.6, the failure probability estimator is:

$$\widetilde{P_F} = \Pi_{i=1}^{m} \widetilde{P_i} \tag{2.7}$$

More detailed explanations of subset simulation method can be found in (Au 2001) [4].

2.7. Directional Sampling (DS)

Directional simulation reduces the dimension of the limit state probability integral by identifying a set of directions for integration, integrating either in closed-form or by approximation in those directions, and estimating the probability as a weighted average of the directional integrals. Most existing methods identify these directions by a set of points distributed on the unit hypersphere. The accuracy of the directional simulation depends on how the points are identified. When the limit state is highly nonlinear, or the inherent failure probability is small, a very large number of points may be required, and the method can become inefficient (Nie et al. 2000 [16]). The directional simulation method involves generating uniformly distributed direction vectors and performing a one-dimensional integration along each direction (Nie et al. 2000 [16]). A brief depiction of directional sampling is displayed in figure 2.2.

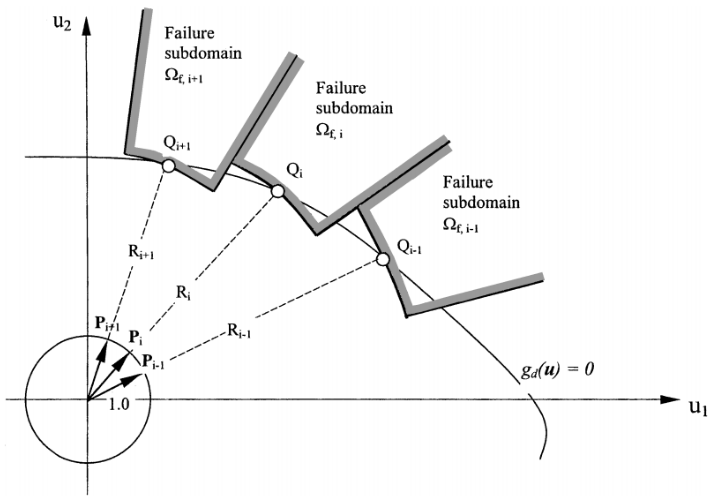


Figure 2.2: Spherical segments approximation of a limit state G(u)=0 by directional sampling (Nie et al. 2000 [16])

More details about directional sampling can be found in (Nie et al. 2000 [16]) and (Merchelrs et al. 2018 [18]).

2.8. Adaptive Directional Sampling (ADS)

Adaptive directional sampling is an improvement of directional sampling. The algorithm in ADS is based on a directional simulation scheme in which the most important directions are sampled more exact by means of a response surface approach (Grooteman 2010 [13]). The improvement is aimed to efficiently determine the optimal β-sphere that is excluded from the sampling domain, and drastically reduce the required number of simulations compared to the crude Monte Carlo method (Grooteman 2008 [12]). More detailed explanation regarding ADS can be found in (Grooteman 2010 [13]), (Grooteman 2008 [12]) and (OpenTurns 2021 [22]).

2.9. Importance Sampling (IS)

Importance Sampling is a technique to improve the Monte Carlo method for probability integration (Melchers 1989 [17]). The difference compared to the normal Monte Carlo method is that the realization are selected in a smarter way, preferably in the area on the edge of failure and non failure (Probabilistic Toolkit 2021 [8]). The main idea of Importance Sampling is choose a distribution which "encourages" the important results, which are the samples that lead to failure (hence it is called "importance" sampling). Furthermore, this new "biased" distribution will be weighted to correct the bias, as opposed to the Normal Monte Carlo where each realization weight the same. A general depiction of Importance Sampling can be observed in figure 2.3.

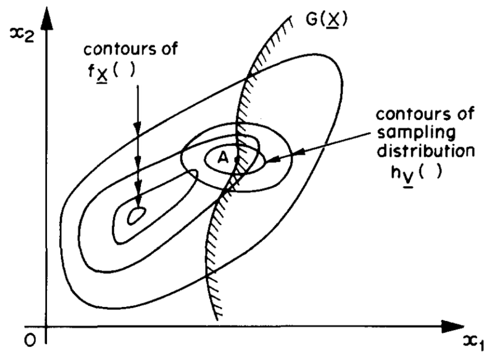


Figure 2.3: A general depiction of Importance Sampling Method with two-variable problem (after Melchers 1989 [17]).

In Probabilistic Toolkit Software, each realization (u) is translated to importance sampling realization (u_{imp}) in the standard normal space. The translation for each variable is separately supported as equation 2.8.

$$u_{imp,var} = f_{var} \cdot u_{var} + s_{var} \tag{2.8}$$

A correction is applied in the calculation of failure to incorporate the translation by giving a realization a weight, which is calculated as equation 2.9 (the multiplication with f_{var} is performed to compensate for the dimensionality).

$$w_{var} = \frac{f_{var} \cdot \Phi(u_{imp,var})}{\Phi(u_{var})} \tag{2.9}$$

and

$$W_{realization} = \Pi_{variables} w_{var} \tag{2.10}$$

where:

f_{var}	: variance factor per variable.
s_{var}	: variance shift per variable.
Φ	: the standard normal probability density function.
w_{var}	: the weight factor per variable per realization.
$W_{realization}$	: the weight factor of the realization.

Therefore, corresponding to equation 2.3, the probability of failure for Importance Sampling is defined in equation 2.11.

$$P_f = \frac{\sum_{failing\ realizations} W}{\sum_{all\ realizations} W} \tag{2.11}$$

Further explanation about Importance Sampling can be found in (Melchers 1989 [17]).

2.10. Adaptive Importance Sampling (AIS)

An adaptive importance sampling methodology is proposed to compute the multidimensional integrals encountered in reliability analysis. It is based on an improved Markov simulation algorithm (Metropolis et al. 1953 [2]). In the proposed methodology, samples are simulated as the states of a Markov chain and are distributed asymptotically according to the optimal importance sampling density. These Markov chain samples are then used to construct a kernel sampling density to provide a good approximation to the optimal importance sampling density. The Markov chain samples populate the region of higher probability density in the failure region and so the kernel sampling density approximates the optimal importance sampling density for a large variety of shapes of the failure region (hence it is called ”adaptive” importance sampling). An elaborated explanation of the method can be found in (Au 1999 [3]).

2.11. Latin Hypercube Sampling (LHS)

Latin Hypercube Sampling is a sampling method that enables better cover of the domain of the input variables variations due to a stratified sampling strategy. The sampling procedure is based on dividing the range of each variable into several intervals of equal probability. The sampling is undertaken as the followings (OpenTurns 2021 [21]):

- Step 1: The range of each input variable is stratified into isoprobabilistic cells.
- Step 2: A cell is uniformly chosen among all the available cells.
- Step 3: The random number is obtained by inverting the Cumulative Density Function locally in the chosen cell.
- Step 4: All the cells having a common strate with the previous cell are put apart from the list of available cells.

The probability of failure (P_f) is estimated by:

$$P_{f,LHS}^N = \frac{1}{N}\,\Sigma_{i=1}^N 1_{(g(X^i,d)\leq 0)} \tag{2.12}$$

Where the sample of $\{X^i, i = 1, ..., N\}$ is obtained as described previously. Further explanation regarding LHS can be seen in (Helton et al. 2002 [14]).

2.12. Active Learning Kriging-based Monte Carlo Simulation (AK-MCS)

AK-MCS is one of a Kriging-based metamodelling method. Metamodelling (surrogate modelling) relies in constructing models that acts as surrogates of complex problem (Teixeira et al. 2021 [26]). In their most fundamental form, metamodels are easily understood as black-box functions that relate an input variable x to an output $Y(x)$, allowing cheap evaluation of $Y(x)$ at any input value x, as can be seen in Figure 2.4 (Teixeira et al. 2021 [26]).

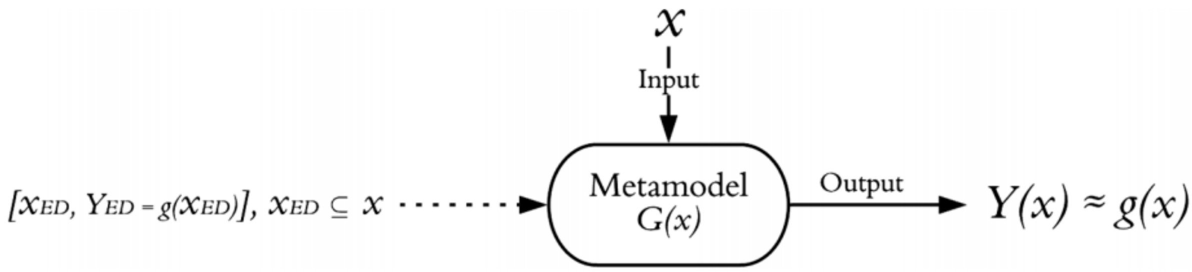


Figure 2.4: Generic description of a metamodel as a black-box function defined on a support ED (experiment design) [26]
.

In this research, the applied AK-MCS method is after (Echard et al. 2011 [10]). It consists of an active-learning reliability method combining Kriging and Monte Carlo Simulation. Kriging is based on the idea that the performance function $G(x)$ is seen as the realization of a stochastic field $g(x)$. However, instead of approximating the limit state in the whole space and therefore making expensive evaluations of limit state G for points with very weak densities of probabilities, here it is preferred to focus on a Monte Carlo population. AK-MCS classifies a Monte Carlo population of N_{MC} points without evaluating N_{MC} times the performance function. To avoid the evaluations, Kriging is used for its exact interpolation characteristic and for its interesting proprieties in active learning methods. Moreover, this research applies AK-MCS with simple kriging (0 order, AK-MCS 0) and universal kriging with linear basis function (1 order, AK-MCS 1). Furthermore, the maximum realization number for each simulation is set to 100 for both AK-MCS 0 and AK-MCS 1.

Due to its complex nature, a more detailed explanation regarding the implemented AK-MCS can be obtained from (Echard et al. 2011 [10]).

2.13. Numerical Integration (NI)

It is the most time consuming, but most exact way to calculate the failure probability. A step size, minimum, and maximum values of the input variables are needed for the evaluation. The minimum and maximum value for which the integration will run, are defined in the u-space. The numerical integration will fill up the left over space between these values and -8 and 8 by additional cells, so the whole integration domain will be covered always (Probabilistic Toolkit 2021 [8]). An example of numerical integration can be observed in Figure 2.5, which is made based on a simple slope stability with 2 modes of failure as discussed in section 3.4.

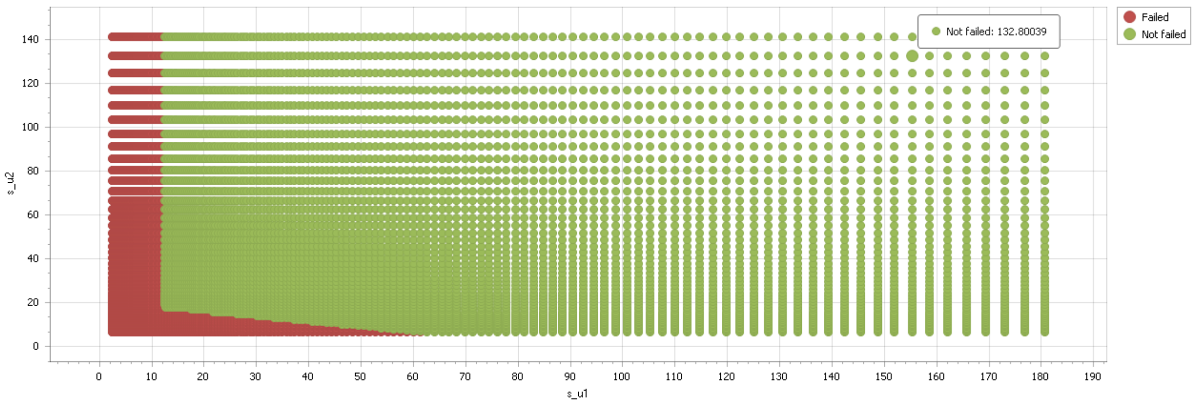


Figure 2.5: A numerical integration application on a simple slope stability problem.

Due to the method's nature, the method is currently suitable for a problem with 2 independent variables (applied using Probabilistic Toolkit software) and is not including in the ranking system (section 4). Therefore, the method is only applied to the simple slope stability problem which is discussed in section 3.4.

2.14. Numerical Bisection (NB)

Numerical bisection is similar to numerical integration, however, the integration cells are generated by bisection of the complete domain. As soon as an integration cell has same qualitative results (fail, not fail, not counting) on its corner points, we assume that all values inside the cell would deliver the same qualitative result and no points will be calculated within the call nor on its sides. As long as cells exist with different qualitative results, the cell will be split and new corner points will be calculated. This process ends when the cells, which don't have a similar result, represent a probability lower than an accepted allowed difference in the reliability (Probabilistic Toolkit 2021 [8]). An example of numerical bisection application can be seen in Figure 2.6, where it is applied to a simple slope stability with 2 modes of failure as discussed in section 3.4.

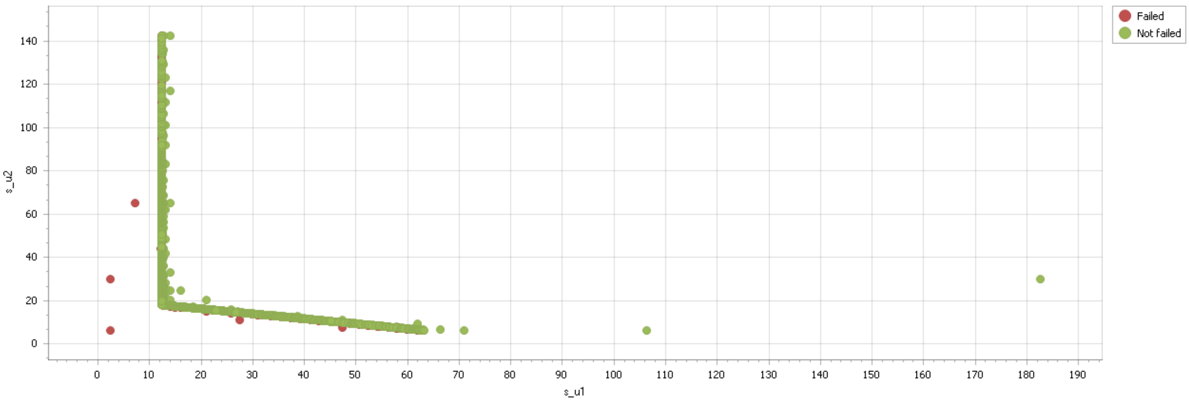


Figure 2.6: A numerical bisection application on a simple slope stability problem.

Due to the same reasoning as numerical integration, numerical bisection is only applied to the simple slope stability problem discussed in section 3.4 and is not included in the ranking system of section 4.

2.15. Convergence Criterion

Convergence criterion (among others) is needed to determine when to stop the reliability analysis process. It determines the desired level of accuracy. The convergence criterion is reached when the probability of failure P_f variation coefficient (COV) reaches a certain limit. For example, the COV of the MCS method is defined as equation 2.13

$$COV = \begin{cases} P_f < \frac{1}{2} & \frac{\sigma_p}{P_f} = z\sqrt{\frac{1-P_f}{NP_f}} \\ P_f \geq \frac{1}{2} & \frac{\sigma_p}{1-P_f} = z\sqrt{\frac{P_f}{N(1-P_f)}} \end{cases} \tag{2.13}$$

where:

σ_p : standard deviation of P_f.
z : quantile of the standard normal distribution corresponding with confidence level (z = 1 PTK).
N : The total number of realizations

The other sampling-based method follows similar rules in determining convergence criterion (depending on the COV of P_f), where more detailed explanation can be seen on each method's forementioned references. In this research, a minimum convergence of 0.1 is chosen. Therefore, for each method (especially for sampling-based methods), the reliability analysis process is stopped when the current step obtained reliability COV reach 0.1 (or the determined maximum N is reached). When this convergence is not obtained, then the minimum sample/ step will be increased.

However for gradient-based reliability method like FORM, convergence is reached when the difference between u_i (the point at the last ith iteration) is close enough to the predicted value of u_pred (where both u_i and u_pred are from standard normal space). In this research, this difference is limited to 10^{-2} for PTK methods, and 10^{-3} to OT and DIY methods.

An example of convergence history can be observed in figure 2.7, which is taken from a Monte Carlo Simulation of a geotechnical engineering problem.

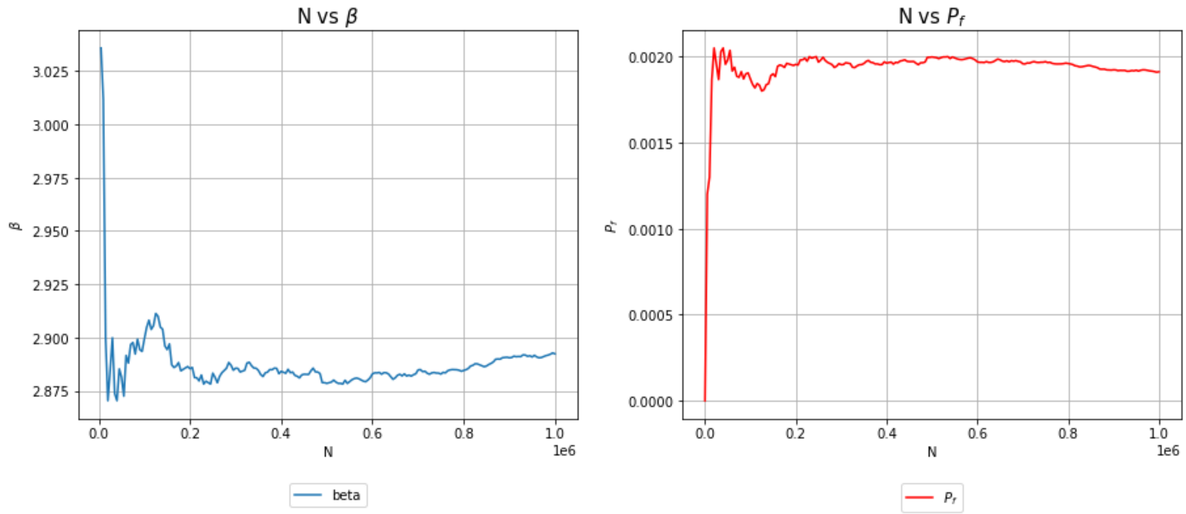


Figure 2.7: N vs β and Pf (Monte Carlo Simulation).

Based on Figure 2.7, it can be observed that β and P_f value fluctuate (and converging) throughout the increasing number of realization N. Moreover, the convergence of the P_f COV can also be displayed as in Figure 2.8.

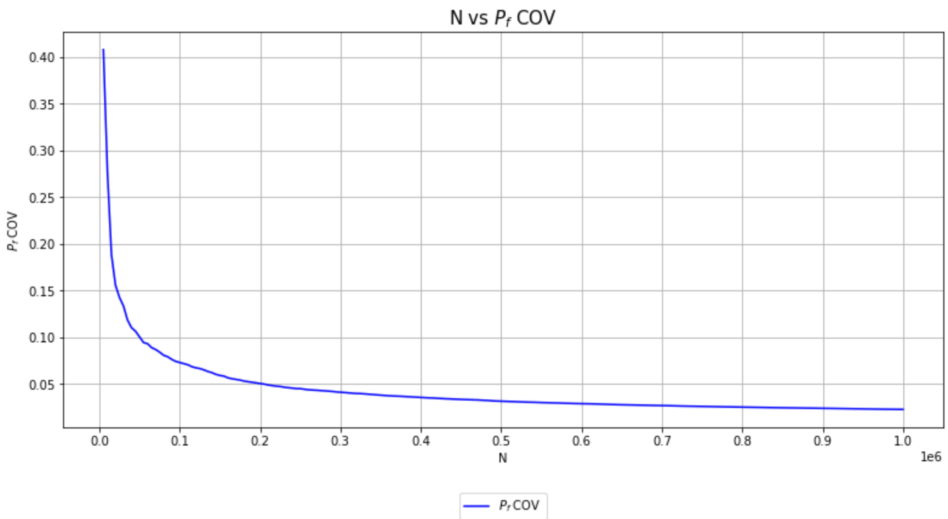


Figure 2.8: P_f COV of an MCS example.

It can be observed from Figure 2.8 that after N > 0.05 x 10^6, the P_f COV reaches values lower than 0.1. Moreover, after approximately N > 0.7 x 10^6, P_f COV is reaching a constant value. Therefore, the analysis will be concluded once the P_f COV has reach the tolerated value (0.1). However, for MCS (OT) and MCS (DIY), the simulation will be performed until N reaches approximately 10^6 (like displayed in Figure 2.8) since it will be used as references. Consequently, for Monte Carlo Analysis (and other sampling-based methods), a sufficient value of N (or step) is needed to get the required convergence criterion (P_f COV = 0.1).

3
Case Analysis

This section will discuss the types of problem analyzed in this research. The results will be discussed in Section 4.

3.1. Geotechnical engineering problem to be analyzed

Geotechnical Safety Network (GEOSNet) is an international open collaborative platform that serves to promote, coordinate, and support activities relating to geotechnical safety (GEOSNet 2021 [11]). This research analyze 9 different geotechnical engineering problems with various modes of failure (MoF) and number of independent variables, 8 of which are taken from GEOSNet reliability benchmarking examples (GEOSNet 2021 [19]). The details regarding the problems could be found in Appendix B. These geotechnical engineering problems can be summarized into table 3.1.

Table 3.1: Geotechnical engineering problems to be analyzed in this research.

Problems	No. of Mode of Failure	No. of variables	P_f references	β references
A Simple Slope Stability	2	2	2.44E-02	1.97
GEOSNet EX1	1	6	5.73E-02	1.58
GEOSNet EX2	2	9	6.39E-02	1.52
GEOSNet EX3	3	4	0.18E-02	2.91
GEOSNet EX4	1	5	0.39E-02	2.66
GEOSNet EX5	1	7	9.44E-02	1.32
GEOSNet EX7	1	4	2.63E-02	1.94
GEOSNet EX8	3	12	8.17E-02	1.39
GEOSNet EX9	1	9	0.08E-02	3.17

Due to high uncertainties and data inconsistencies, MCS (OT) P_f value will be used as the reference for the simple slope stability, GEOSNet example 5, 7, and 8 problems. Moreover, the rest of the problem use GEOSNet's MCS result. Both GEOSNet and OpenTurn's MCS reference results are obtained with $N = 10^6$.

3.2. Model Fluctuations

In sampling-based methods, it is almost impossible to obtain the exact same result for each simulation (unlike gradient-based method i.e. FORM). Sometimes the convergence criterion is not achieved when the maximum step/realization has been surpassed (like in the case of AK-MCS 0 order in this research). In that case, the resulting β of P_f values might differ considerably from the reference (or MCS) result. In this research, the particular problem often happens in the application of directional sampling (DS) and AK-MCS 0 order. Thus for these methods and the rest of the sampling-based methods, the simulations are performed 100 times therefore the obtained average β (or P_f) is more depended on realizations where the convergence criterion is achieved. These additional simulations are only applied to the sampling-based methods other than MCS (LHS, IS, DS, SS, ADS, AIS, and both AK-MCS).

The different β values obtained from 100 simulations of directional sampling and 0 order AK-MCS can be seen in Figure 3.1 and 3.2. For Figure 3.2, each simulation has a different realizations number (N) depending on whether it reaches convergence (or not) before the maximum realization number/step is achieved. Moreover, each simulation in Figure 3.1 is based on a maximum N of 100.

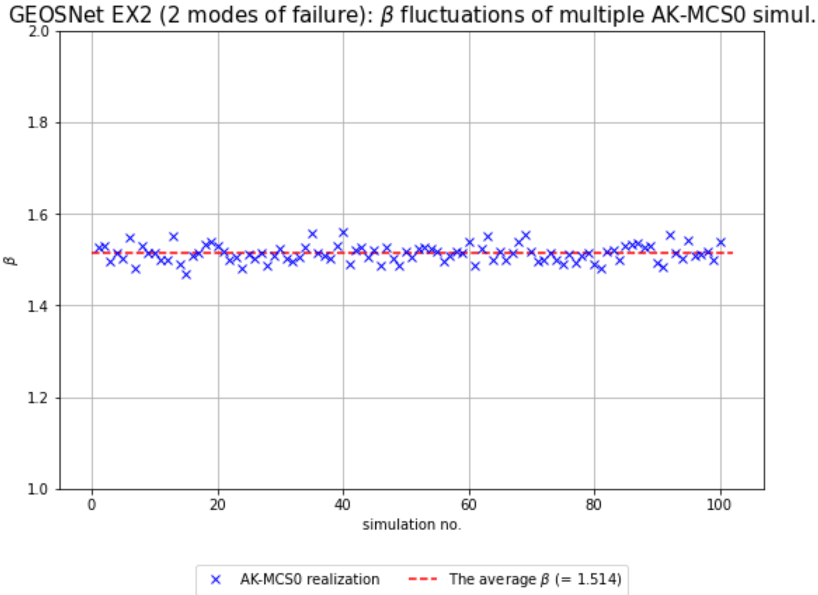


Figure 3.1: Different β value for each realization in AK-MCS 0.

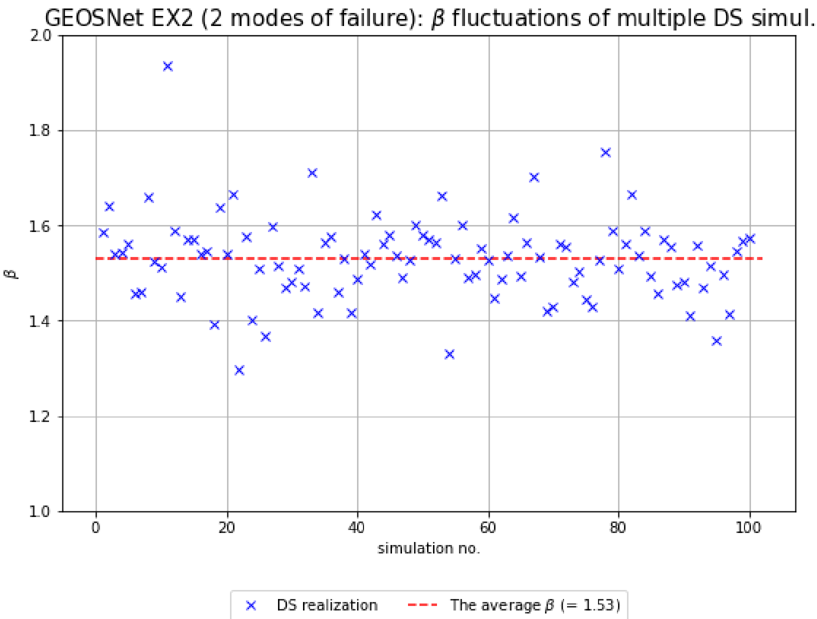


Figure 3.2: Different β value for each realization in DS.

It can be observed from Figure 3.1 and 3.2 that for the same problems, sample-based reliability methods could vary between each realization. These differences could be considerably noticeable despite the average of the realizations are considerably similar. This can be seen in Figure 3.2 where one of the realization gives a β value of more than 1.9 despite the β average is 1.53. Therefore, one must avoid performing only 1 simulation when dealing with a sampling-based method (note that each simulation has its own N value) since it could give a false sense of accuracy.

3.3. A problem with 1 mode of failure

The first example discussed here is the 1^{st} exercise of the GEOSNet example. The example simulates a simple infinite slope with six independent random variables. The problem can be depicted in Figure 3.3.

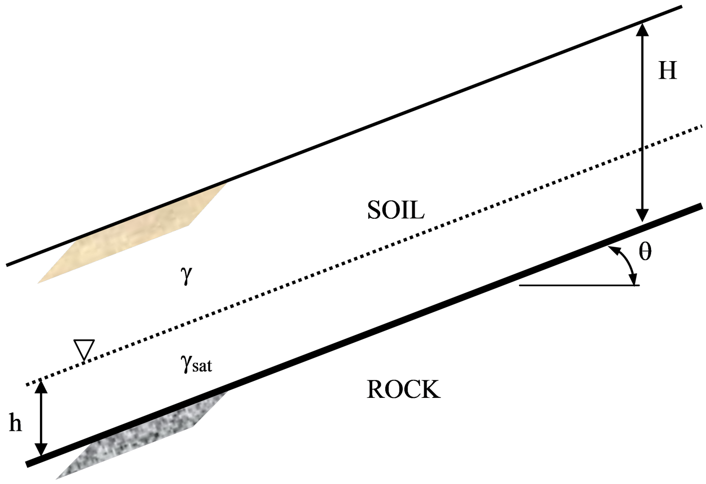


Figure 3.3: Example 1 scheme. (Source: GEOSNet, www.geoengineer.org).

The performance function for the problem is formulated in equation (3.1) below.

$$P = \frac{[\gamma(H-h) + h(\gamma_{sat} - \gamma_w)]cos\theta tan\phi}{[\gamma(H-h) + h\gamma_{sat}]sin\theta} - 1 \tag{3.1}$$

where:

$$\gamma = \gamma_w(G_s + 0.2e)/(1+e) \tag{3.2}$$

and

$$\gamma_{sat} = \gamma_w(G_s + e)/(1+e) \tag{3.3}$$

with:

H	= depth of soil above bedrock
h	= U x H, height of groundwater table above bedrock
U	= a factor of groundwater level determination
γ and γ_{sat}	= moist unit weight and saturated unit weight of the surficial soil, respectively
γ_w	= unit weight of water (9.81 kN/m^3)
ϕ	= effective stress internal friction angle
θ	= slope inclination

The moist and saturated soil unit weights are not independent, because they are related to the specific gravity of the soil solids (G_s) and the void ratio (e). The uncertainties in γ and γ_{sat} are characterized by modeling G_s and e as two independent uniform random variables. There are six independent random variables in this problem (H, U, ϕ, θ, e, and G_s). The distribution of the independent variables are displayed in table 3.2.

Table 3.2: GEOSNet Example 1 Result Comparison (GEOSNet result by Jianye Ching and Yi-Hung Hsieh).

Variable	Distribution	Statistics	Units
H	Uniform	[2, 8]	m
U	Uniform	[0, 1]	-
ϕ	Lognormal	μ = 35, COV = 8%	deg.
θ	Lognormal	μ = 20, COV = 5%	deg.
G_s	Uniform	[2.5, 2.7]	-
e	Uniform	[0.3, 0.6]	-

Furthermore, it can be seen from the performance function that there is only one mode of failure in this problem. The reliability index (β), probability of failure (Pf), number of realizations (N), P_f error, and absolute β error from each method is compared in Table 3.3 (β reference = 1.578 and P_f reference = 0.0573).

Table 3.3: GEOSNet example 1 results comparison with P_f reference = 0.0573 based on GEOSNet MCS result with N=10^6 by Jianye Ching and Yi-Hung Hsieh [23].

Methods	β	P_f	N	**Abs. β error**	**P_f error**
FORM (PTK)	1.427	0.0768	42	9.57%	34.0%
LHS (PTK)	1.578	0.0573	10001	0.00%	0.0%
MCS (PTK)	1.574	0.0578	163120	0.25%	0.9%
IS (PTK)	1.570	0.0583	37312	0.51%	1.7%
DS (PTK)	1.574	0.0577	40955	0.25%	0.7%
FORM (OT)	1.426	0.0770	144	9.66%	34.4%
SORM (OT)	1.544	0.0613	252	2.18%	7.0%
MCS (OT)	1.573	0.0579	1000000	0.31%	1.0%
SS (OT)	1.574	0.0578	20000	0.24%	0.8%
DS (OT)	1.551	0.0605	522.5	1.74%	5.6%
IS (OT)	1.562	0.0592	514.41	1.03%	3.3%
LHS (OT)	1.570	0.0582	6486.47	0.53%	1.7%
ADS (OT)	1.524	0.0638	39950.3	3.43%	11.3%
FORM (DIY)	1.426	0.0769	71	9.62%	34.2%
FOSM (DIY)	1.397	0.0812	13	11.46%	41.7%
MCS (DIY)	1.575	0.0576	900000	0.19%	0.6%
SS (DIY)	1.576	0.0575	4905	0.11%	0.3%
DS (DIY)	1.578	0.0573	369.2	0.01%	-0.1%
IS (DIY)	1.541	0.0617	286.5	2.35%	7.6%
AIS (DIY)	1.574	0.0577	907.8	0.24%	0.7%
AKMCS-0	1.584	0.0566	51	0.38%	-1.2%
AKMCS-1	1.532	0.0627	29	2.90%	9.5%

It can be seen from Table 3.3 that for the GEOSNet example 1 problem (with a simple performance function and only has one mode of failure), the most accurate reliability method is LHS (PTK) with P_f error of 0.0%, which mean it is giving the exact same P_f value as GEOSNet MCS' despite of only using 10001 realizations (N). Moreover, it can also be seen from Table 3.3 that FOSM (DIY) requires the least N while also giving the least accurate P_f value.

The result (β and P_f) of each method can be compared together into a single graph of N vs β (since P_f values are normally small, therefore it is more convenient to present it as β value, as will then be displayed in the result graphs of this report). A lower required N means a more efficient method, and a closer beta to the point of reference means a more accurate method. GEOSNet's Monte Carlo Result will be used as a reference, as the Monte Carlo Simulation used 10^6 realizations and is considered as the most robust/accurate result (in spite of the efficiency disadvantages). The comparison can be observed in Figure 3.4.

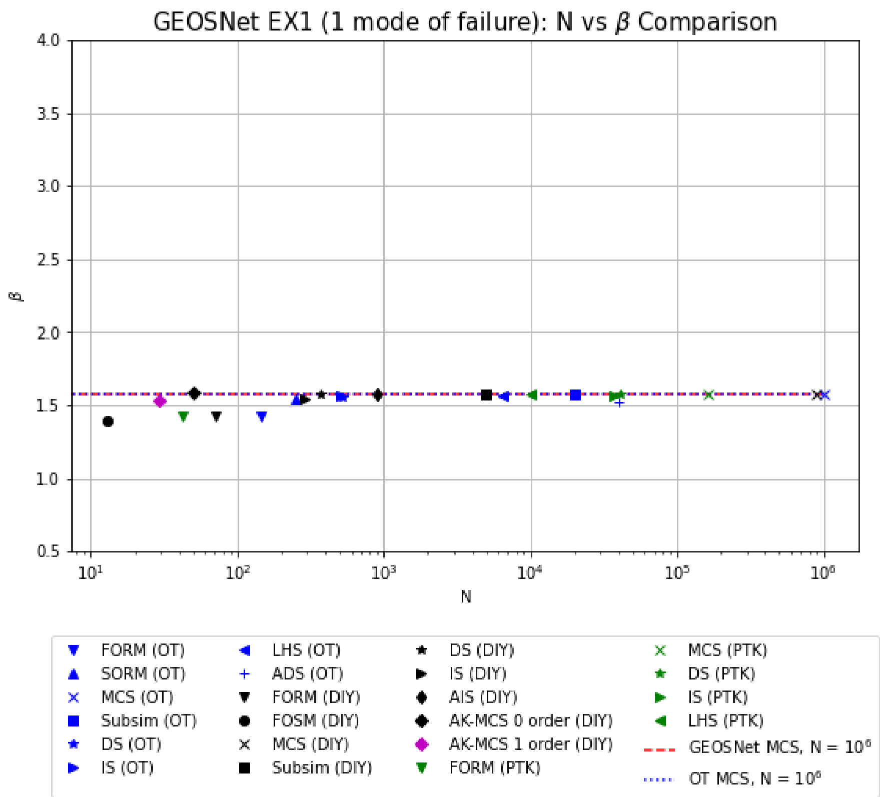


Figure 3.4: N vs Beta comparison of GEOSNet Example 1.

It can be observed that all of the results are close to the reference line, with the largest β deviation (FOSM (DIY) result) is less than 0.5 (or P_f error of 41.7%).

3.4. A problem with more than 1 mode of failure

In this example, a simple slope stability with two undrained layers is analyzed (Figure 3.5. The two layers have independent stochastic soil strength parameters $s_{u,1}$ and $s_{u,2}$ with all other variables are deterministic.

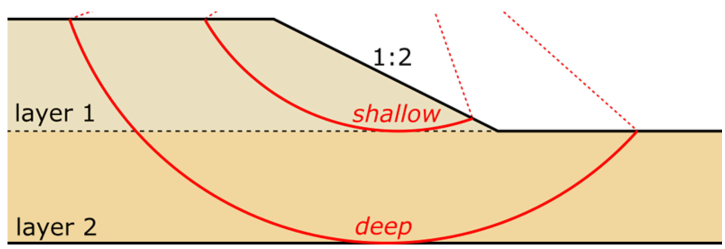


Figure 3.5: A simple slope stability problem with 2 independent stochastic variables.

The following deterministic parameters were used:

Slope heights	: 4 m
Slope inclination	: 1:2 (26.57 deg.)
Layer thickness	: 4 m ; 4 m
Unit weight	: 16 kN/m^3

By using Bishop's limit equilibrium method, the 2 modes of failure (either a deep or a shallow slip circle) can be found as a critical slip circle with the lowest factor of safety FS. The FS is depending on just the relative undrained shear strength of the 2 layers. The limit state function $g_{sys}(\mathbf{X})$ can thus be formulated as the combination of the component limit state functions $g_s(\mathbf{X})$ and $g_d(\mathbf{X})$ for shallow and deep failure modes respectively.

$$g_s(\mathbf{X}) = FS_s - 1 \tag{3.4}$$

$$g_d(\mathbf{X}) = FS_d - 1 \tag{3.5}$$

with:

$$FS_s = c_{s1}.s_{u,1} \tag{3.6}$$

$$FS_d = c_{d1}.s_{u,1} + c_{d2}.s_{u,2} \tag{3.7}$$

Therefore:

$$g_{sys}(\mathbf{X}) = min(g_s, g_d) \tag{3.8}$$

Where X_i is the stochastic variable corresponding to the undrained shear strength s_u for layer i. Using a limit equilibrium software with Bishop's method, the coefficient of the performance function above are calibrated as the following.

c_{s1} = 0.004072890235
c_{d1} = 0.000554854
c_{d2} = 0.002397665

Moreover, the independent variable distribution are displayed in Table 3.4.

Table 3.4: Parameter distribution of the simple slope stability problem.

Variable	Distribution	Statistics	Units
$s_{u,1}$	Lognormal	μ = 21.8, σ = 6.0	kN/m^2
$s_{u,2}$	Lognormal	μ = 30.4, σ = 6.0	kN/m^2

The limit state function can be plotted like in Figure 3.6.

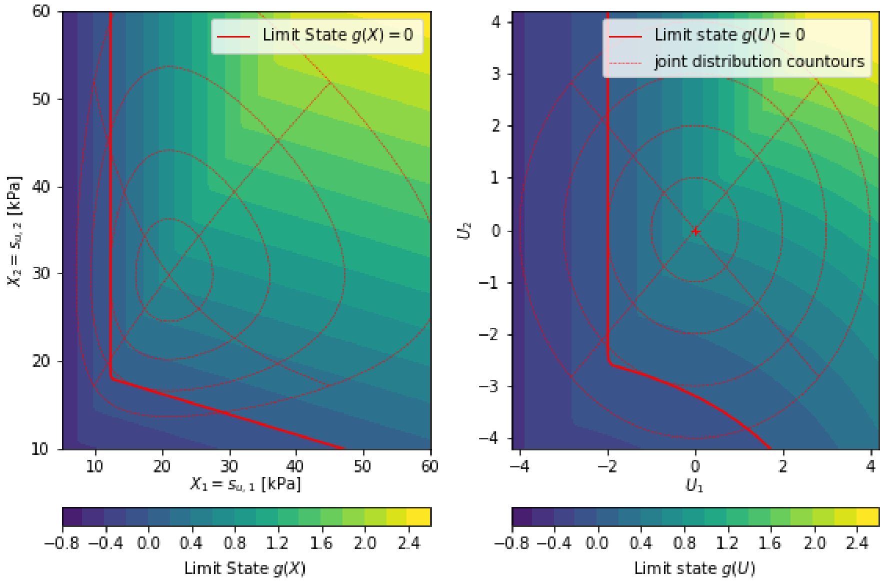


Figure 3.6: Limit state function of a slope stability problem with 2 mode of failulres.

From Figure 3.6, it can be seen that the limit state function at $g_{sys}(X)$ = 0 is consisted of 2 lines that are combined together. These lines represent the failure of shallow and deep sliding surface. These lines can be plotted separately as displayed in Figure 3.7 and 3.8 for shallow and deep sliding surfaces respectively.

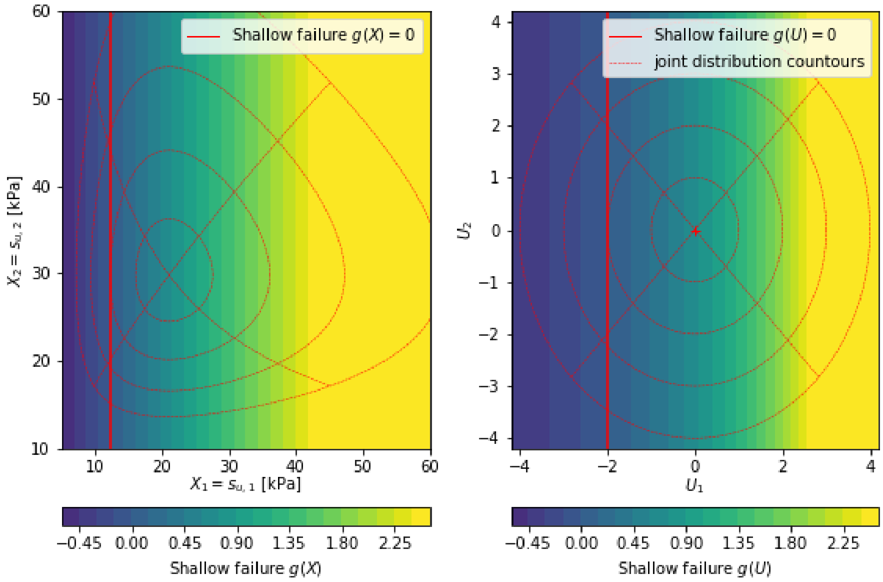


Figure 3.7: Limit state function of the shallow slip surface.

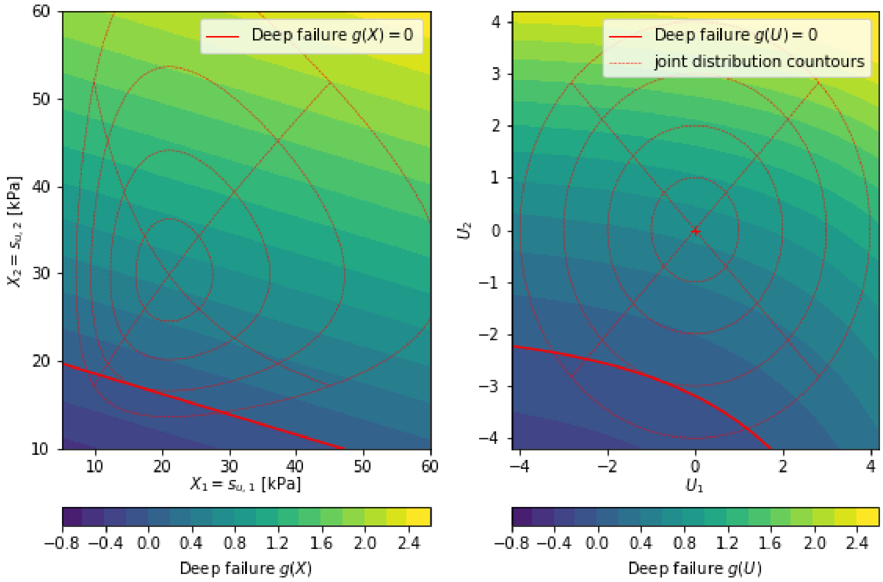


Figure 3.8: Limit state function of the deep slip surface.

The distance of the design points from the origin of the standard normal space as displayed in Figure 3.7 and 3.8 are certainly different between each other ($\beta_{shallow} = 1.97$ and $\beta_{deep} = 3.00$). Therefore, there are 2 possibilities of design point in $g_{sys}(X) = 0$, as displayed in Figure 3.9).

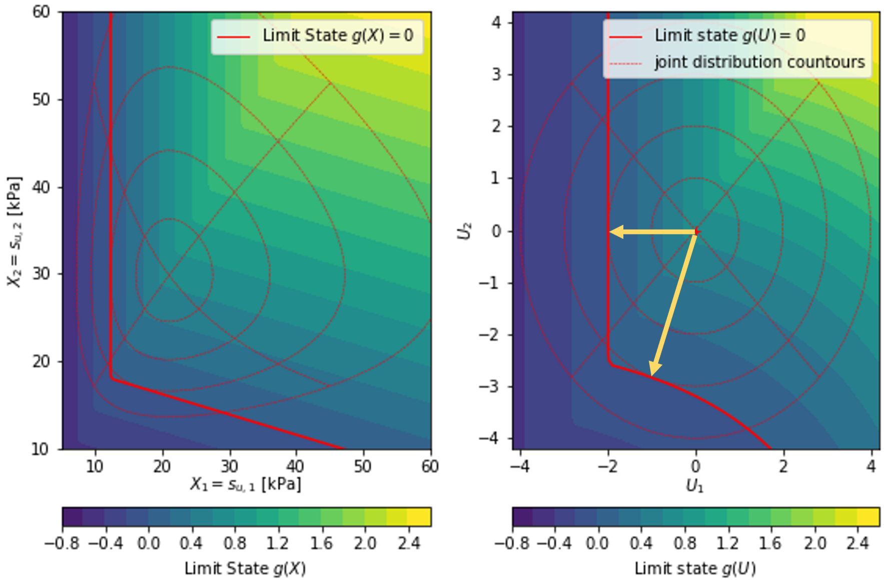


Figure 3.9: Two possibilities of β for $g_{sys}(X) = 0$.

This multiple design points can also be observed when analyzed using Monte Carlo Simulation (Figure 3.10).

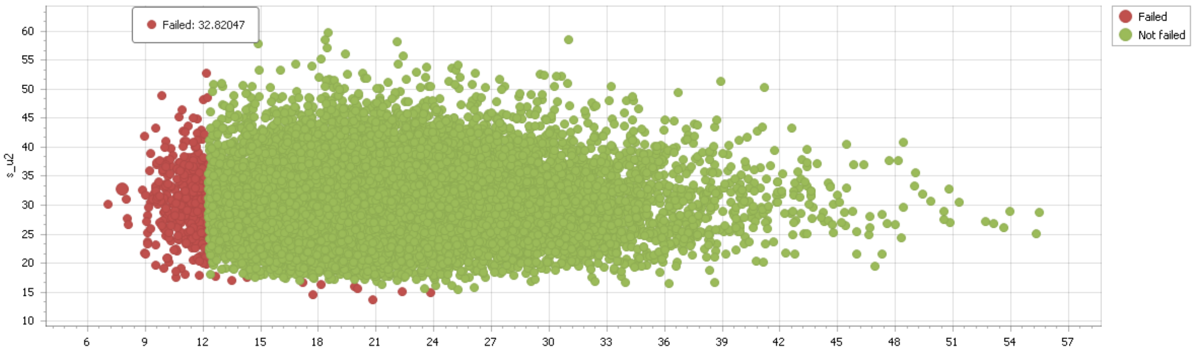


Figure 3.10: Two possible design points as observed from MCS (PTK) result.

Judging by the definition of reliability index (the shortest distance from the origin to the limit state function in a standard normal space), clearly the correct reliability index of $g_{sys}(X)$ should be $\beta = 1.97$, or the shortest distance between the origin to the limit state function of the shallow slip in standard normal space. However, when evaluated in a combined limit state function $g_{sys}(X)$, FORM (OT and or DIY version) gives a reliability index $\beta = 3.00$, which is clearly overestimating the reliability of $g_{sys}(X)$. Similarly, this also happens with the SORM (OT) method due to their similar approach in determining the reliability index (SORM (OT) gives $\beta = 3.04$). Despite a slightly more accurate re-creation of a design point in a limit state function's curve, SORM (OT) method still unable to detect the closest design point when the function has more than 1 design point.

Consequently, for a function with a multiple mode of failures (thus having multiple design points), it is advised to model each mode of failure separately (as applied in PTK analysis), and then analyze the reliability index of each design point. Furthermore, the system reliability can be estimated based on each of its design point's reliability, as defined in equation 3.9.

$$\beta_{sys} \approx \Phi^{-1}[\Phi(\beta_{shallow}) + \Phi(\beta_{deep})] = 1.97 \tag{3.9}$$

This problem can be avoided by using other reliability methods (e.g. sampling-based methods). The application of other reliability methods is presented and compared in Figure 3.11.

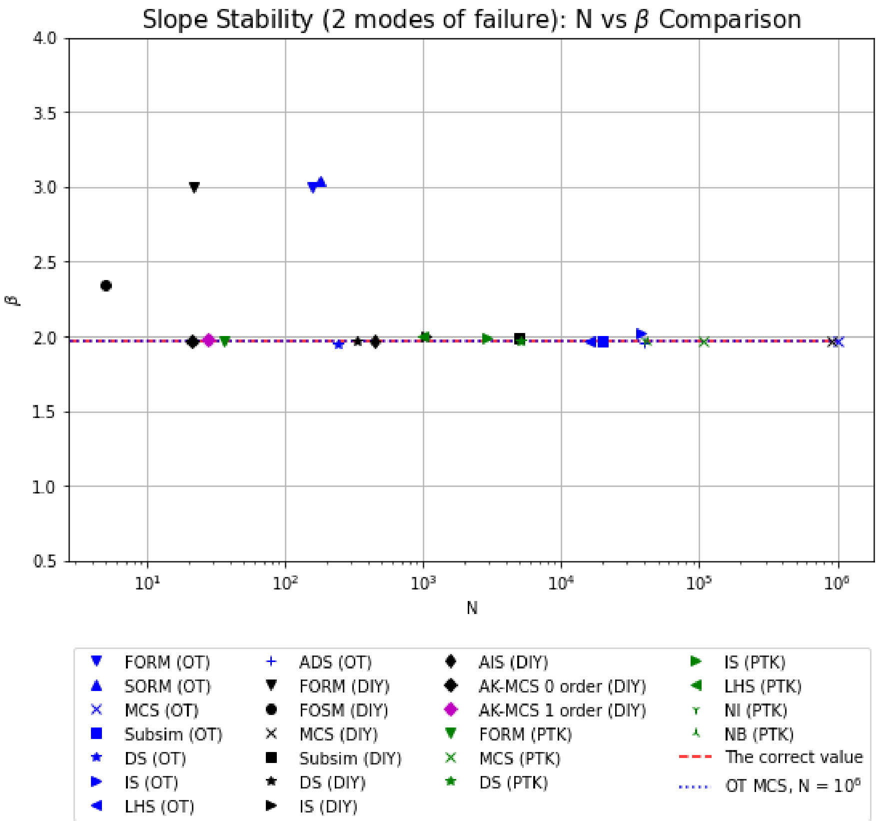


Figure 3.11: Result comparison of a simple slope stability problem with 2 design points.

Figure 3.11 shows that gradient-based reliability methods such as FORM (OT and DIY) and SORM (OT) method gives β = 3.00, while the sampling-based reliability methods approximately give β = 1.97 (which is the correct β value). Moreover, despite being a gradient-based reliability method, FORM (PTK) accurately predicts the correct β value by analyzing each mode of failure separately. Furthermore, if each design point is evaluated separately as displayed in Figure 3.7 and 3.8, all methods except for FOSM (DIY) give the same result as displayed in Figure 3.12 (for shallow slip surface) and Figure 3.13 (for deep slip surface).

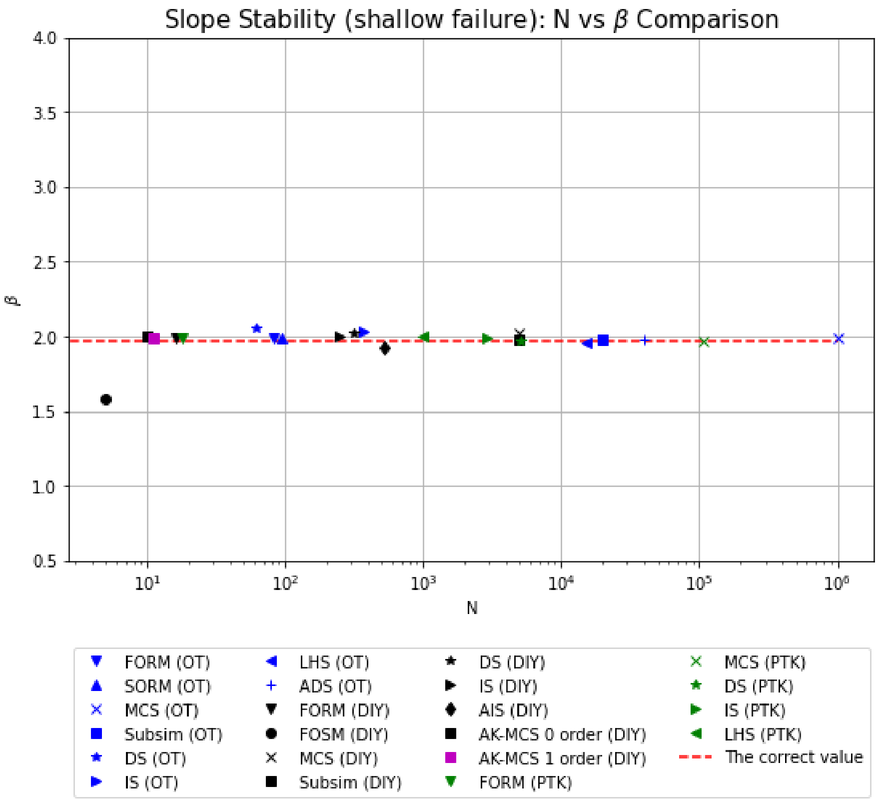


Figure 3.12: Result comparison of the shallow mode of failure.

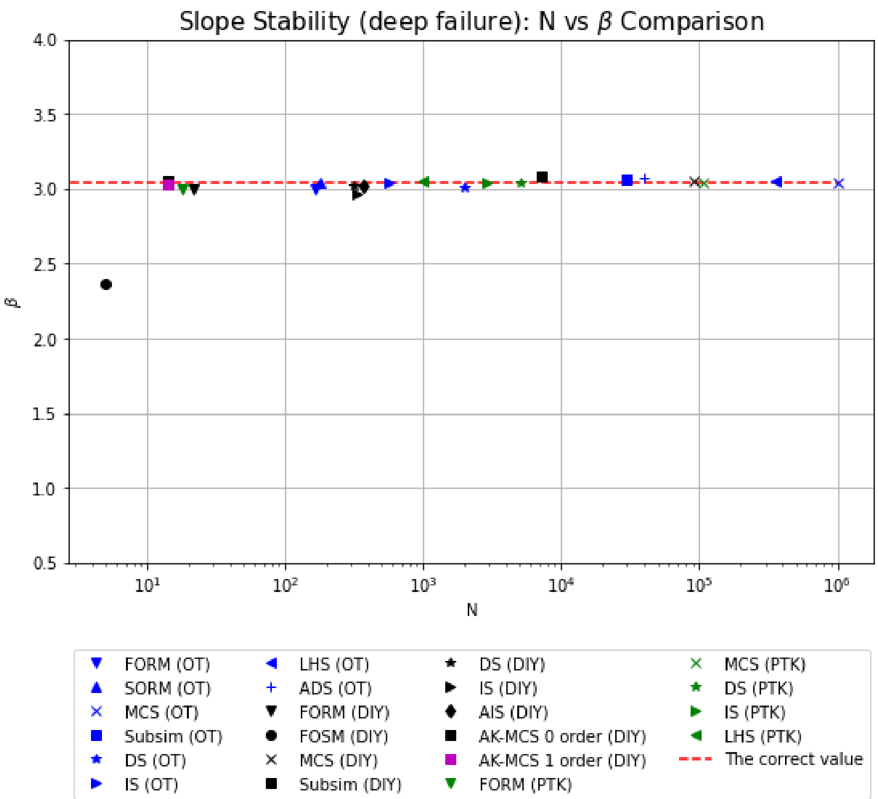


Figure 3.13: Result comparison of the deep mode of failure.

It can be observed from Figure 3.12 and 3.13 that all of the reliability methods (except FOSM) give the same result. Therefore in this case, it can be concluded that a gradient-based reliability method (in this case is FORM) could be misleading for a function with more than 1 failure of mode (or more than 1 design point). However, this problem can be avoided if the problem's modes of failure are evaluated separately.

Moreover, this problem can also be avoided by applying different initial/ starting point. Normally FORM starts from the mean values of each parameter or origin of the standard normal space (as shown in Figure 3.9). Therefore in this case, if the starting point is moved higher (constant X_1 and higher X_2), the obtained β value would be 1.98 (similar to the result displayed in Equation 3.9). This can be briefly explained by Figure 3.14.

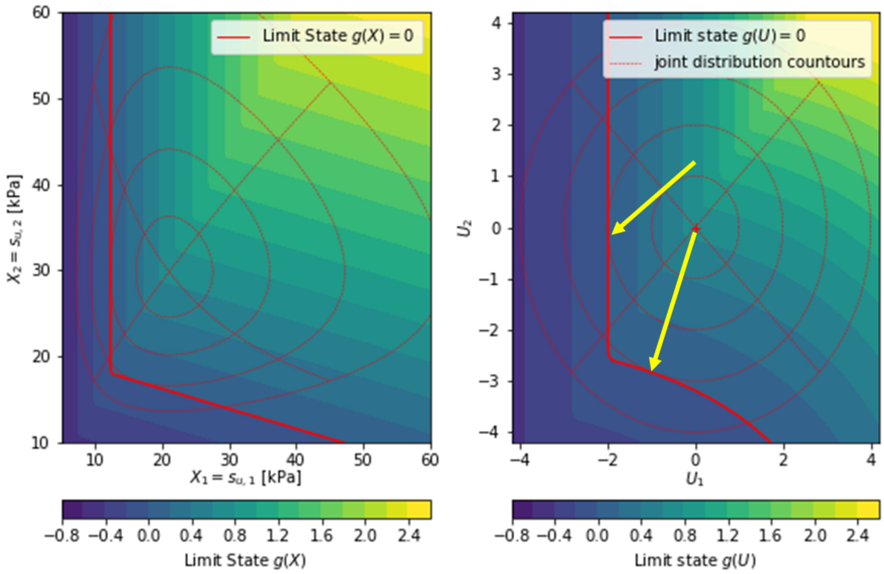


Figure 3.14: Two possible different FORM β values with different starting points.

3.5. Analysis Results

The results of multiple reliability analysis methods on a slope stability problem and GEOSNet examples are presented from Figure 3.15 to Figure 3.23 (note that Figure 3.11 and 3.4 are included again as Figure 3.15 and 3.16 respectively for practical purpose). The gradient-based methods applied are the basic type as discussed in Section 2.3, therefore their shortcomings in problems with multiple modes of failure can be better acknowledged (except for FORM (PTK) where each mode of failure is analyzed separately in each problem). Moreover, every sampling-based method for each problem in the following results is based on averaged result of 100 simulations (as mentioned in Section 3.2).

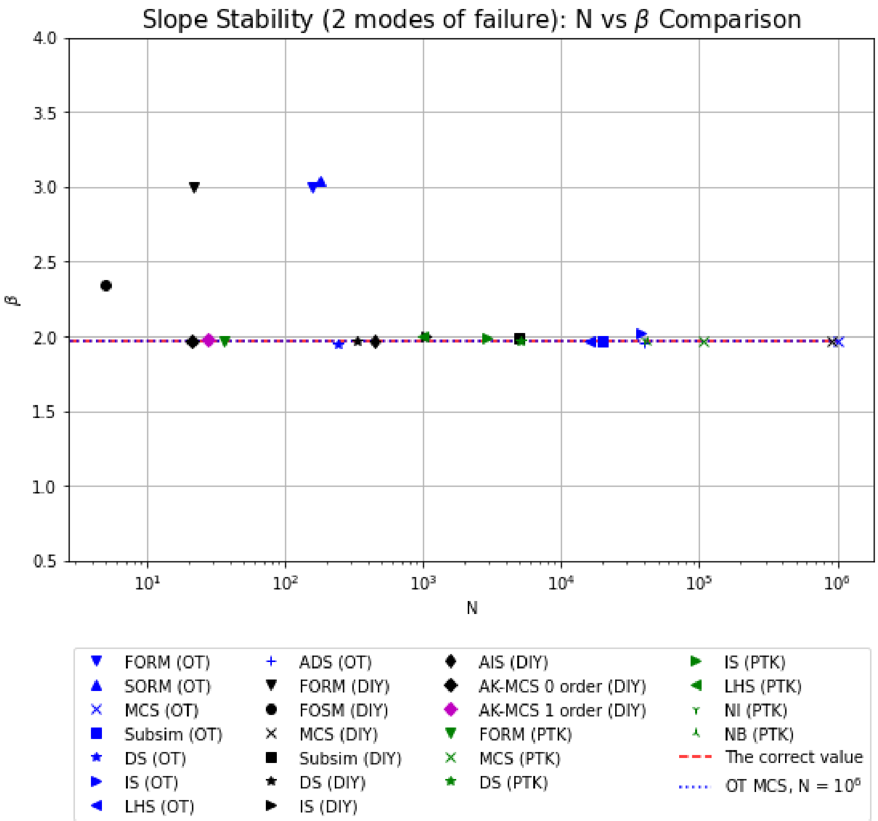


Figure 3.15: Result comparison of the slope stability problem (note that this graph is identical as Figure 3.11).

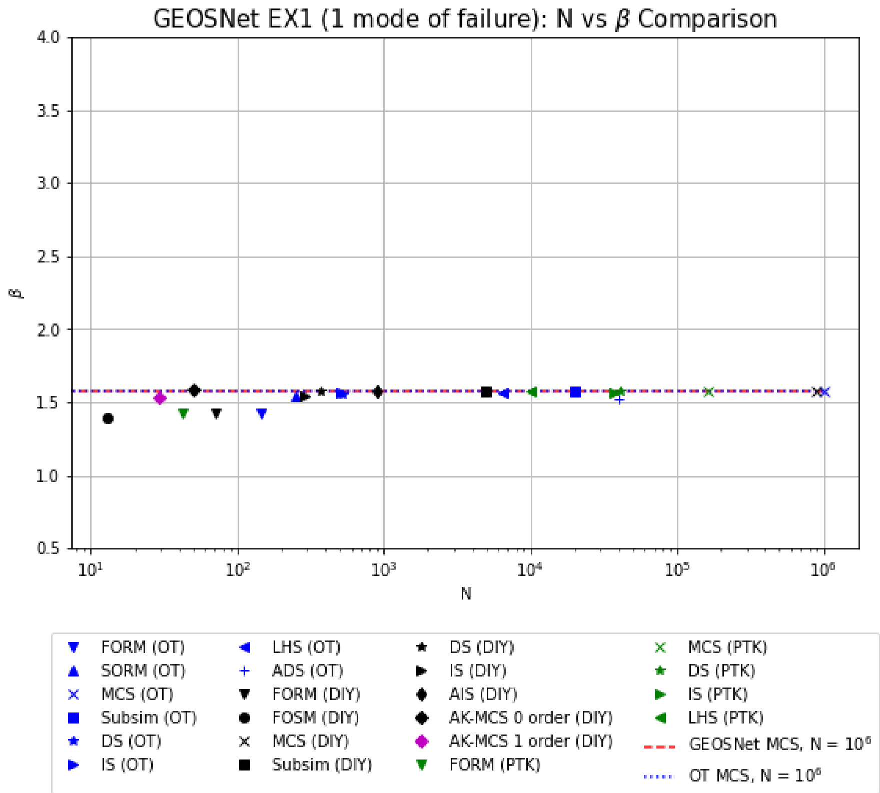


Figure 3.16: Result comparison of the GEOSNet example 1 (note that this graph is identical with Figure 3.4)

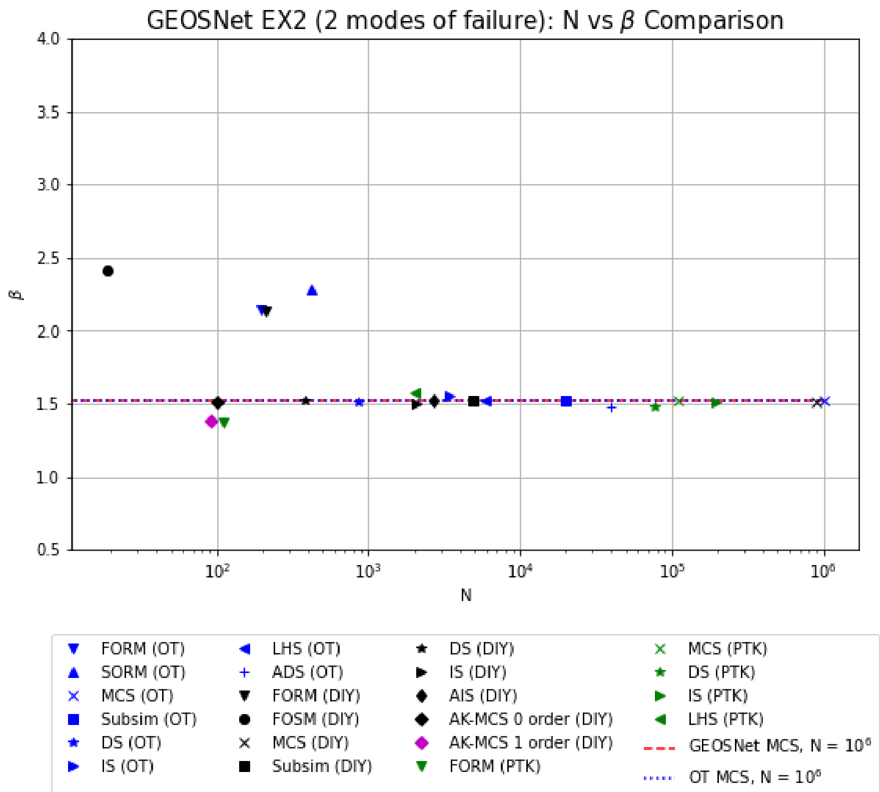


Figure 3.17: Result comparison of the GEOSNet example 2

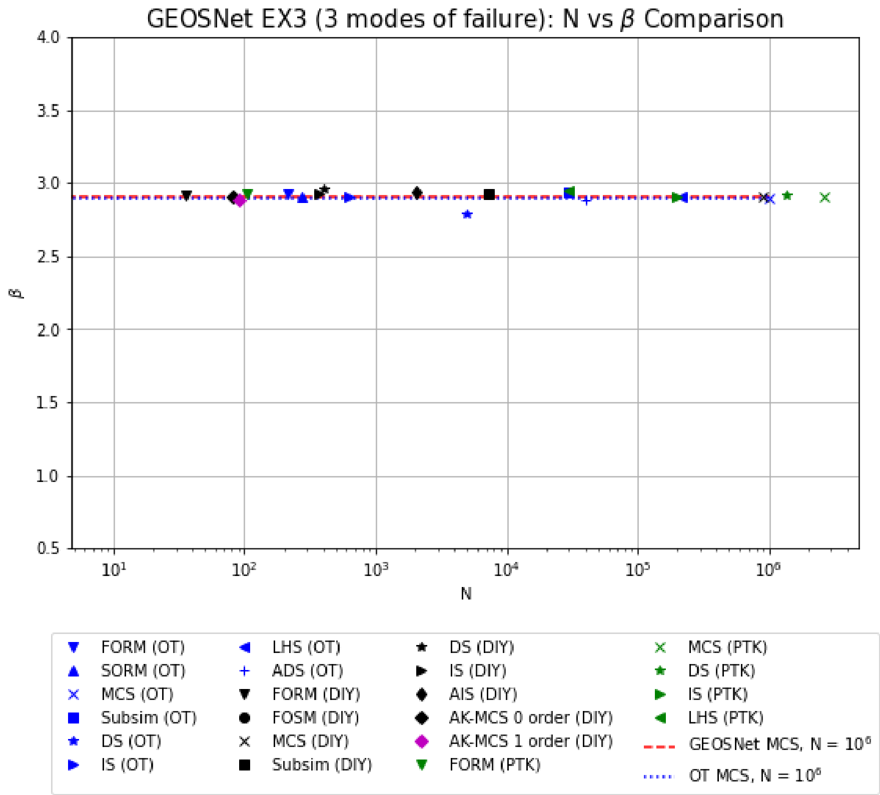


Figure 3.18: Result comparison of the GEOSNet example 3

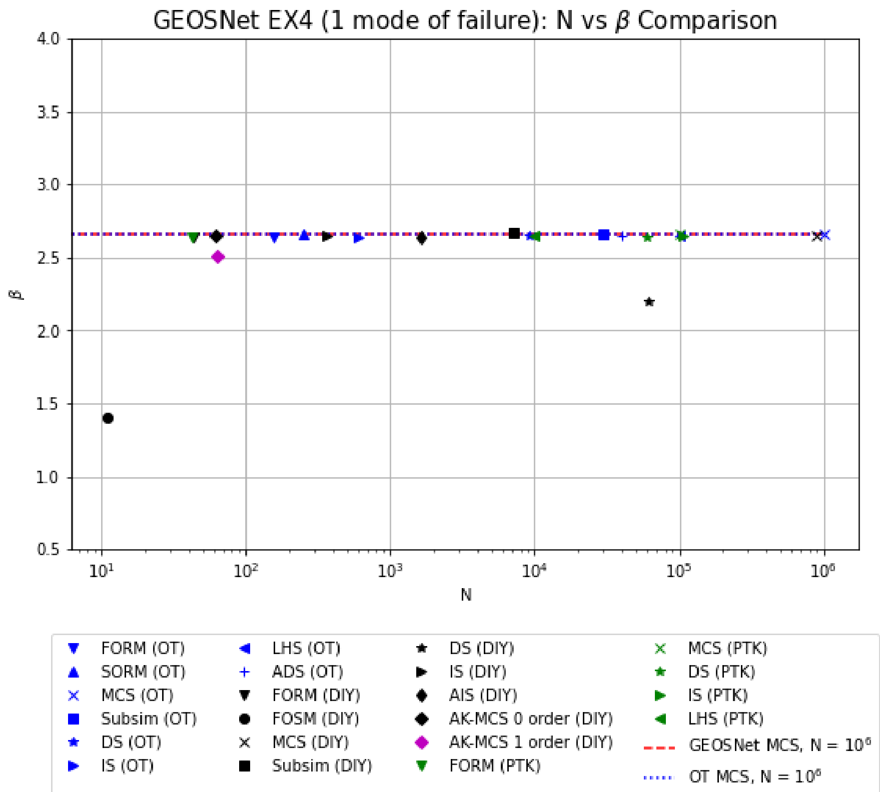


Figure 3.19: Result comparison of the GEOSNet example 4

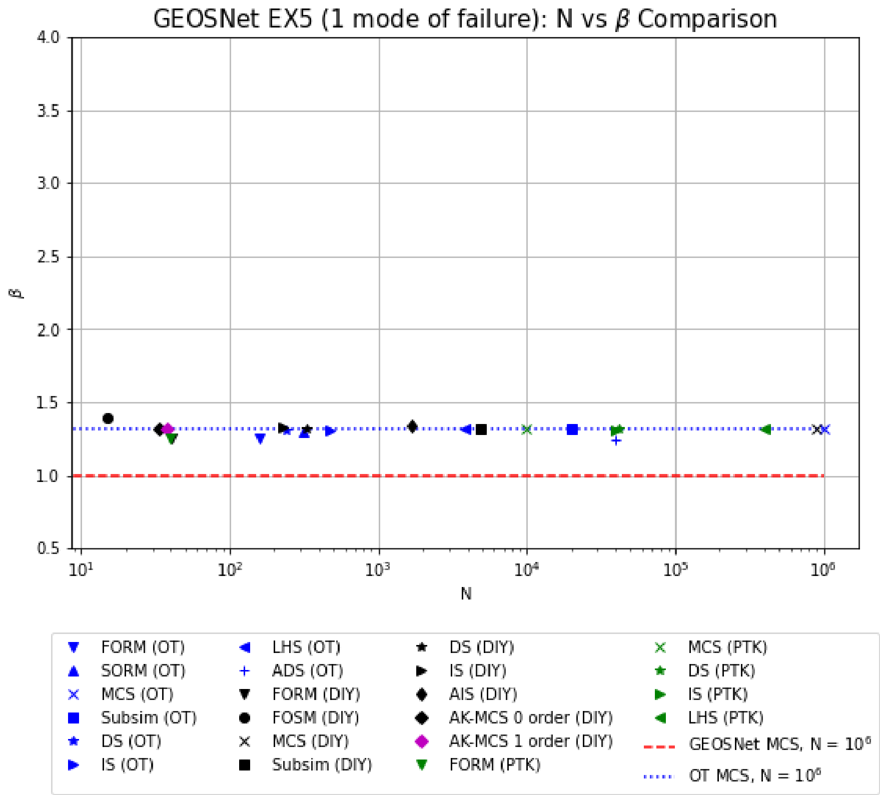


Figure 3.20: Result comparison of the GEOSNet example 5

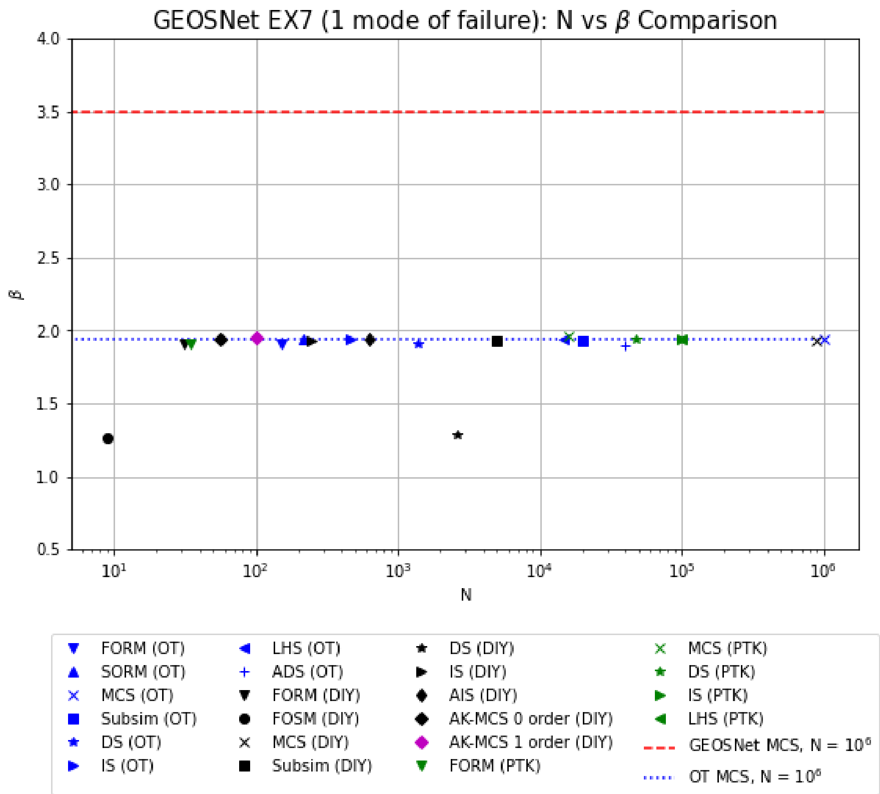


Figure 3.21: Result comparison of the GEOSNet example 7

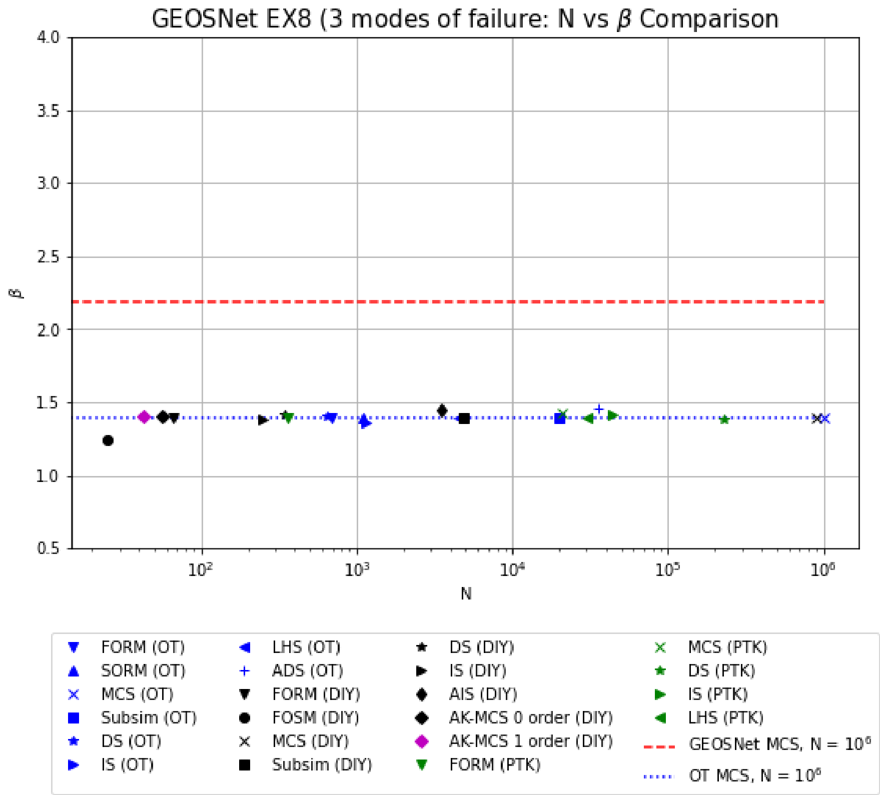


Figure 3.22: Result comparison of the GEOSNet example 8

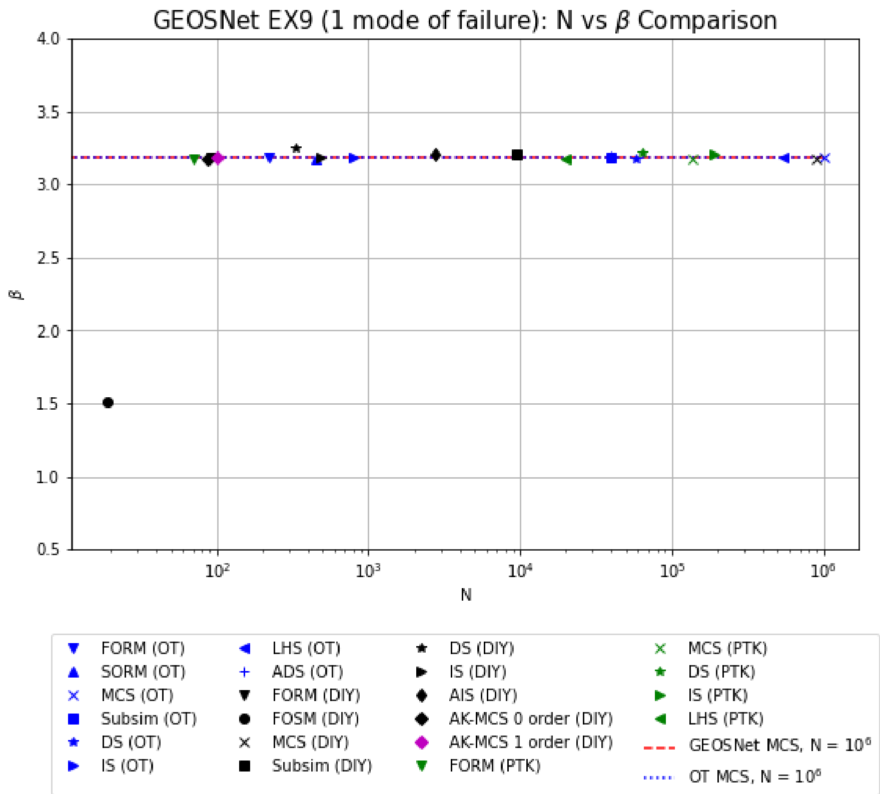


Figure 3.23: Result comparison of the GEOSNet example 9

4

Discussion

This section will discuss the results found in Chapter 3.

4.1. Result Summary

Based on Figure 3.15 to Figure 3.23 in Section 3.5, the following points can be observed.

The basic gradient-based reliability method such as FORM (OT), FORM (DIY), SORM (OT), and FOSM (DIY) failed to identify β (the closest design point to the standard space origin) in the simple slope stability and GEOSNet EX2 problems, which have more than 1 mode of failure/ design point (as can be seen in Figure 3.15 and 3.17). However, with the same problems, FORM (PTK) had a closer prediction of β (as can be observed in Figure 3.15 and 3.17 respectively) due to separate analysis of each mode of failure. Therefore, these gradient-based methods (except for FORM (PTK)) tend to overestimate β value (as discussed in Section 3.4 and shown in Figure 3.15 and 3.17). However, all of these gradient-based methods successfully identified the correct β values in GEOSNet EX3 and EX8 (Figure 3.18 and 3.22 respectively). Therefore under specific conditions, gradient-based methods could still accurately determine the correct β value. The rest of the reliability methods successfully determined the correct β values (with some variation in accuracy and efficiency) despite needing higher N values, especially with the sampling-based reliability methods.

FOSM (DIY) method often fails to determine the correct β values, even for problems with only 1 mode of failure (Figure 3.19, 3.21, and 3.23). Moreover, sampling-based methods often give accurate estimation of β values (Figure 3.15 to 3.23), however they need much more computational effort (higher N), especially when they are averaged to avoid result fluctuations (as shown in Section 3.2). Since PTK analyzed each mode of failure/ design point separately, it further increased the computational effort. This can be observed in Figure 3.18 (EX3 with 3 MoF) where the required N from PTK in MCS (PTK) and DS (PTK) exceed the required N = 10^6 of MCS (OT), also in Figure 3.15 to Figure 3.23 where FORM (PTK) needs higher N than the rest of the gradient-based methods.

The results of the reliability methods from each problem can be presented as its P_f errors (differences) from GEOSNet's Monte Carlo Analysis probability of failure (P_f). Therefore, the smaller the P_f differences (P_f errors), the more accurate the reliability method is with the corresponding problem. Moreover, the less realization (N) it needs, the more efficient the method is. The P_f errors are summarized in Table 4.1 and the N of each cases is summarized in Table 4.3. For practical purpose, the absolute β error of each method in each problem will also be displayed in Table 4.2. However, the complete P_f and β results can be found in Table A.1 and Table A.2 in the Appendix (A). The relation between P_f and β can be defined by equation 2.2.

Due to the high uncertainties in the model and incomplete data, the attempt to re-create the same performance function used in the GEOSNet reports was unsuccessful (as can be seen from Figure 3.20 to 3.22). Therefore, problems from GEOSNet Example 5, 7, and 8 will use the MCS (OT) P_f and

β with N=10^6 as references.

Table 4.1: Probability of failure (P_f) errors for each method and problem (GEOSNet EX5, EX7, and EX8 use MCS (OT) P_f as reference while the rest are based on GEOSNet's MCS result with N = 10^6).

Problems: **No. of DP:**	**Slope** **(2)**	**EX1** **(1)**	**EX2** **(2)**	**EX3** **(3)**	**EX4** **(1)**	**EX5** **(1)**	**EX7** **(1)**	**EX8** **(3)**	**EX9** **(1)**
P_f reference: β reference:	0.0246 1.97	0.0573 1.58	0.0639 1.52	0.0018 2.91	0.0039 2.66	0.0944 1.32	0.0263 1.94	0.0817 1.39	0.0007 3.17
FORM (PTK)	-0.1%	34.0%	33.3%	-0.2%	7.4%	11.2%	5.7%	0.5%	-2.9%
LHS (PTK)	-5.8%	0.000%	-11.0%	-0.3%	2.6%	-0.4%	-1.1%	0.5%	-3.3%
MCS (PTK)	0.7%	0.9%	0.2%	0.0%	1.3%	-0.2%	-5.7%	-5.8%	-2.1%
IS (PTK)	-4.2%	1.7%	2.3%	0.1%	-5.4%	0.4%	-0.8%	-4.4%	-10.8%
DS (PTK)	-0.1%	0.7%	7.8%	-0.1%	6.9%	-0.5%	0.0%	3.4%	-14.0%
FORM (OT)	-94.5%	34.4%	-74.9%	-4.8%	5.2%	11.5%	6.0%	0.5%	-2.8%
SORM (OT)	-95.1%	7.0%	-82.6%	0.3%	0.8%	3.2%	-0.3%	-0.2%	-1.2%
MCS (OT)	0.0%	1.0%	-0.2%	4.5%	0.6%	0.0%	0.0%	0.0%	-3.6%
SS (OT)	-0.8%	0.8%	0.4%	-8.4%	0.9%	-1.0%	0.7%	-0.1%	-4.3%
DS (OT)	4.9%	5.6%	1.8%	45.4%	3.7%	2.0%	8.1%	-1.2%	0.8%
IS (OT)	-11.5%	3.3%	-6.1%	1.6%	5.3%	1.4%	0.3%	5.3%	-5.5%
LHS (OT)	-0.6%	1.7%	0.6%	1.7%	2.1%	-0.4%	0.3%	-0.3%	-2.8%
ADS (OT)	2.1%	11.3%	9.9%	7.5%	2.0%	12.8%	9.3%	-11.0%	-5.7%
FORM (DIY)	-94.5%	34.2%	-74.5%	-3.3%	7.3%	11.5%	6.0%	0.2%	-2.8%
FOSM (DIY)	-61.1%	41.7%	-87.5%	-100%	1942%	-13.0%	294.7%	30.2%	8642%
MCS (DIY)	-0.7%	0.6%	1.3%	0.8%	4.2%	-0.2%	0.7%	-0.8%	-1.0%
SS (DIY)	-4.3%	0.3%	0.1%	-6.6%	-1.4%	-1.5%	3.1%	0.3%	-9.5%
DS (DIY)	1.0%	-0.1%	1.1%	-13.9%	259%	-1.2%	280.3%	-3.3%	-23.0%
IS (DIY)	-6.9%	7.6%	4.1%	-5.6%	3.7%	-2.6%	1.9%	1.6%	-2.1%
AIS (DIY)	-0.4%	0.7%	-0.3%	-8.6%	6.3%	-4.5%	-0.6%	-9.0%	-11.5%
AKMCS-0	-0.7%	-1.2%	2.3%	-0.1%	2.2%	-0.7%	-1.6%	-1.1%	0.9%
AKMCS-1	-2.5%	9.5%	30.2%	7.0%	56.4%	-1.4%	-2.0%	-1.1%	-3.9%
NI (PTK)	-2.6%	-	-	-	-	-	-	-	-
NB (PTK)	-0.5%	-	-	-	-	-	-	-	-

Table 4.2: Reliability index (β) absolute errors for each method and problem (GEOSNet EX5, EX7, and EX8 use MCS (OT) β as reference while the rest are based on GEOSNet's MCS result with N = 10^6).

Problems: **No. of DP:**	**Slope** **(2)**	**EX1** **(1)**	**EX2** **(2)**	**EX3** **(3)**	**EX4** **(1)**	**EX5** **(1)**	**EX7** **(1)**	**EX8** **(3)**	**EX9** **(1)**
P_f reference:	0.0246	0.0573	0.0639	0.0018	0.0039	0.0944	0.0263	0.0817	0.0007
β reference:	1.97	1.58	1.52	2.91	2.66	1.32	1.94	1.39	3.17
FORM (PTK)	0.01%	9.57%	10.05%	0.75%	0.90%	4.88%	1.46%	0.25%	0.32%
LHS (PTK)	1.53%	0.00%	3.74%	1.43%	0.53%	0.45%	0.09%	0.25%	0.32%
MCS (PTK)	0.01%	0.25%	0.20%	0.11%	0.15%	0.45%	1.12%	2.62%	0.32%
IS (PTK)	1.02%	0.51%	0.85%	0.06%	0.60%	0.31%	0.09%	1.90%	1.26%
DS (PTK)	0.01%	0.25%	2.82%	0.40%	0.90%	0.45%	0.09%	0.97%	1.58%
FORM (OT)	52.35%	9.66%	40.71%	0.63%	0.77%	4.73%	1.30%	0.18%	0.41%
SORM (OT)	54.21%	2.18%	50.14%	0.07%	0.22%	1.37%	0.07%	0.09%	0.26%
MCS (OT)	0.00%	0.31%	0.06%	0.37%	0.20%	0.00%	0.00%	0.00%	0.48%
SS (OT)	0.20%	0.24%	0.13%	1.04%	0.20%	0.43%	0.14%	0.04%	0.61%
DS (OT)	1.04%	1.74%	0.62%	3.99%	0.59%	0.86%	1.74%	0.48%	0.08%
IS (OT)	2.61%	1.03%	2.07%	0.07%	0.78%	0.58%	0.08%	2.02%	0.66%
LHS (OT)	0.13%	0.53%	0.22%	0.08%	0.39%	0.18%	0.06%	0.10%	0.41%
ADS (OT)	0.46%	3.43%	3.20%	0.67%	0.38%	5.23%	2.00%	4.44%	0.68%
FORM (DIY)	52.38%	9.62%	40.27%	0.46%	1.02%	4.73%	1.30%	0.08%	0.41%
FOSM (DIY)	19.06%	11.46%	58.29%	87.73%	47.17%	5.86%	34.98%	10.61%	52.38%
MCS (DIY)	0.16%	0.19%	0.45%	0.02%	0.65%	0.07%	0.16%	0.32%	0.24%
SS (DIY)	0.95%	0.11%	0.03%	0.84%	0.05%	0.63%	0.67%	0.11%	1.06%
DS (DIY)	0.22%	0.01%	0.38%	1.70%	17.51%	0.52%	33.88%	1.30%	2.52%
IS (DIY)	1.54%	2.35%	1.37%	0.72%	0.59%	1.11%	0.41%	0.63%	0.34%
AIS (DIY)	0.09%	0.24%	0.08%	1.06%	0.90%	1.97%	0.13%	3.63%	1.26%
AKMCS-0	0.15%	0.38%	0.78%	0.11%	0.41%	0.28%	0.35%	0.42%	0.06%
AKMCS-1	0.54%	2.90%	9.14%	0.63%	5.92%	0.59%	0.45%	0.44%	0.51%
NI (PTK)	0.52%								
NB (PTK)	0.01%								

Table 4.3: The number of realization N for each problem's reliability methods.

Problems: **No. of DP:**	**Slope** **(2)**	**EX1** **(1)**	**EX2** **(2)**	**EX3** **(3)**	**EX4** **(1)**	**EX5** **(1)**	**EX7** **(1)**	**EX8** **(3)**	**EX9** **(1)**
P_f reference:	0.0246	0.0573	0.0639	0.0018	0.0039	0.0944	0.0263	0.0817	0.0007
β reference:	1.97	1.58	1.52	2.91	2.66	1.32	1.94	1.39	3.17
FORM (PTK)	36	42	110	104	42	40	35	364	70
LHS (PTK)	1001	10001	2002	30030	10001	400001	100001	30003	200001
MCS (PTK)	1E+05	2E+05	1E+05	3E+06	1E+05	1E+04	2E+04	2E+04	1E+05
IS (PTK)	2880	37312	195814	200848	105968	40016	100016	43872	189200
DS (PTK)	5E+03	4E+04	8E+04	1E+06	6E+04	4E+04	5E+04	2E+05	6E+04
FORM (OT)	157	144	198	215	157	158	152	694	218
SORM (OT)	179	252	415	272	253	314	215	1089	452
MCS (OT)	1E+06	1E+06	1E+06	1E+06	1E+06	1E+06	1E+06	1E+06	1E+06
SS (OT)	20000	20000	20000	30000	30000	20000	20000	20000	40000
DS (OT)	239	523	863	4966	9094	240	1380	649	57341
IS (OT)	38218	514	3450	622	599	471	470	1153	805
LHS (OT)	16118	6486	5839	218704	100317	3870	14811	4519	549733
ADS (OT)	39998	39950	39815	39996	39992	39920	39992	35614	39871
FORM (DIY)	22	71	211	36	43	41	31	66	91
FOSM (DIY)	5	13	19	9	11	15	9	25	19
MCS (DIY)	9E+05	9E+05	9E+05	9E+05	9E+05	9E+05	9E+05	9E+05	9E+05
SS (DIY)	4905	4905	4905	7228	7228	4905	4905	4905	9552
DS (DIY)	329	369	384	400	60659	328	2596	344	330
IS (DIY)	1061	287	2061	373	362	230	246	246	479
AIS (DIY)	447	908	2698	2034	1631	1686	634	3532	2756
AKMCS-0	21	51	100	82	62	34	56	56	88
AKMCS-1	28	29	92	91	64	38	100	43	100
NI (PTK)	40001	-	-	-	-	-	-	-	-
NB (PTK)	42369	-	-	-	-	-	-	-	-

The required N values displayed in Table 4.3 for the sampling-based methods are based on the average of N from 100 simulations. These additional simulations are only applied to the sampling-based methods other than MCS (LHS, IS, DS, SS, ADS, AIS, and both AK-MCS).

Moreover, Table 4.1 can also be presented as a graph based on P_f errors and the number of independent variable throughout different β, as displayed in Figure 4.1 and 4.3 below. Furthermore, the range of the results are displayed in Figure 4.2 and Figure 4.4 for P_f errors throughout different β and number of independent variables respectively. Due to the extreme differences in FOSM (DIY), Figure 4.2 and Figure 4.4 display the range by including and excluding FOSM (DIY) result.

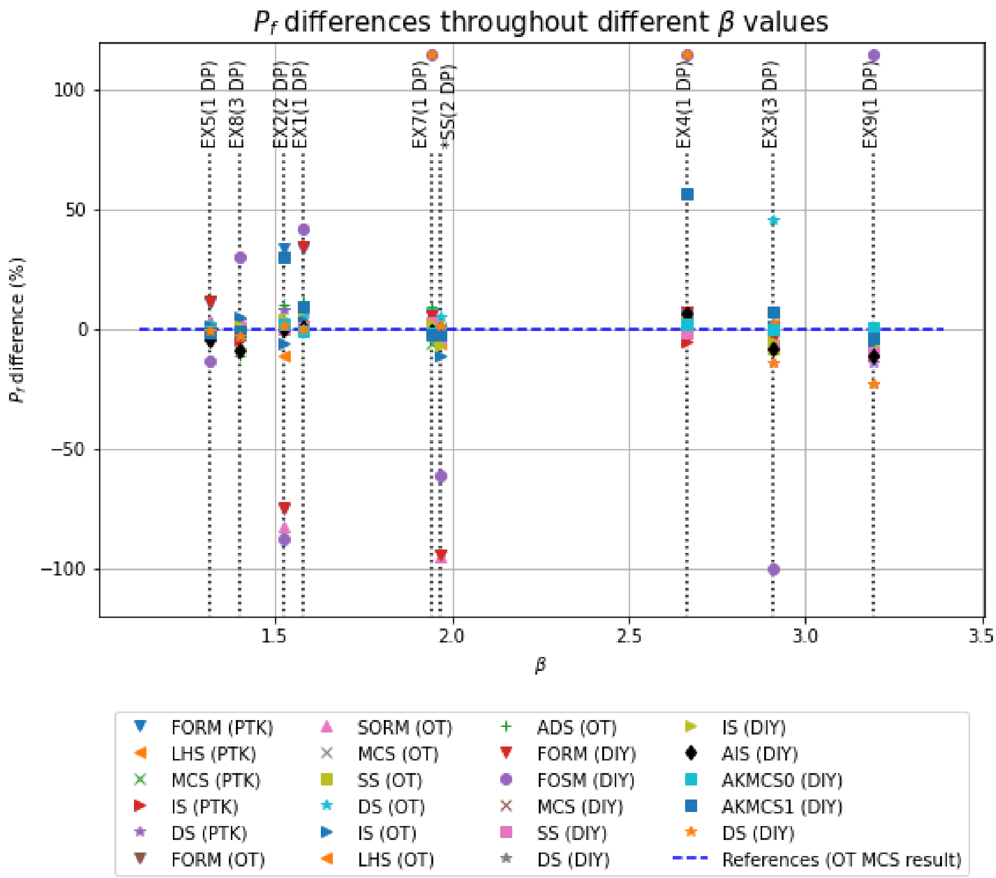


Figure 4.1: β differences throughout different β values (*SS = Slope stability problem).

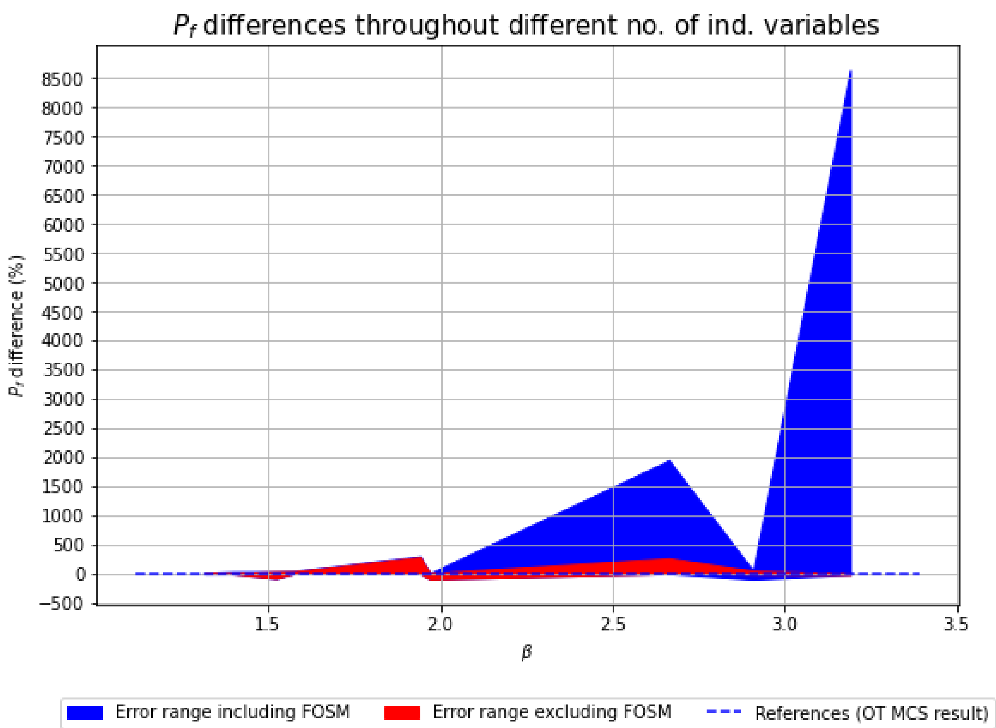


Figure 4.2: The range of β differences throughout different β values.

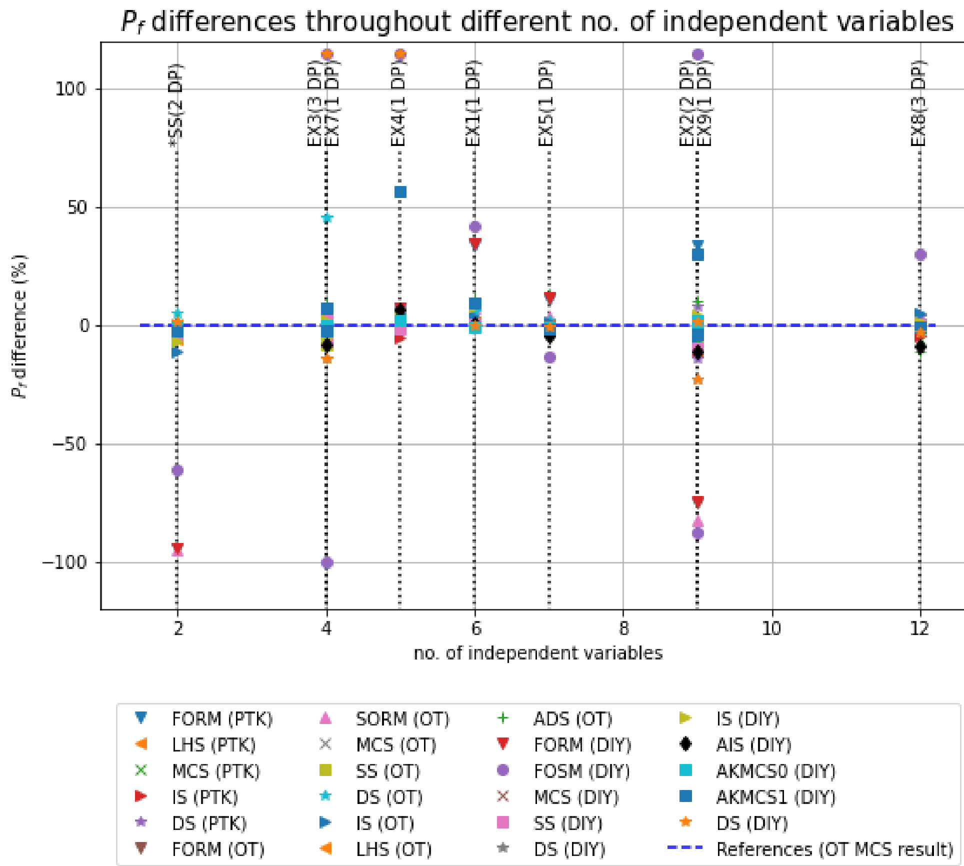


Figure 4.3: β differences throughout different no. of independent variables (*SS = Slope stability problem).

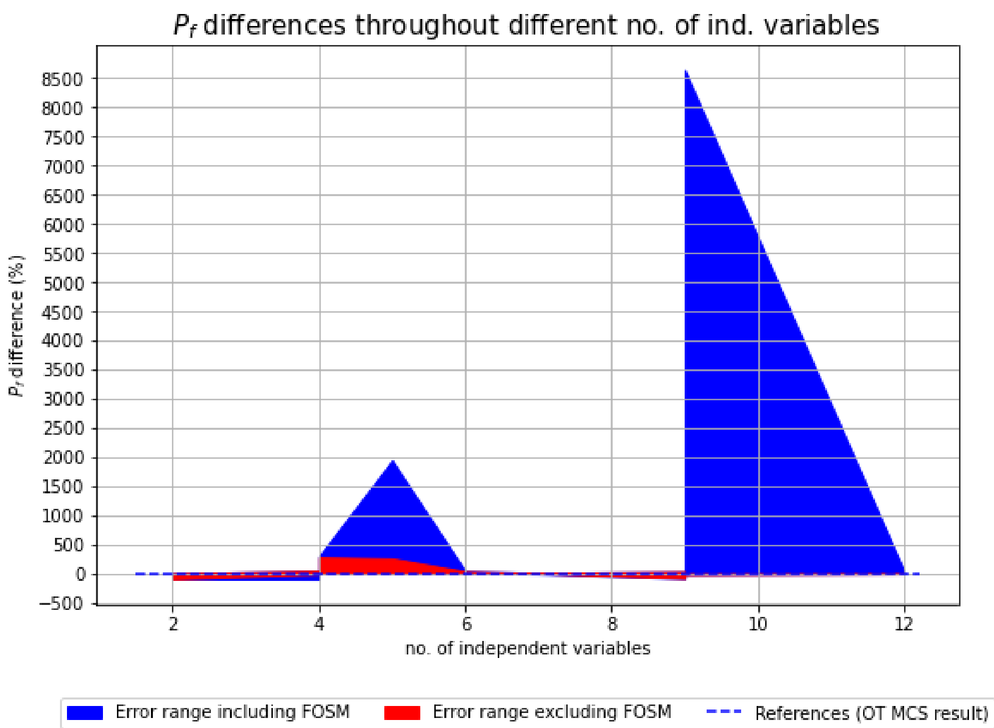


Figure 4.4: The range of β differences throughout different no. of independent variables.

It can be observed from Figure 4.1 and 4.3 that generally the problems with more than 1 mode of failure or design points (DP) have considerably higher P_f errors compared to the problems with 1 DP. It can also be observed that FOSM (DIY) and DS (DIY) has the highest P_f errors (more than 100% in some examples. For practicality purpose, Figure 4.1 and 4.3 only display differences within 100% range, therefore any error higher than 100% is displayed at the border of the graphs. It is worth mentioning that for β value of 2.64 (EX4) and 3.19 (EX9) with number of independent variables of 5 & 9, FOSM(DIY) P_f errors reach a staggering value of 1942% and 8642% respectively.

Figure 4.2 and 4.4 show the ranges of all the P_f differences from each reliability method. Since FOSM (DIY) errors are extremely high and could be considered as outliers, Figure 4.2 and 4.4 also include the error ranges without FOSM (DIY).

Moreover, despite there is a tendency for FOSM (DIY) to get higher P_f errors with higher β (or lower P_f) and higher number of independent variables, there is no clear indication for the rest of the reliability methods to follow the same pattern based on 4.1 and 4.4, or Table 4.1. More examples with different β references and varying number of independent variables are needed to correctly correlate between β and P_f errors or between number of independent variables and P_f errors.

4.2. Rankings

To further compare the reliability index β and P_f errors between each method, a ranking system will be implemented. Within each problem, the methods will be ranked according to it's absolute P_f errors, where the higher the absolute P_f error means the higher the rank (the lower rank is the better). For example, based on Table 4.1, LHS (PTK) obtained the most accurate result for GEOSNet EX1, therefore it is ranked as the 1st in term accuracy. Similarly, FOSM (DIY) is ranked as the least accurate (22nd) for the same problem since it has the highest P_f error (Table 4.1). The same principle is also applied according to the number of required N in each method. Since numerical integration (NI) and numerical bisection (NB) are only applied in the slope stability problem, these methods will be excluded in the ranking system.

The ranks based on each method's absolute P_f errors (accuracy) are presented in Table 4.4 and the ranks for the required N (efficiency) are presented in Table 4.5.

Table 4.4: The rank of each method based on its absolute P_f errors (the lower the rank the better).

Problems: **No. of DP:**	**Slope** **(2)**	**EX1** **(1)**	**EX2** **(2)**	**EX3** **(3)**	**EX4** **(1)**	**EX5** **(1)**	**EX7** **(1)**	**EX8** **(3)**	**EX9** **(1)**
P_f reference: β reference:	0.0246 1.97	0.0573 1.58	0.0639 1.52	0.0018 2.91	0.0039 2.66	0.0944 1.32	0.0263 1.94	0.0817 1.39	0.0007 3.17
FORM (PTK)	3	19	18	5	19	18	16	7	10
LHS (PTK)	16	1	16	7	9	6	10	7	11
MCS (PTK)	7	8	2	1	4	3	15	19	6
IS (PTK)	13	12	11	3	15	4	9	17	18
DS (PTK)	2	5	14	2	17	7	2	16	20
FORM (OT)	20	21	20	13	13	20	18	9	7
SORM (OT)	22	15	21	6	2	16	4	4	4
MCS (OT)	1	9	3	12	1	1	1	1	12
SS (OT)	9	7	5	18	3	9	7	2	14
DS (OT)	15	14	9	21	11	14	19	13	1
IS (OT)	18	13	13	9	14	11	5	18	15
LHS (OT)	5	11	6	10	7	5	3	5	8
ADS (OT)	11	18	15	17	6	21	20	21	16
FORM (DIY)	21	20	19	11	18	19	17	3	9
FOSM (DIY)	19	22	22	22	22	22	22	22	22
MCS (DIY)	8	4	8	8	12	2	8	10	3
SS (DIY)	14	3	1	15	5	13	14	6	17
DS (DIY)	10	2	7	20	21	10	21	15	21
IS (DIY)	17	16	12	14	10	15	12	14	5
AIS (DIY)	4	6	4	19	16	17	6	20	19
AKMCS-0	6	10	10	4	8	8	11	11	2
AKMCS-1	12	17	17	16	20	12	13	12	13

Table 4.5: The rank of each method based on its N (the lower the better).

Problems: **No. of DP:**	**Slope** **(2)**	**EX1** **(1)**	**EX2** **(2)**	**EX3** **(3)**	**EX4** **(1)**	**EX5** **(1)**	**EX7** **(1)**	**EX8** **(3)**	**EX9** **(1)**
P_f reference:	0.0246	0.0573	0.0639	0.0018	0.0039	0.0944	0.0263	0.0817	0.0007
β reference:	1.97	1.58	1.52	2.91	2.66	1.32	1.94	1.39	3.17
FORM (PTK)	5	3	4	5	2	4	3	7	2
LHS (PTK)	11	15	10	15	13	20	19	17	19
MCS (PTK)	20	20	19	22	19	15	15	16	17
IS (PTK)	13	17	20	17	20	18	20	19	18
DS (PTK)	15	19	18	21	16	19	18	20	16
FORM (OT)	6	6	5	6	6	6	6	9	6
SORM (OT)	7	7	8	7	7	9	7	10	8
MCS (OT)	22	22	22	20	22	22	22	22	22
SS (OT)	17	16	16	14	14	16	16	15	14
DS (OT)	8	11	9	12	12	8	11	8	15
IS (OT)	18	10	13	10	9	11	9	11	10
LHS (OT)	16	14	15	18	18	13	14	13	20
ADS (OT)	19	18	17	16	15	17	17	18	13
FORM (DIY)	3	5	6	2	3	5	2	4	4
FOSM (DIY)	1	1	1	1	1	1	1	1	1
MCS (DIY)	21	21	21	19	21	21	21	21	21
SS (DIY)	14	13	14	13	11	14	13	14	12
DS (DIY)	9	9	7	9	17	10	12	6	7
IS (DIY)	12	8	11	8	8	7	8	5	9
AIS (DIY)	10	12	12	11	10	12	10	12	11
AKMCS-0	2	4	3	3	4	2	4	3	3
AKMCS-1	4	2	2	4	5	3	5	2	5

The ranks (based on absolute P_f errors and N values) of each method can be further summed, therefore the final rank will be based on the total sum (the lower the total sum is, the higher the final rank is). For example, the total P_f error ranks of FORM (PTK) is (21 + 19 + 17 + ... + 11 =). This can be presented in Table 4.6.

Table 4.6: The cummulative ranks based on P_f errors and N values (the lower the better).

Method	**Tot. P_f error rank (a)**	**Tot. N value rank (b)**	**Tot. rank (c = a+b)**	**Final rank (d)**
FORM (PTK)	115	35	150	2
LHS (PTK)	83	139	222	16
MCS (PTK)	65	163	228	17
IS (PTK)	102	162	264	21
DS (PTK)	85	162	247	19
FORM (OT)	141	56	197	7
SORM (OT)	94	70	164	3
MCS (OT)	41	196	237	18
SS (OT)	74	138	212	13
DS (OT)	117	94	211	11
IS (OT)	116	101	217	15
LHS (OT)	60	141	201	8
ADS (OT)	145	150	295	22
FORM (DIY)	137	34	171	5
FOSM (DIY)	195	9	204	9
MCS (DIY)	63	187	250	20
SS (DIY)	88	118	206	10
DS (DIY)	127	86	213	14
IS (DIY)	115	76	191	6
AIS (DIY)	111	100	211	11
AKMCS-0	70	28	98	1
AKMCS-1	132	32	164	3

Finally, the best reliability methods according to the P_f errors, N, and total performance (P_f errors + N values) are presented in Table 4.7

Table 4.7: The final ranks (the lower the better).

Ranks	**P_f error ranks**	**N value rank**	**Final ranks**
1	MCS (OT)	FOSM (DIY)	**AK-MCS 0**
2	LHS (OT)	AK-MCS 0	**FORM (PTK)**
3	MCS (DIY)	AK-MCS 1	**SORM (OT)**
4	MCS (PTK)	FORM (DIY)	**AK-MCS 1**
5	AK-MCS 0	FORM (PTK)	**FORM (DIY)**
6	SS (OT)	FORM (OT)	**IS (DIY)**
7	LHS (PTK)	SORM (OT)	**FORM (OT)**
8	DS (PTK)	IS (DIY)	**LHS (OT)**
9	SS (DIY)	DS (DIY)	**FOSM (DIY)**
10	SORM (OT)	DS (OT)	**SS (DIY)**
11	IS (PTK)	AIS (DIY)	**AIS (DIY)**
12	AIS (DIY)	IS (OT)	**DS (OT)**
13	FORM (PTK)	SS (DIY)	**SS (OT)**
14	IS (DIY)	SS (OT)	**DS (DIY)**
15	IS (OT)	LHS (PTK)	**IS (OT)**
16	DS (OT)	LHS (OT)	**LHS (PTK)**
17	DS (DIY)	ADS (OT)	**MCS (PTK)**
18	AK-MCS 1	DS (PTK)	**MCS (OT)**
19	FORM (DIY)	IS (PTK)	**DS (PTK)**
20	FORM (OT)	MCS (PTK)	**MCS (DIY)**
21	ADS (OT)	MCS (DIY)	**IS (PTK)**
22	FOSM (DIY)	MCS (OT)	**ADS (OT)**

It can be observed from Table 4.6 that DS (OT) and AIS (DIY) obtained the same final rank of 11. However, AIS (DIY) has less score in P_f error rank (111) compared to DS (OT) P_f error score (117). Therefore, in Table 4.7, AIS (DIY) is assigned the higher final rank (11) compared to DS (OT) final rank (12).

5

Conclusion and Reflection

5.1. Conclusions

Based on the things discussed in Chapter 3 and 4, the following things can be concluded:

- It can be observed from Table 4.7 that MCS (OT) with N=10^6 is ranked as the most accurate method despite its shortcoming in efficiency where it is also ranked as the most inefficient method at the same time. However, the accuracy rank is highly influenced by its P_f use as the reference for GEOSNet example 5, 7, and 9. Moreover, its inaccuracy is highly influence by its N value which was kept constant at 10^6.
- Furthermore, Monte Carlo Simulation methods are ranked as the most inefficient methods in this research since Table 4.7 shows that all of the 3 MCS methods are positioned at the bottom 3 in the efficiency ranking. This is due to the fact that MCS (OT) and MCS (DIY) are having fixed N numbers of 10^6 and 9x10^5 respectively. However, MCS (PTK)'s number of N is obtained after reaching the determined convergence value (COV = 0.1) therefore its N values are lower than MCS (OT) and MCS (DIY). It can be concluded that Monte Carlo Simulation is the most inefficient method in this research.
- It can also be observed from Table 4.7 that the FOSM (DIY) method obtained the most efficient method despite it is also ranked as the least accurate method at the same time. It is also shown in Table 4.1 and Figure 4.2 that FOSM (DIY) accuracy further decreases for β value larger than 2 (or P_f value lower than 2.275%). Therefore, FOSM (DIY) is the most inaccurate method in this research.
- Combining the accuracy and efficiency, AK-MCS 0 is ranked as the best reliability method in this research. It is ranked 5^{th} in term of accuracy and 2^{nd} in term of efficiency.
- Gradient-based reliability methods such as FORM (OT), FORM (DIY), SORM (OT), and FOSM (DIY) failed to identify the correct design point for the simple slope stability and GEOSNet EX2 problems (which have more than 1 mode of failure), as can be seen in Table 4.1, Figure 3.15, and 3.17. However, FORM (PTK) accurately predicts the correct β value by analyzing each mode of failure separately.
- However in some cases, the gradient-based reliability methods could still accurately determine the correct β values for problems with more than 1 mode of failure. This can be observed in in GEOSNet EX3 and EX8 (Figure 3.18 and 3.22 respectively). Therefore, it is recommended to perform at least 1 sampling-based reliability method for confirmation purpose when one performs a gradient-based reliability method.
- Except for FOSM (DIY), Figure 4.1, Figure 4.2, and Table 4.1 show that the higher the problem's β value (or lower P_f value) does not necessarily give a wider range of P_f errors.
- Figure 4.4 and Table 4.1 show that the increase in number of independent variable does not necessarily give a wider range of P_f errors.
- Although SORM (OT) and FORM (DIY), and FORM (OT) method obtained relatively high final ranks based on 4.7 (final rank of 3, 5, and 7 respectively), they must be implemented very

carefully due to their limitation in correctly identifying the P_f of problems with more than 1 design points (as can be seen in Table 4.1).

- For problems with multiple modes of failure (multiple design points), analyzing and implementing reliability analysis for each mode of failure is strongly recommended when implementing FORM method (as discussed in Section 3.4). However, this takes more computational efforts despite it is still much less than the sampling-based methods' computational efforts. This can be seen in FORM (PTK) cases based on Table 4.3.
- Based on Table 4.1 and 4.7, it can be observed that sampling-based reliability methods (where the P_f, β, and N results are averaged from 100 simulations except for Monte Carlo Simulation) have higher accuracy compared to the gradient-based methods. However, gradient-based methods have higher efficiency.
- Sampling-based methods tend to fluctuate between simulations, therefore it is recommended to perform multiple simulations and get the average β, P_f, and N values (as shown in Section 3.2).
- The accuracy and efficiency of the reliability analysis results are dependant on convergence criteria, where a higher accuracy criteria (lower COV threshold and P_f errors) needs higher N values (as shown in Section 2.15).
- Based on Table 4.1 and Figure 4.1, it can be seen that every reliability method has a different accuracy and efficiency for each problem. Therefore, it is highly recommended to take at least more than 1 type of reliability method for each problem to give a better prediction of the problem's β value.

5.2. Reflection

The following are limitations and considerations that could have further improved the research:

- More up-to-date reliability methods may give more accurate β estimate compared to this research, e.g. improved FORM method that could differentiate multiple design points for problems with more than 1 design point (e.g. Der Kiureghian et al. 1998 [9]).
- Despite this research is only limited to the 9 selected problems, geotechnical engineering has a very wide range of problem types with far more complex performance functions and design points (including FEM-based models). Therefore, the result of this research only limited to the discussed problems.
- FEM-based application is widely used in geotechnical engineering practice. Unfortunately, this research does not include FEM-based problem into consideration due to time limitation.

References

[1] Anne Dutfoy et al. *OpenTURNS*. URL: `https://openturns.github.io/www/`.

[2] Nicholas Metropolis et al. “Equation of State Calculations by Fast Computing Machines”. In: *The Journal of Chemical Physics* 21 (1953), Issue 6.

[3] S. K. Au and J.L. Beck. “A new adaptive importance sampling scheme for reliability calculations”. In: *Structural Safety* 21 (1999), pp. 135–158.

[4] Siu-Kui Au and James L. Beck. “Estimation of small failure probabilities in high dimensions by subset simulation”. In: *Probabilistic Engineering Mechanics* 16 (2001), pp. 263–277.

[5] Kuang-Hua Chang. *e-Design, Computer-Aided Engineering Design*. -. Elsevier, 2015. ISBN: 978-0-12-382038-9.

[6] C. A. Cornell. “Bounds on the reliability of structural systems”. In: *Journal of the structural division, ASCE* 93(1) (1967), pp. 171–200.

[7] Deltares. *Probabilistic Toolkit (PTK)*. URL: `https://www.deltares.nl/en/software/probabilistic-toolkit-ptk/`.

[8] Deltares. *Probabilistic Toolkit Software User Manual Guide*. Version: 2.0, SVN Revision: 00. Deltares, 2021.

[9] Dakessian T. Der Kiureghian A. “Multiple design points in first and second-order reliability”. In: *Structural Safety* 20 (1998), pp. 37–49.

[10] B. Echard, N. Gayton, and M. Lemaire. “AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation”. In: *Structural Safety* 33 (2011), pp. 145–154.

[11] Geoengineer.org. *Geotechnical Safety network*. URL: `https://www.geoengineer.org/geosnet/`.

[12] Frank Grooteman. “Adaptive radial-based importance sampling method for structural reliability.” In: *Structural Safety* 30 (2008).

[13] Frank Grooteman. “An adaptive directional importance sampling method for structural reliability”. In: *Probabilistic Engineering Mechanics* 26 (2011), pp. 134–141.

[14] J. Helton and F.J. Davis. “Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems”. In: *SAND 2001-0417* (2002).

[15] Michael A. Hicks and Kristinah Samy. “Influence of heterogeneity on undrained clay slope stability”. In: *Quarterly Journal of Engineering Geology and Hydrogeology* 35 (2008), pp. 41–49.

[16] $BruceR.EllingwoodJinsuoNie$. “Directional methods for structural reliability analysis”. In: *Structural Safety* 22 (2000), pp. 233–249.

[17] R. E. Melchers. “Importance Sampling in Structural Systems”. In: *Structural Safety 6: 3-10* (1989).

[18] R. E. Melchers and A. T. Beck. *Structural Reliability Analysis and Prediction*. Chichester John Wiley & Sons, 2018.

[19] Geotechnical Safety network. *Reliability Benchmarking*. URL: `http://140.112.12.21/issmge/tc304.htm`.

[20] OpenTURNS. *OpenTURNS’ SORM application*. URL: `https://openturns.github.io/openturns/latest/auto_reliability_sensitivity/reliability/plot_estimate_probability_form.html`.

[21] OpenTurns.github.io. *Latin Hypercube Simulation*. URL: `https://openturns.github.io/openturns/latest/theory/reliability_sensitivity/lhs_simulation.html`.

[22] OpenTurns.github.io. *Use the Adaptive Directional Stratification Algorithm*. URL: https://openturns.github.io/openturns/latest/auto_reliability_sensitivity/reliability/plot_estimate_probability_adaptive_directional_sampling.

[23] KK Phoon. "GEOSNet Reliability Benchmarking - Example 1". In: (2008).

[24] Rüdiger Rackwitz and Bernd Flessler. "Structural reliability under combined random load sequences". In: *Computers & Structures vol. 9 iss. 5* 9.5 (1978), pp. 489–494. DOI: 10.1016/0045-7949(78)90046-9.

[25] Timo Schweckendiek. "Structural reliability applied to deep excavations. Coupling reliability methods with finite elements." In: *TU Delft MSc Thesis* (2006).

[26] Rui Teixeira, Maria Nogal, and Alan O'Connor. "Adaptive approaches in metamodel-based reliability analysis: A review". In: *Structural Safety* 89 (2021).

[27] Wikipedia. *Monte Carlo Method*. URL: https://en.wikipedia.org/wiki/Monte_Carlo_method.

A

The results of β and P_f

This section complete the results as mentioned in Section 4.1.

Table A.1: Reliability index (β) of each problem.

Problems: **No. of DP:**	**Slope** **(2)**	**EX1** **(1)**	**EX2** **(2)**	**EX3** **(3)**	**EX4** **(1)**	**EX5** **(1)**	**EX7** **(1)**	**EX8** **(3)**	**EX9** **(1)**
P_f reference:	0.0246	0.0573	0.0639	0.0018	0.0039	0.0944	0.0263	0.0817	0.0007
β reference:	1.97	1.58	1.52	2.91	2.66	1.32	1.94	1.39	3.17
FORM (PTK)	1.9700	1.4270	1.3700	2.9300	2.6400	1.2500	1.9100	1.3900	3.1800
LHS (PTK)	2.0000	1.5780	1.5800	2.9500	2.6500	1.3200	1.9400	1.3900	3.1800
MCS (PTK)	1.9700	1.5740	1.5200	2.9114	2.6600	1.3200	1.9600	1.4300	3.1800
IS (PTK)	1.9900	1.5700	1.5100	2.9100	2.6800	1.3100	1.9400	1.4200	3.2100
DS (PTK)	1.9700	1.5740	1.4800	2.9200	2.6400	1.3200	1.9400	1.3800	3.2200
FORM (OT)	3.0011	1.4256	2.1430	2.9266	2.6435	1.2519	1.9130	1.3910	3.1829
SORM (OT)	3.0376	1.5437	2.2866	2.9103	2.6580	1.2961	1.9396	1.3948	3.1782
MCS (OT)	1.9698	1.5731	1.5239	2.8975	2.6586	1.3141	1.9382	1.3935	3.1853
SS (OT)	1.9737	1.5742	1.5211	2.9385	2.6586	1.3197	1.9354	1.3940	3.1892
DS (OT)	1.9494	1.5506	1.5135	2.7923	2.6482	1.3028	1.9045	1.4003	3.1725
IS (OT)	2.0212	1.5617	1.5545	2.9062	2.6432	1.3065	1.9368	1.3654	3.1910
LHS (OT)	1.9725	1.5697	1.5196	2.9060	2.6536	1.3165	1.9371	1.3950	3.1830
ADS (OT)	1.9608	1.5238	1.4743	2.8887	2.6538	1.2454	1.8995	1.4555	3.1917
FORM (DIY)	3.0017	1.4261	2.1364	2.9217	2.6367	1.2519	1.9130	1.3924	3.1830
FOSM (DIY)	2.3453	1.3972	2.4108	5.4597	1.4074	1.3911	1.2602	1.2456	1.5097
MCS (DIY)	1.9729	1.5751	1.5161	2.9087	2.6468	1.3150	1.9352	1.3979	3.1777
SS (DIY)	1.9887	1.5763	1.5225	2.9326	2.6654	1.3224	1.9252	1.3920	3.2036
DS (DIY)	1.9655	1.5782	1.5172	2.9576	2.1975	1.3209	1.2815	1.4117	3.2498
IS (DIY)	2.0002	1.5409	1.5021	2.9292	2.6483	1.3287	1.9302	1.3848	3.1808
AIS (DIY)	1.9717	1.5742	1.5242	2.9391	2.6400	1.3399	1.9408	1.4441	3.2099
AKMCS-0	1.9728	1.5840	1.5111	2.9114	2.6532	1.3178	1.9450	1.3994	3.1720
AKMCS-1	1.9806	1.5322	1.3839	2.8899	2.5064	1.3218	1.9470	1.3996	3.1861
NI (PTK)	1.9800	-	-	-	-	-	-	-	-
NB (PTK)	1.9700	-	-	-	-	-	-	-	-

Table A.2: Probability of failure (P_f) of each problem.

Problems: **No. of DP:**	**Slope (2)**	**EX1 (1)**	**EX2 (2)**	**EX3 (3)**	**EX4 (1)**	**EX5 (1)**	**EX7 (1)**	**EX8 (3)**	**EX9 (1)**
P_f reference:	0.0246	0.0573	0.0639	0.0018	0.0039	0.0944	0.0263	0.0817	0.0007
β reference:	1.97	1.58	1.52	2.91	2.66	1.32	1.94	1.39	3.17
FORM (PTK)	0.0244	0.0768	0.0852	0.0017	0.0042	0.1050	0.0278	0.0821	0.0007
LHS (PTK)	0.0230	0.0573	0.0569	0.0016	0.0040	0.0940	0.0260	0.0821	0.0007
MCS (PTK)	0.0246	0.0578	0.0640	0.0018	0.0040	0.0942	0.0248	0.0770	0.0007
IS (PTK)	0.0234	0.0583	0.0654	0.0018	0.0037	0.0948	0.0261	0.0781	0.0007
DS (PTK)	0.0244	0.0577	0.0689	0.0018	0.0042	0.0939	0.0263	0.0845	0.0006
FORM (OT)	0.0013	0.0770	0.0161	0.0017	0.0041	0.1053	0.0279	0.0821	0.0007
SORM (OT)	0.0012	0.0613	0.0111	0.0018	0.0039	0.0975	0.0262	0.0815	0.0007
MCS (OT)	0.0244	0.0579	0.0638	0.0019	0.0039	0.0944	0.0263	0.0817	0.0007
SS (OT)	0.0242	0.0578	0.0641	0.0016	0.0039	0.0935	0.0265	0.0817	0.0007
DS (OT)	0.0256	0.0605	0.0651	0.0026	0.0040	0.0963	0.0284	0.0807	0.0008
IS (OT)	0.0216	0.0592	0.0600	0.0018	0.0041	0.0957	0.0264	0.0861	0.0007
LHS (OT)	0.0243	0.0582	0.0643	0.0018	0.0040	0.0940	0.0264	0.0815	0.0007
ADS (OT)	0.0249	0.0638	0.0702	0.0019	0.0040	0.1065	0.0287	0.0728	0.0007
FORM (DIY)	0.0013	0.0769	0.0163	0.0017	0.0042	0.1053	0.0279	0.0819	0.0007
FOSM (DIY)	0.0095	0.0812	0.0080	0.0000	0.0796	0.0821	0.1038	0.1064	0.0656
MCS (DIY)	0.0243	0.0576	0.0647	0.0018	0.0041	0.0943	0.0265	0.0811	0.0007
SS (DIY)	0.0234	0.0575	0.0639	0.0017	0.0038	0.0930	0.0271	0.0820	0.0007
DS (DIY)	0.0247	0.0573	0.0646	0.0016	0.0140	0.0933	0.1000	0.0790	0.0006
IS (DIY)	0.0227	0.0617	0.0665	0.0017	0.0040	0.0920	0.0268	0.0831	0.0007
AIS (DIY)	0.0243	0.0577	0.0637	0.0016	0.0041	0.0901	0.0261	0.0744	0.0007
AKMCS-0	0.0243	0.0566	0.0654	0.0018	0.0040	0.0938	0.0259	0.0808	0.0008
AKMCS-1	0.0238	0.0627	0.0832	0.0019	0.0061	0.0931	0.0258	0.0808	0.0007
NI (PTK)	0.0238	-	-	-	-	-	-	-	-
NB (PTK)	0.0243	-	-	-	-	-	-	-	-

A Microsoft Excel file of the table (and all results) is available in: **click here** or https://bit.ly/32Hu65v or by an email request to rayyan8818@gmail.com.

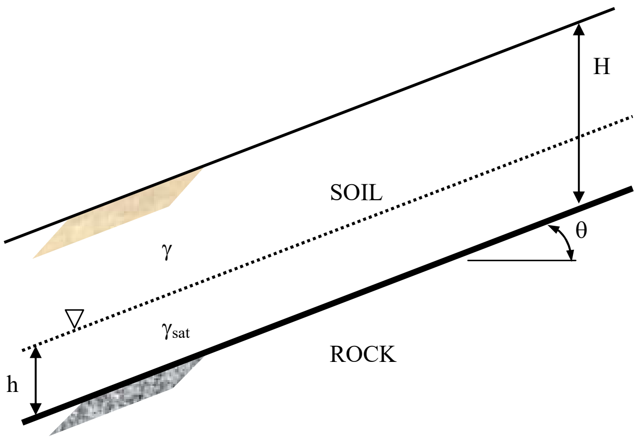
B

GEOSNet Example problems

B.1. GEOSNet problems performance functions and distributions

geotechnical safety network

TASK GROUP 3 – RELIABILITY BENCHMARKING

Example No.	1
Author(s)	KK Phoon
Date	11 April 2008 (ver. 1)
Brief description	Infinite slope problem with 6 independent random variables. The purpose is to show that “independent” variables must be carefully selected so that they do not violate physics, e.g. height of water table and soil unit weights shown below.
Figure	
Performance function	$P = \frac{[\gamma(H-h)+h(\gamma_{sat}-\gamma_w)]\cos\theta\tan\phi}{[\gamma(H-h)+h\gamma_{sat}]\sin\theta} - 1$ H = depth of soil above bedrock h = height of groundwater table above bedrock γ and γ_{sat} = moist unit weight and saturated unit weight of the surficial soil, respectively γ_w = unit weight of water (9.81 kN/m^3) ϕ = effective stress friction angle θ = slope inclination Note that the height of the groundwater table (h) can not exceed the depth of surficial soil (H) and can not be negative. Hence, it is modeled by h = H×U, in which U = standard uniform variable. The

moist and saturated soil unit weights are not independent, because they are related to the specific gravity of the soil solids (G_s) and the void ratio (e). The uncertainties in γ and γ_{sat} are characterized by modeling G_s and e as two independent uniform random variables. There are six independent random variables in this problem (H, U, ϕ, β, e, and G_s)

Inputs

Variable	Description	Distribution	**Statistics**
H	**Depth of soil above bedrock**	**Uniform**	[2,8] m
$h = H \times U$	**Height of water table**	**U is uniform**	[0, 1]
ϕ	**Effective stress friction angle**	**Lognormal**	mean = 35° cov = 8%
θ	**Slope inclination**	**Lognormal**	mean = 20° cov = 5%
γ	**Moist unit weight of soil**	*	*
γ_{sat}	**Saturated unit weight of soil**	**	**
γ_w	Unit weight of water	Deterministic	9.81 kN/m^3

* $\gamma = \gamma_w (G_s + 0.2e)/(1+e)$ (assume degree of saturation = 20% for "moist").
** $\gamma_{sat} = \gamma_w (G_s + e)/(1+e)$ (degree of saturation = 100%).
Assume specific gravity of solids = G_s = uniformly distributed [2.5, 2.7] and void ratio = e = uniformly distributed [0.3, 0.6].

Solution methods

FORM, SORM, simulation using MATLAB

Results (optional)

Deterministic (based on origin of standard normal space):
H = 5 m, h = 2.5 m, ϕ = 35°, θ = 20°, G_s = 2.6, e = 0.45 (γ_{sat} = 20.6 kN/m^3, γ = 18.2 kN/m^3)

$$FS = \frac{[\gamma(H-h) + h(\gamma_{sat} - \gamma_w)]\cos\theta\tan\phi}{[\gamma(H-h) + h\gamma_{sat}]\sin\theta} = 1.43$$

Probabilistic (based on distributions):

	FORM (EXCEL)	FORM	SORM	Simulation ($n = 10^6$)
β	1.426	1.426	1.544	1.579
p_f	0.0769	0.0769	0.0613	0.0572
% error in p_f	34.4	34.4	7.2	-

Code URL (optional)

http://jyching.twbbs.org/reliability_benchmark/Pfun_case1.m

References (optional)	Phoon K. K. (2008). *Reliability-Based Design in Geotechnical Engineering: Computations and Applications*, Chapter 1, Taylor & Francis, London.
Reviewers	By Jianye Ching and Yi-Hung Hsieh

Solution method	FOSM	FORM[(1)]	SORM[(2)]	MCS	Subsim[(3)]
β	1.083	1.426	1.544	1.578	1.571
P_F	0.1394	0.0769	0.0614	0.0573	0.0581[(4)]
% error in p_f	143.28	34.21	7.16	-	1.40
# of evaluation of P function (optional)	13	57	112	10^6	2800
Estimator cov (optional)	-	-	-	0.41%	10.68%[(5)]

(1) Gradient Projection algorithm is taken
(2) Algorithm by Der Kiureghian and Stefano (1991) is taken
(3) 1000 samples taken in each stage
(4) average of 100 runs of Subsim
(5) cov estimated from 100 runs of Subsim

By Kieu Le T.C. & Honjo Y., using Subset MCMC simulation in Fortran

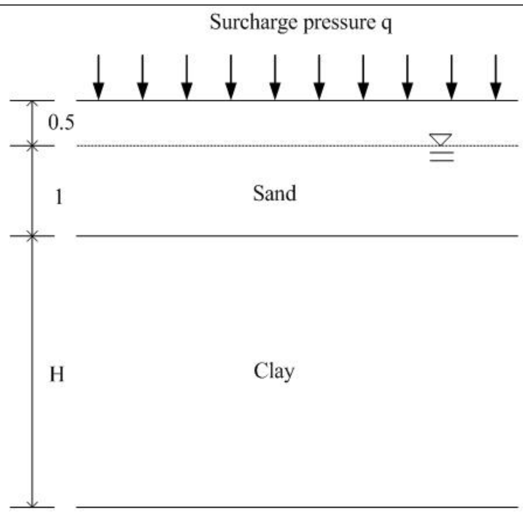
Solution method	Nt = 100	Nt = 50
p_f	0.02~0.1741	0.0199~0.1754
COV(p_f)	0~0.87	0.02~0.98
Log(mean p_f)/Log(p_{f_True}*)	0.61~1.37	0.61~1.37

* p_{f_True} is chosen as p_f obtained by simulation with Nt = 10^6 (KK Phoon)

The optimum input for Nt, Ns & Nf:
Nt = 50 ~ 100
Nf/Nt = 0.02~0.05
Ns/Nt = 0.02~0.04

geotechnical safety network

TASK GROUP 3 – RELIABILITY BENCHMARKING

Example No.	2
Author(s)	Jianye Ching and Yi-Chu Chen
Date	15 April 2008 (ver. 1)
Brief description	A consolidation problem with 8 independent random variables, modified from the example in P.372 in Ang and Tang (1984). The purpose of this benchmark example is to examine the robustness of each reliability method for problems with non-differentiable performance functions.
Figure	
Performance function	$P = S_{allow} - \frac{H}{1+e^{clay}} \times \left[C_r \log_{10}\left(\frac{\sigma'_p}{\sigma'_0}\right) + C_c \log_{10}\left(\frac{\sigma'}{\sigma'_p}\right) \right] \quad if\ \sigma' > \sigma'_p$ $P = S_{allow} - \frac{H}{1+e^{clay}} \times C_r \log_{10}\left(\frac{\sigma'}{\sigma'_0}\right) \quad if\ \sigma' < \sigma'_p$ S_{allow} = allowable settlement H = thickness of the clay layer e^{clay} = initial void ratio of the clay layer σ'_p = average pre-consolidation stress of the clay

$\sigma_0^{'}$ = average overburden stress of clay before construction

$\sigma^{'}$ = average overburden stress of clay after construction

C_r = re-compression index

C_c = compression index

where

$$\sigma_0^{'} = 0.5\gamma^{sand} + 1\left(\gamma_{sat}^{sand} - \gamma_w\right) + \frac{H}{2}\left(\gamma_{sat}^{clay} - \gamma_w\right)$$

$$\sigma^{'} = \sigma_0^{'} + q \qquad \sigma_p^{'} = OCR \cdot \sigma_0^{'} \qquad C_r = \alpha C_c$$

γ_w = unit weight of water (9.81 kN/m^3)

The moist soil unit weight γ and saturated soil unit weight γ^{sat} are not independent, because they are related to the specific gravity of the soil solids (G_s) and the void ratio (e):

$$\gamma_{sat}^{sand} = \gamma_w \frac{G_s^{sand} + e^{sand}}{1 + e^{sand}} \qquad \gamma_{sat}^{clay} = \gamma_w \frac{G_s^{clay} + e^{clay}}{1 + e^{clay}}$$

$$\gamma^{sand} = \gamma_w \frac{G_s^{sand} + 0.2e^{sand}}{1 + e^{sand}}$$

where a degree of saturation = 20% is assumed for "moist".

There are nine independent random variables in this problem (q, H, e^{clay}, e^{sand}, G_s^{clay}, G_s^{sand}, OCR, C_c and α)

Inputs

Variable	Description	Distribution	Statistics
S_{allow}	Allowable settlement	Deterministic	0.05 m
q	Surcharge pressure	Lognormal	Mean = 20 kN/m^2 cov = 20%
H	Thickness of clay	Gaussian	mean = 4 m cov = 10%
e^{clay}	Initial void ratio of clay	Lognormal	mean = 1.2 cov = 15%
e^{sand}	Void ratio of sand	Lognormal	mean = 0.8 cov = 15%
G_s^{clay}	Specific gravity of clay solids	Uniform	[2.5, 2.7]
G_s^{sand}	Specific gravity of sand solids	Uniform	[2.5, 2.7]
OCR	Over-consolidation ratio of clay	Uniform	[1.5, 2.5]
C_c	Compression index	Lognormal	mean = 0.4 cov = 25%
α	$\alpha = C_r/C_c$	Uniform	[0.1, 0.2]

Solution methods (optional)

FOSM, FORM, SORM, MCS, Subsim

Results (optional)

Deterministic (based on mean values):

q = 20 kN/m^2, H = 4m, e^{clay} = 1.2, e^{sand} =0.8, G_s^{clay} = 2.6, G_s^{sand} = 2.6, OCR = 2, C_c =0.4, α = 0.15

$$FS = S_{allow} \Big/ \left(\frac{H}{1+e^{clay}} \times C_r \log_{10}\left(\frac{\sigma'}{\sigma'_0} \right) \right) = 2.0935\ (2.1873)$$

Probabilistic (based on distributions):

Solution method	FOSM	FORM[(1)]	SORM[(3)]	MCS	Subsim[(4)]
β	2.935	1.431 2.149[(2)]	1.547 2.287	1.523	1.527
P_F	1.67e-3	7.62e-2 1.58e-2	6.09e-2 1.11e-2	6.39e-2	6.34e-2[(5)]
% error in p_f	-97.39	19.25 -75.27	-4.69 -82.63	-	-0.78
# of evaluation of P function (optional)	19	152	261	10^6	1900
Estimator cov (optional)	n/a	n/a	n/a	0.4%	10.2%[(6)]

(1) Gradient Projection algorithm is taken
(2) There are two solutions by FORM, the frequency of the real solution is 21%, the frequency of the fake solution is 79%
(3) Algorithm by Der Kiureghian and Stefano (1991) is taken
(4) 1000 samples taken in each stage
(5) average of 100 runs of Subsim
(6) cov estimated from 100 runs of Subsim

Code URL (optional)

Matlab code for the performance function can be download via http://jyching.twbbs.org/reliability_benchmark/Pfun_case2.m

References (optional)

1. Ang, A. H.-S. and Tang, W.H. (1984). Probability Concepts in Engineering Planning and Design. Volume II: Decision, Risk and Reliability.
2. Der Kiureghian, A. and Stefano, M.D. (1991). Efficient algorithm for second-order reliability analysis. ASCE Journal of Engineering Mechanics, 117(12), 2904-2923.

Reviewers

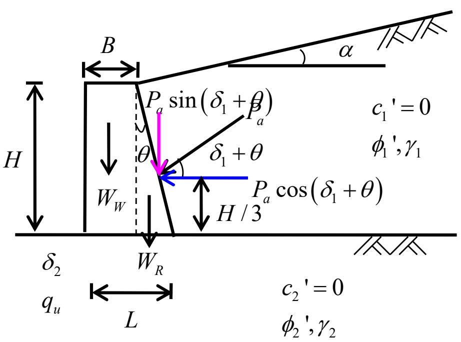
Kok-Kwang Phoon

	FORM (EXCEL)	FORM	SORM	Simulation ($n = 10^6$)
β	1.432 2.187*	1.451** 2.088	1.565 2.223	1.524
p_f	0.0761 0.0144	0.0734 0.0184	0.0588 0.0131	0.0638
% error in p_f	19.3 -77.4	15.1 -71.2	-7.8 -79.5	-

*wrong answer occurs more frequently in SOLVER
**correct answer occurs in MATLAB using origin as initial point

geotechnical safety network

TASK GROUP 3 – RELIABILITY BENCHMARKING

Example No.	3
Author(s)	Jianye Ching & Yi-Hung Hsieh
Date	16 September 2011 (ver. 2)
Brief description	A retaining wall problem with 4 independent random variables. We consider three performance functions including sliding, overturning and bearing capacity. The purpose of this benchmark example is to examine the robustness of each reliability method for problems with multiple failure modes.
Figure	
Performance function	$P=\min\begin{pmatrix}\left[P_a\sin(\delta_1+\theta)+W_W+W_R\right]\tan\delta_2-P_a\cos(\delta_1+\theta)\\ q_u\bar{L}-\left(W_W+W_R+P_a\sin(\delta_1+\theta)\right)\\ M_R-M_O\end{pmatrix}$ P_a = active force δ_1 = friction angle between the backfill and the back of the wall δ_2 = friction angle between the foundation soil and base of the wall ϕ_1 = friction angle of the backfill soil ϕ_2 = friction angle of the foundation soil W_W+W_R = the weight of the retaining wall B = top width of the retaining wall L = bottom width of the retaining wall α = slope of the backfill soil

q_u = bearing capacity of foundation soil
H = height of the retaining wall
θ = back angle of the retaining wall
M_R = resisting moment
M_O = overturning moment
where
$W_W = 23.58 \cdot B \cdot H$ (unit weight of the retaining wall: 23.58kN/m^3)
$W_R = 23.58 \cdot (L-B) \cdot H/2$
$\theta = \tan^{-1}\left[(L-B)/H\right]$
$\delta_1 = \frac{2}{3}\phi_1 \quad \delta_2 = \frac{2}{3}\phi_2 \quad P_a = \frac{1}{2}K_a\gamma_1 H^2$

$$K_a = \frac{\cos^2(\phi_1 - \theta)}{\cos^2(\theta)\cos(\delta_1+\theta)\cdot\left[1+\sqrt{\frac{\sin(\delta_1+\phi_1)\cdot\sin(\phi_1-\alpha)}{\cos(\delta_1+\theta)\cdot\cos(\theta-\alpha)}}\right]^2}$$

$$q_u = \frac{1}{2}\gamma_2 \bar{L} N_\gamma F_{\gamma i} \quad N_\gamma = 2(N_q+1)\tan\phi_2 \quad N_q = e^{\pi\tan\phi_2}\tan^2\left(45+\frac{\phi_2}{2}\right)$$

$F_{\gamma i} = (1-\beta/\phi_2)^2$ ($F_{\gamma i}$ from Hanna and Meyerhof (1981))
where β is the inclination angle of the total foundation loading:

$$\beta = \tan^{-1}\left(P_a\cos(\delta_1+\theta)/\left(P_a\sin(\delta_1+\theta)+W_W+W_R\right)\right)$$

$$M_R = W_W\cdot\frac{B}{2}+W_R\cdot\left(B+\frac{1}{3}(L-B)\right)+P_a\sin(\delta_1+\theta)\cdot\left(B+\frac{2}{3}(L-B)\right)$$

$$M_O = P_a\cos(\delta_1+\theta)\cdot\frac{H}{3}$$

$$\bar{L} = L-2e_L \quad e_L = \left|\bar{x}-\frac{L}{2}\right| \quad \bar{x} = \frac{M_R-M_O}{W_W+W_R+P_a\sin(\delta_1+\theta)}$$

There are four independent random variables in this problem ($\gamma_1, \gamma_2, \phi_1, \phi_2$)

Inputs

Variable	Description	Distribution	Statistics
B	Top width of the wall	Deterministic	2.5 m
H	Height of the wall	Deterministic	4 m
L	Bottom width of the wall	Deterministic	3.5 m
α	Slope of the backfill soil	Deterministic	5°
γ_1	Unit weight of backfill soil	Gaussian	mean = 19 kN/m^3 cov = 10%
γ_2	Unit weight of foundation soil	Gaussian	mean = 17 kN/m^3 cov = 10%
ϕ_1	Friction angle of backfill soil	Gaussian	mean = 35° cov = 10% (truncated at α, i.e.: PDF = 0 when $\phi_1 < \alpha$)
ϕ_2	Friction angle of foundation soil	Gaussian	mean = 35° cov = 10%

Solution methods

MCS

Results (optional)

Deterministic (based on mean values):

γ_1 = 19 kN/m^3, γ_2 = 17 kN/m^3, ϕ_1 = 35°, ϕ_2 =35°

$FS_1 = \left\{\left[P_a \sin(\delta_1+\theta)+W_W+W_R\right]\tan\delta_2\right\} / P_a \cos(\delta_1+\theta)$ =2.9425>1.5 (ok)

$FS_2 = q_u \bar{L} / \left(W_W + W_R + P_a \sin(\delta_1+\theta)\right)$=6.7214>3 (ok)

$FS_3 = M_R / M_O$ = 8.6919>2.5 (ok)

$FS_4 = (L/6)/e_L$ =2.3662 <L(ok)

Probabilistic (based on distributions):

Solution method	MCS
β	2.9083
P_F	0.0018
% error in p_f	-
# of evaluation of P function (optional)	10^6
Estimator cov (optional)	2.35%

Code URL (optional)

Matlab code for the performance function can be download via http://140.112.12.21/issmge/reliability_benchmark/Pfun_case3.zip

References (optional)

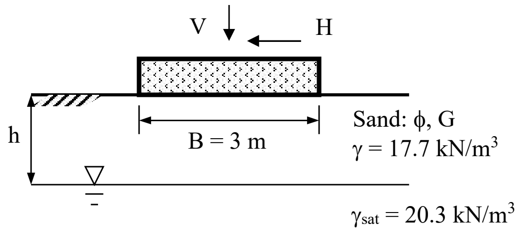
1. Hanna, A.M. and Meyerhof, G.G. (1981). Experimental evaluation of bearing capacity of footings subjected to inclined loads. Canadian Geotechnical Journal, 18(4), 599-603.

Reviewers

Kok-Kwang Phoon

geotechnical safety network

TASK GROUP 3 – RELIABILITY BENCHMARKING

Example No.	4
Author(s)	KK Phoon
Date	25 April 2008 (ver. 1)
Brief description	Footing problem subjected to inclined loading with 5 independent random variables. The purpose is to show that load variables can be both favourable and unfavourable in the same problem. The vertical load is unfavourable in the usual load context and favourable in the bearing capacity inclination factor.
Figure	

Performance function

$P = (0.5\ B\ \gamma^*\ N_\gamma\ \zeta_{\gamma s}\ \zeta_{\gamma i}\ \zeta_{\gamma r})\ B^2 - V$

$N_q = \exp(\pi \tan\phi) \tan^2(45^\circ + \phi/2)$
$N_\gamma = 2\ (N_q + 1) \tan\phi$

$\gamma' = \gamma_{sat} - \gamma_w = 20.3 - 9.81 = 10.5$ kN/m^3
γ^* (kN/m^3) $= \gamma = 17.7$ $\qquad h > B$
$= \gamma' + (\gamma - \gamma')h/B = 10.5 + 7.2h/B$ $\qquad B > h > 0$

$\zeta_{\gamma s} = 1 - 0.4\ (B/L) = 0.6$

$$\zeta_{\gamma i} = \left(1 - \frac{H}{V}\right)^{2.5}$$

Rigidity index, $I_r = G / (\sigma'_a \tan\phi)$

Reduced rigidity index, $I_{rr} = I_r / (1 + I_r\ \Delta)$

	$\Delta = 0.00025\,(45 - \phi)(\sigma'_a / 100\text{ kPa})$ (Note: ϕ in degrees) σ'_a (kPa) $= 0.5B\gamma = 26.55$ $h > B/2$ $= h\gamma + (0.5B-h)\gamma' = 7.2h + 5.25B$ $B/2 > h > 0$ $I_{rc} = 0.5\exp[(3.30 - 0.45\,B/L)\cot(45^o - \phi/2)]$ $I_{rr} > I_{rc} \Rightarrow$ General shear failure $I_{rr} < I_{rc} \Rightarrow$ Local/punching shear failure $\zeta_{\gamma r} = \exp\{[(-4.4 + 0.6\,B/L)\tan\phi] + [(3.07\sin\phi)(\log_{10} 2I_{rr})/(1 + \sin\phi)]\}$ $I_{rr} < I_{rc}$ $= 1$ otherwise h = depth of groundwater table below ground surface γ and γ_{sat} = moist unit weight and saturated unit weight of sand, respectively γ_w = unit weight of water (9.81 kN/m^3) ϕ = effective stress friction angle of sand G = shear modulus of sand V = vertical dead load H = horizontal live load There are five independent random variables in this problem (h, ϕ, G, V, H)
Inputs	see table below

Variable	Description	Distribution	**Statistics**
h	**Depth of water table**	**Lognormal**	mean = 2 m cov = 50%
ϕ	**Effective stress friction angle**	**Lognormal**	mean = 35^o cov = 8%
G	**Shear modulus**	**Lognormal**	mean = 20 MPa cov = 50%
V	**Vertical dead load**	**Normal**	mean = 1500 kN cov = 5%
H	**Horizontal live load**	**Extreme Type I**	mean = 150 kN cov = 20%
B	**Footing width**	**Deterministic**	3 m
γ	**Moist unit weight of soil**	**Deterministic**	17.7 kN/m^3
γ_{sat}	**Saturated unit weight of soil**	**Deterministic**	20.3 kN/m^3
γ_w	Unit weight of water	Deterministic	9.81 kN/m^3

$\gamma = \gamma_w (G_s + 0.2e)/(1+e)$ (assume degree of saturation = 20% for "moist") = 17.7 kN/m^3

$\gamma_{sat} = \gamma_w (G_s + e)/(1+e)$ (degree of saturation = 100%) = 20.3 kN/m^3

Assume specific gravity of solids = G_s = 2.6 and void ratio = e = 0.5

<table>
<tr><td>Solution methods</td><td>FORM, SORM, simulation using MATLAB</td></tr>
<tr><td>Results (optional)</td><td>

Deterministic (based on origin of standard normal space):
h = 1.8 m, ϕ = 35°, G = 18 MPa, V = 1500 kN, H = 155 kN

$FS = (0.5\,B\,\gamma^*\,N_\gamma\,\zeta_{\gamma s}\,\zeta_{\gamma i}\,\zeta_{\gamma r})\,B^2/V = 2.87$

Probabilistic (based on distributions):

<table>
<tr><th></th><th>FORM (EXCEL)</th><th>FORM</th><th>SORM</th><th>Simulation ($n = 10^6$)</th></tr>
<tr><td>β</td><td>2.612</td><td>2.593</td><td>2.653</td><td>2.664</td></tr>
<tr><td>p_f</td><td>0.0045</td><td>0.0048</td><td>0.0040</td><td>0.0039</td></tr>
<tr><td>% error in p_f</td><td>15.4</td><td>23.1</td><td>2.56</td><td>-</td></tr>
</table>

</td></tr>
<tr><td>Code URL (optional)</td><td>http://jyching.twbbs.org/reliability_benchmark/Pfun_case4.m</td></tr>
<tr><td>References (optional)</td><td></td></tr>
<tr><td>Reviewers</td><td>

Jianye Ching and Yi-Hung Hsieh

<table>
<tr><th>Solution method</th><th>FOSM</th><th>FORM(1)</th><th>SORM(2)</th><th>MCS</th><th>Subsim(3)</th></tr>
<tr><td>β</td><td>1.365</td><td>2.454</td><td>2.514</td><td>2.5121</td><td>2.5241</td></tr>
<tr><td>P_F</td><td>0.0861</td><td>0.0071</td><td>0.0060</td><td>0.0060</td><td>0.0058(4)</td></tr>
<tr><td>% error in p_f</td><td>13.35</td><td>18.33</td><td>0</td><td>-</td><td>-3.3</td></tr>
<tr><td># of evaluation of P function (optional)</td><td>11</td><td>176</td><td>217</td><td>10^6</td><td>2800</td></tr>
<tr><td>Estimator cov (optional)</td><td>n/a</td><td>n/a</td><td>n/a</td><td>1.29%</td><td>28.98%(5)</td></tr>
</table>

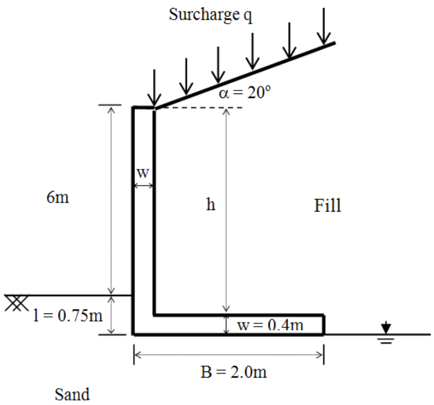
(1) Gradient Projection algorithm is taken
(2) Algorithm by Der Kiureghian and Stefano (1991) is taken
(3) 1000 samples taken in each stage
(4) average of 100 runs of Subsim
(5) cov estimated from 100 runs of Subsim

Der Kiureghian, A. and Stefano, M.D. (1991). Efficient algorithm for second-order reliability analysis. ASCE Journal of Engineering Mechanics, 117(12), 2904-2923.

</td></tr>
</table>

c

TASK GROUP 3 – RELIABILITY BENCHMARKING

Example No.	5
Author(s)	T.C. Kieu Le & Y. Honjo
Date	November 27, 2008
Brief description	The procedure McDeva based on the combination of Design Value Method, and using Subset Markov Chain Monte Carlo Simulation (Subset MCMC), has been used to determine partial factors for a gravity retaining wall under sliding failure mode. The optimum combination of some parameters inputted into Subset Subset MCMC is also proposed.
Figure	
Performance function	$P = R - S$ Where: $R = \frac{1}{2} l^2 \gamma_s^{'} \tan^2\left(45^o + \frac{\phi_s^{'}}{2}\right) + \left\{(h+B) w \gamma_c + \left[(B-w)h + \frac{1}{2}(B-w)^2 \tan\alpha\right]\gamma_f^{'} + q\frac{(B-w)}{\cos\alpha}\right\}\tan\phi_{bs}^{'}$ $S = \frac{1}{2}\left[(B-w)\tan\alpha + h + w\right]^2 \gamma_f^{'} \tan^2\left(45^o - \frac{\phi_f^{'}}{2}\right)\cos\alpha$ Fill soil behind the wall and soil beneath the wall are sand. Properties of sand beneath the wall: internal friction angle $\phi_s^{'}$, unit weight $\gamma_s^{'}$ Properties of fill behind the wall: internal friction angle $\phi_f^{'}$, unit

weight γ'_f
q: surcharged load behind the wall
Groundwater level is at depth below the base of the wall
Thickness of retaining wall is w.

Inputs

Variable	Description	Distribution	Statistics
γ'_f (kN/m^3)	Unit weight of fill sand above the wall	Lognormal	Mean = 20 COV = 0.05
γ'_s (kN/m^3)	Unit weight of sand beneath the wall	Lognormal	Mean = 19 COV = 0.05
γ_c (kN/m^3)	Unit weight of the wall	Lognormal	Mean = 25 COV = 0.05
$tan\phi'_f$	Tangent of internal friction angle of fill sand	Lognormal	Mean = 0.781 COV = 0.27
$tan\phi'_s$	Tangent of internal friction angle of sand beneath the wall	Lognormal	Mean = 0.675 COV = 0.26
$tan\phi'_{bs}$	Tangent of friction angle between the wall and sand beneath	Lognormal	Mean = 0.577 COV = 0.24
q (kN/m^2)	Surcharge load	Lognormal	Mean = 15 COV = 0.3
w (m)	Thickness of the wall	Deterministic	0.4
B (m)	Width of wall base	Deterministic	2.0
α (°)	Inclination angle of fill behind the wall	Deterministic	20
h (m)		Deterministic	6.35
l (m)		Deterministic	0.75

Solution methods (optional)

* **McDeva** (Markov Chain Monte Carlo simulation based on Design Value Method)
Performance function has the form of $P(\mathbf{X}) = R(\mathbf{X}) - S(\mathbf{X})$
The resistance component $R(\mathbf{X})$ and the load component $S(\mathbf{X})$ which are combinations of basic variables X_i, $(i = 1,\ldots, n)$ are also random variables.

Step 1: Define PDF's and probabilistic parameters of basic variables X_i $(i = 1,\ldots, n)$

Step 2: Carry out Subset MCMC in N_{run} times to obtain the failure probability p_f, its standard deviation and the location of the design points.
If one can group candidate design points to several groups, it may suggest there are several failure modes, and therefore

several design points.

Step 3: Among N_{run} design points obtained in the previous step, one should choose the best estimated design points $A\left(X_1^*,...,X_n^*\right)$.

Step 4: From $A\left(X_1^*,...,X_n^*\right)$, the corresponding estimated design point in two dimensional co-ordinate *(R,S)*, i.e. (R^*,S^*), may also be obtained. Then the sensitivity factors (α_R and α_S) and load and resistance factors (γ_R and γ_S) may be calculated.

Verification of the Robustness of McDeva by Ordinary Monte Carlo (OMC) and obtain the optimum combination of N_t, N_s and N_f

Where:
- The total samples generated for each subset: N_t
- The number of seeds used for generating a subset: N_s
- The minimum number of samples falling in the failure region so that the Subset MCMC algorithm will be satisfied and stopped (*cut-off criteria*): N_f

McDeva has been carried out with different combination of N_t, N_s, and N_f as follows:

N_t = 50, 100, 150
N_s/N_t = 0.02, 0.04, 0.10, 0.20, 0.50
N_f/N_t = 0.02, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50

Results (optional)

*** Probability of failure (P_f) and reliability index β**

	McDeva	OMC
β	2.70	2.76
Probability of failure (p_f)	0.25×10^{-2}	0.29×10^{-2}
COV(p_f)	1.4	

*** Load and resistance factors**

	McDeva[(1)]				OMC[(2)]			
	Normal		Lognormal		Normal		Lognormal	
	R	S	R	S	R	S	R	S
α	-0.54	0.84	-0.54	0.84	-0.50	0.87	-0.50	0.87
γ	0.63	1.71	0.64	1.79	0.65	1.76	0.65	1.86

(1) 1000 runs of Subset MCMC
(2) 1,000,000 samples taken in each stage

*** Optimum combination of N_t, N_s, and N_f:**
N_t = 50 to 150, N_s/N_t = 0.1 to 0.2, N_f/N_t = 0.1 to 0.2

Code URL (optional)

References (optional)

Kieu Le, T.C. (2008). *Code Calibration Procedure Based on Monte Carlo Simulation for Geotechnical Design.* PhD Thesis, Gifu University, Japan

Reviewers

Jianye Ching

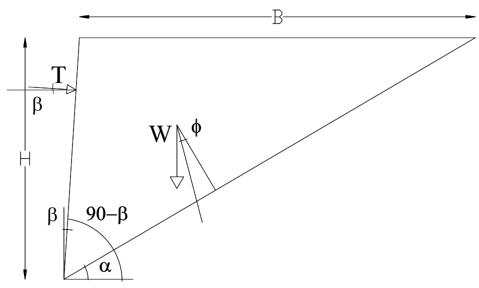
Solution method	FOSM	FORM*	SORM**	MCS	Subsim***
β	0.9428	0.9555	0.9812	0.9978	0.9943
p_f	0.1729	0.1697	0.1633	0.1592	0.1600****
% error in p_f	8.61	6.60	2.58	-	0.5025
# of evaluation of P function (optional)	15	120	191	10^6	1000
Estimator cov(optional)	-	-	-	0.23%	7.21%*****

* Gradient Projection algorithm is taken
** Algorithm by Der Kiureghian and Stefano (1991) is taken
*** 1000 samples taken in each stage
**** average of 100 runs of Subsim
***** cov estimated from 100 runs of Subsim

Der Kiureghian, A. and Stefano, M.D. (1991). Efficient algorithm for second-order reliability analysis. ASCE Journal of Engineering Mechanics, 117(12), 2904-2923.

geotechnical safety network

TASK GROUP 3 – RELIABILITY BENCHMARKING

Example No.	7
Author(s)	C. Cherubini & G. Vessia
Date	24/11/2008
Brief description	The equilibrium of a sliding wedge stabilized by means of an anchorage is investigated in terms: 1)Factor of safety: $FS^{+}=\frac{W\cos(\alpha)\tan(\varphi)+T\sin(\alpha+\beta)\tan(\varphi)+T\cos(\alpha+\beta)}{W\sin(\alpha)}$ (1) $FS^{-}=\frac{W\cos(\alpha)\tan(\varphi)+T\sin(\alpha+\beta)\tan(\varphi)}{W\sin(\alpha)-T\cos(\alpha+\beta)}$ (2) 2)Probability of failure by means of reliability approaches: $P=W\cos(\alpha)\tan(\varphi)+T\sin(\alpha+\beta)\tan(\varphi)+T\cos(\alpha+\beta)-W\sin(\alpha)$ (3) 2)LRFD applied to limit state design: $W\cos(\alpha)\tan(\varphi)+T\sin(\alpha+\beta)\tan(\varphi)+T\cos(\alpha+\beta)-W\sin(\alpha)>0$ (4) The purpose is to evaluate how much variability of geotechnical design variables affects the stability estimation carried out by means of deterministic approach and the partial factor approach applied to the limit state design.
Figure	
Performance function	$P=W\cos(\alpha)\tan(\varphi)+T\sin(\alpha+\beta)\tan(\varphi)+T\cos(\alpha+\beta)-W\sin(\alpha)$

where $W = wedge\ weight$ is computed as follow:

$$W = A \cdot \gamma = \frac{H \cdot B}{2}\gamma = \frac{H\left(H\tan\left(\frac{\pi}{2}-\alpha\right) - H\tan(\beta)\right)}{2}\gamma$$

$\alpha = sliding\ surface\ slope$
$\beta = rock\ face\ slope$
$\gamma = unit\ weight$
$\phi = internal\ friction\ angle\ of\ rock\ mass$
$T = anchorage\ pull$
$H = slope\ height$
$B = wedge\ width$
$A = wedge\ area$

Inputs

Variable	Distribution type	Variation coefficient	Min value	Max value
α [°]	Normal	10%		
β [°]	Uniform		5	10
ϕ [°]	Lognormal	30, 40, 50%		
γ [kN/m^3]	Normal	2%		
T [kN]	Constant	-		
H[m]	Constant	-		
B[m]	Constant	-		

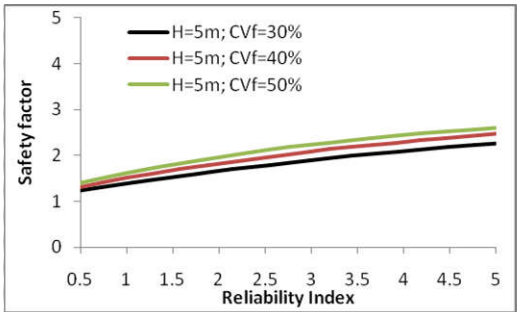
Variable	Mean value	Standard deviation	Characteristic value
α [°]	40	4	38
β [°]	7.5	-	7.5
ϕ [°]	35	10.5	29.8
γ [kN/m^3]	23	0.5	22.8
H[m]	5 10 15		
B[m]	5.1 10.2 15.2		

Solution methods (optional)

First Order Reliability Method (FORM), Second Order Reliability Method (SORM) and Monte Carlo Simulation Adaptive sampling (ASAM) implemented in COMREL code (1997).

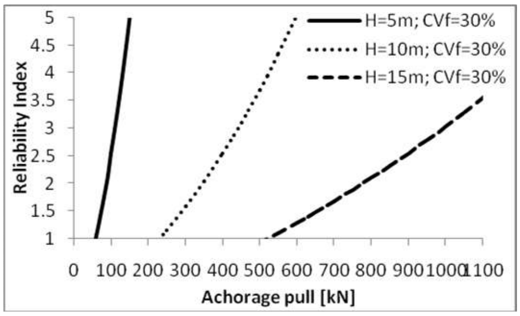
Results (optional)

1) Safety factor FS$^+$ versus reliability index (FORM) for different coefficients of variation of rock mass internal friction angle

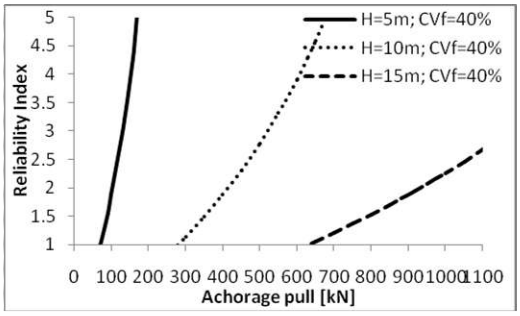


2) Reliability index (FORM) versus anchorage pull for:

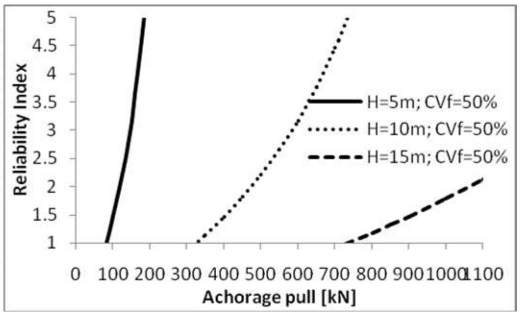
CV_ϕ=30%



CV_ϕ=40%

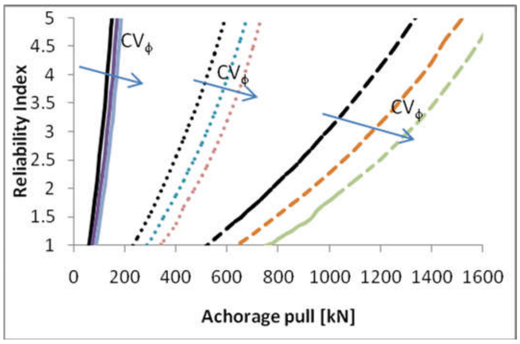


CV_ϕ=50%

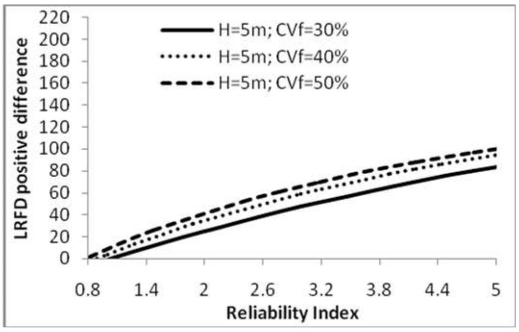


<table>
<tr><td></td><td>

3) Reliability index (MCS) versus anchorage pull for: CV_ϕ=30, 40, 50% and H=5m (continuous line); H=10m (dotted line); H=15m (dashed line).



4) LRFD positive difference calculated according to Italian Combination2 for global equilibrium condition (see Eq. 4) versus reliability index for slope height H=5m and CV_ϕ=30%, 40% and 50%.



</td></tr>
<tr><td>Code URL (optional)</td><td></td></tr>
<tr><td>References (optional)</td><td>COMREL 1997. Reliability Consulting Programs RCP GmbH, Munchen, Germany.</td></tr>
<tr><td>Reviewers</td><td></td></tr>
</table>

<table>
<tr><td>Example No.</td><td></td></tr>
<tr><td>Authors(s)</td><td>Dianqing Li., Xiaosong Tang</td></tr>
<tr><td>Date</td><td>January 23, 2009</td></tr>
<tr><td>Brief description</td><td>For the concrete dam shown in Fig. 1, three different failure modes may occur as follows (Novak et al. 2001): sliding (failure mode 1), overstress at upstream heel (failure mode 2), and overstress at downstream toe (failure mode 3). For illustrative purpose, only the normal operation is considered herein. With the identified failure modes, performance functions can be formulated for each stability mode as follows.</td></tr>
<tr><td>Figure</td><td>

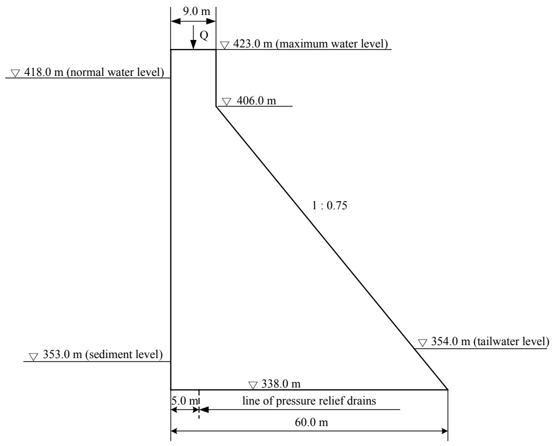


Figure 1. Main cross-section of concrete dam</td></tr>
<tr><td>Performance function</td><td>

$$g_1 = \left(2499\gamma_h + Q + \left(0.375H_x^2 - 30\alpha H_s - 27.5H_x - 2.5H_s\right)\gamma_w\right) f' + 60c' - 0.5\gamma_w\left(H_s^2 - H_x^2\right) - 0.5\gamma_n H_n^2 \tan^2(45^\circ - \theta_n / 2) \quad (1)$$

$$g_2 = \sigma_t + 85.7225\gamma_h - 0.0003H_s^3 - 0.9168\alpha H_s + 0.0003H_x^3 - 0.0102H_x^2 + 0.9168\alpha H_x - H_x - 0.0003\gamma_n H_n^3 \tan^2(45^\circ - \theta_n / 2) + 0.0592Q - 10.381 \quad (2)$$

$$g_3 = \sigma_c + 2.4225\gamma_n - 0.0003H_s^3 - 0.0832\alpha H_s + 0.0003H_x^3 - 0.0227H_x^2 - 0.0832\alpha H_x + H_x - 0.0003\gamma_n H_n^3 \tan^2(45^\circ - \theta_n / 2) + 0.0258Q - 4.9644 \quad (3)$$

</td></tr>
<tr><td>Inputs</td><td>

Table 1. Summary statistics of basic random variables in the dam stability model

Variable	Description	Mean	Standard deviation	COV	Distribution

</td></tr>
</table>

	H_s (m)	Water level for upstream face	80.00	4.80	0.06	Normal
	H_x (m)	Water level for downstream face	16.00	0.96	0.06	Normal
	H_n (m)	Sediment depth	15.00	2.25	0.15	Normal
	γ_h(kN/m^3)	Unit weight of dam concrete	24	0.72	0.03	Normal
	γ_n(kN/m^3)	Submerged unit weight of sediment	12	0.60	0.05	Normal
	θ_n (^{0})	Angle of shearing resistance of sediment	15	1.50	0.10	Extreme type I
	α	Reduction coefficient of uplift pressure	0.40	0.10	0.25	Lognormal
	c'(kN/m^2)	Cohesion	62	21.70	0.35	Lognormal
	f'	Friction coefficient	1.00	0.30	0.30	Lognormal
	σ_c(kN/m^2)	Maximum allowable compressive stress of dam concrete	9000	1800	0.20	Lognormal
	σ_t(kN/m^2)	Maximum allowable tension stress of dam concrete	1000	250	0.25	Lognormal
	Q (kN)	Vertical live load	350	122.50	0.35	Extreme type I

Solution methods

The system reliability analyses are performed using the following system reliability methods: the first order multinormal (FOMN) (Hohenbichler and Rackwitz 1983), Cornell's bound (Cornell 1967), Ditlevsen's bound (Ditlevsen 1979), Adaptive importance sampling (AIS) (Melchers 1989), Radius-based importance sampling (ISAMF) (Harbitz 1986), and Monte Carlo simulation (MCS).

Results

Table 2. System reliability indexes of concrete dam using different methods

Solution method	FOMN	Cornell's bound	Ditlevsen's bound	AIS (confidence level=0.95)	ISAMF ($n=10^4$)	MCS ($n=10^6$)
β	2.186	2.172~2.308	2.185	2.177	2.197	2.188
P_f	0.0144	0.0105~0.0149	0.0144	0.0147	0.0140	0.0143
% error in P_f	0.7	-26.6~4.2	0.7	2.8	-2.1	-

References

Cornell, C. A. Bounds on the reliability of structural systems. Journal of Structural Division, 1967, 93(1): 171-200.

Ditlevsen, O. (1979). Narrow reliability bounds for structural systems. Journal of Structural Mechanics, 7(4): 453-472.

Harbitz, A. (1986). An efficient sampling method for probability of failure calculation, Structural Safety, 3: 109-115.

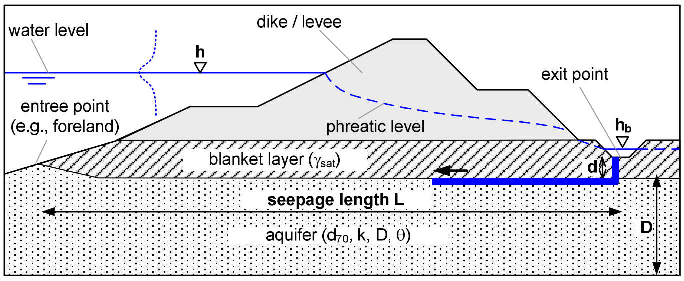
	Hohenbichler, H., Rackwitz, R. (1983). First-order concepts in system reliability. Structural Safety, 1: 177-188. Melchers, R. E. (1989). Importance sampling in structural systems. Structural Safety, 6(1): 3-10. Novak, P., Moffat, A. L. B., Nalluri, C., et al. (2001). Hydraulic Structures, third edition. Taylor & Francis Group.

geotechnical safety network

TASK GROUP 3 – RELIABILITY BENCHMARKING

Example No.	9
Author(s)	Timo Schweckendiek (Deltares & Delft University of Technology, NL)
Date	30 August 2010 (ver. 1)
Brief description	Piping (under-seepage) problem using the (revised) Sellmeijer model. The purpose of this benchmark example is to compare the performance of reliability methods for and internal erosion problem. Furthermore, the relative contribution of each random variable to the total uncertainty is illustrated by means of importance factors (FORM).

Figure



Performance function

$$P = m_p H_c - (h - h_b - 0.3d)$$

Where

$$\frac{H_c}{L} = F_1 F_2 F_3$$

$$F_1 = \frac{\gamma'_p}{\gamma_w}\{\eta \tan(\theta)\}\left(\frac{RD}{RD_m}\right)^{0.35}$$

$$F_2 = \frac{d_{70_m}}{\sqrt[3]{k\,L}}\left(\frac{d_{70}}{d_{70_m}}\right)^{0.4}$$

$$F_3 = 0.91\left(\frac{D}{L}\right)^{\frac{0.28}{\left(\frac{D}{L}\right)^{0.28}-1}+0.04}$$

H_c	critical head difference [m]
γ'_p	effective vol. weight of submerged sand grains [kN/m^3]
γ_w	volumetric weight of water [kN/m^3]
θ	bedding angle of sand grains [º]
η	White's coefficient [-]
RD	relative density [-]
RD_m	reference value relative density (=0.725) [-]
k	permeability of the aquifer [m/s]
d_{70}	70-percentile grain distribution (piping-sensitive layer) [m]
d_{70_m}	reference value of d_{70} (=2.08 e^{-4}) [m]
D	thickness of the sand layer [m]
L	seepage length [m]
m_p	model (uncertainty) factor piping [-]
h	water level, at entry point [m+REF]
h_b	phreatic level at exit point [m+REF]
d	thickness of the blanket layer

The performance function is based on the revised Sellmeijer model for piping (internal erosion, under-seepage), see Knoeff et al. (2009), as it is currently used in the Netherlands. The original Sellmeijer model is described in Sellmeijer (1988). Note that this performance function only considers piping. Often, piping is considered in combination with uplift (of the blanket layer). Together, these mechanisms form can be characterized as a parallel system (i.e., uplift AND piping have to occur for system failure).

Inputs

Variable	**Description**	**Distribution**	**Statistics** (m=mean)
γ'_p	vol. weight grains	Deterministic	17 kN/m^3
γ_w	vol. weight water	Deterministic	9.81 kN/m^3
θ	bedding angle*	Deterministic	37 deg
η	White constant*	Deterministic	0.25
RD	relative density	Normal	m=0.7, cov=10%
RD_m	reference value RD*	Deterministic	0.725
k	permeability aquifer	Lognormal	m=1e-5m/s, cov=1
d_{70}	70-percentile g.s.d.	Lognormal	m=2e-4m, cov=15%
d_{70_m}	reference value d_{70}*	Deterministic	2.08e-4 m
D	thickness aquifer	Normal	m=15.0m, cov=10%
L	seepage length	Normal	m=25.0m, cov=5%
m_p	model factor piping	Lognormal	m=1.0, cov=12%
h	water level entry point	Gumbel	a=1.839, b=0.152
h_b	phreat. level exit point	Normal	m=-1.0m, s=0.1m
d	thickn. blanket layer	Lognormal	m=3.0m, cov=30%

* Note that the uncertainty of the calibrated model parameters is all included in the model factor m_p.

The standard deviations and variation coefficients of most variables are based on default values used in the FLORIS project (Van der Most & Wehrung, 2005).

Solution methods (optional)

FORM (Excel): based on Low & Tang (2007); SORM, MCS, IS, DS, Subsim: using FERUM v4.0, see Sudret & Der Kiureghian, (2000).

The m-file Pfun_case9.m works with PROLIB and produces virtually the same results as FERUM

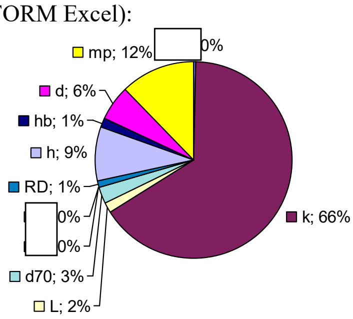
Results (optional)

Method	β	p_f	**error in** p_f
FORM (EXCEL)	3.19	7.2e-4	-4.4%
FORM (n = 106)	3.19	7.2e-4	-4.4%
SORM (n = 106+54 = 160)	3.18	7.5e-4	-0.5%
Simulation (MCS) (n = 10^6)	3.17	7.5e-4	0.0% (cov(p_f)=3.7%)
Importance Sampling[1] (IS) (n = $5*10^3$)	3.19	7.1e-4	-5.3% (cov(p_f)=3.7%)
Directional Sampling (DS) (n = 10^6, 10^5 directions)	3.19	7.2e-4	-3.7%
Subset Simulation (Subsim)[2] (n = $2*10^4$)	3.19	7.1e-4	-5.1%
n = number of performance function evaluations			

[1] importance sampling around design point by FORM
[2] Subsim: 5000 samples per stage

Importance factors (obtained with FORM Excel):

Variable X_i	α_i	α_i^2
RD	0.11	0.01
k	**-0.81**	**0.66**
d_{70}	0.17	0.03
D	-0.05	0.00
L	0.12	0.02
m_p	**0.35**	**0.12**
h	**-0.30**	**0.09**
h_b	0.12	0.01
d	0.24	0.06



The importance factors clearly reveal that the problem is dominated by the uncertainty in de permeability of the aquifer. Besides the model factor, the water level and the thickness of the blanket layer, the remaining uncertainties are practically irrelevant.

All tested reliability methods give answers within roughly 5% error with respect to the result obtained by MCS. Remarkably, the FORM result is very close to MCS, too. The performance function is linear in most of the important random variables except for *k*, the most important one. Thus, the low number of required performance function evaluations makes FORM attractive for this type of problem. Note that obtaining the right answer with Subsim requires either prior knowledge

	of the problem or trial and error (and thus prior knowledge of the p_f) or very generally robust settings (so far not found by the author).
Code URL (optional)	http://jyching.twbbs.org/reliability_benchmark/Pfun_case9.m
References (optional)	Knoeff, H., Sellmeijer, J.B., Lopez, J. & Luijendijk, S. (2009). Hervalidatie Piping (SBW HP1 + HP1.2, in Dutch), Deltares report 1200187-015-GEO-0004. Low, B.K. & Wilson H. Tang (2007). Efficient spreadsheet algorithm for first-order reliability method. Journal of Engineering Mechanics, ASCE, Vol. 133, No. 12, 1378-1387. Sellmeijer, J.B. (1988). On the mechanism of piping under impervious structures. PhD thesis. Delft University of Technology. Sudret, B. & Der Kiureghian, A. (2000). Stochastic Finite Element Methods and Reliability, A State-of-the-Art Report, Report No. UCB/SEMM-2000/08, Department of Civil and Environmental Engineering, University of California, Berkeley. Van der Most, H. & Wehrung, M. (2005). Dealing with Uncertainty in Flood Risk Assessment of Dike Rings in the Netherlands. Natural Hazards 2005 (36), p. 191-206.
Reviewers	Jianye Ching

B.2. GEOSNet example 9 correction

The performance function of GEOSNet example 9 needs some revisions as the following:

$$F_1 = \eta(\frac{\gamma_s}{\gamma_w} - 1)tan\theta \tag{B.1}$$

$$F_2 = \frac{d_{70m}}{\sqrt[3]{\frac{\nu kL}{g}}}(\frac{d_{70}}{d_{70m}})^{0.4} \tag{B.2}$$

$$F_3 = 0.91(D/L)^{\frac{0.28}{(D/L)^{0}.28-1}+0.04} \tag{B.3}$$

$$H_{c,p} = F_1 F_2 F_3 L \tag{B.4}$$

with:

ν	: dynamic viscosity of water at 10 deg. (1.33x10^{-6} Pa.s)
γ_s	: volumetric weight of sand grains (=26.5 kN/m^3)
γ_w	: volumetric weight of water (=10 kN/m^3)
θ	: bedding angle (deg.)
D	: thickness of the aquifer (m)
η	: drag factor coefficient
d_{70}	: 70%-fractile of the grain size distribution (m)
d_{70m}	: reference value for d_{70} (m)
g	: gravitational constant (=9.81 m/s^2)
k	: permeability of the aquifer (m/s)