Reliability Methods for Finite Element Models

Reliability Methods for Finite Element Models

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.dr.ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen

op woensdag 1 april 2009 om 12:30 uur

door

Mohammadreza RAJABALINEJAD
Master in Civil Engineering
Iran University of Science and Technology
geboren Tehran, Iran

Dit manuscript is goedgekeurd door de promotor:
Prof.ir. J.K. Vrijling
Copromotor: Dr.ir. P.H.A.J.M. van Gelder

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof.drs.ir. J.K. Vrijling	Technische Universiteit Delft, promotor
Dr.ir. P.H.A.J.M. van Gelder	Technische Universiteit Delft, copromotor
Prof.dr.ir. F.B.J. Barends	Technische Universiteit Delft
Prof.dr. F. Nadim	Norwegian Geotechnical Institute, Norway
Prof.dr. A. Noorzad	Power and Water University of Technology, Iran
Prof.dr.ir. M.A. Gutirrez	Technische Universiteit Delft
Dr.ir. L.E. Meester	Technische Universiteit Delft
Dr.ir. P. Waarts	TNO, the Netherlands

ISBN 978-1-58603-991-2

Key words: reliability, probabilistic, dynamic bounds, monotonic, monotonicity, Bayesian, Monte Carlo, flood defence, flood defense, dike, finite element.

Cover picture: a typical part of the Dutch dike, divided in sections and modeled by finite elements.

Published and distributed by IOS Press under the imprint Delft University Press

Publisher
IOS Press, Nieuwe Hemweg 6b, 1013 BG Amsterdam, The Netherlands
tel: +31-20-688 3355, fax: +31-20-687 0019
www.iospress.nl, www.dupress.nl
email: info@iospress.nl

PRINTED IN THE NETHERLANDS

In the name of who creates and owns souls and thoughts,
the most precious things that one can think about.
A. Ferdousi, 970 AC

This book is dedicated to
Ali and *Ghadamkheir*, my wonderful parents
Hossein, my lovely son
and *Maryam*

Contents

Summary

Probabilistic techniques in engineering problems are needed because they provide a deeper understanding of failure mechanisms and occurrence probabilities than deterministic techniques. In addition, they draw our attention to the consequences of failure at an early stage in the design process. However, to achieve these advantages, a well-defined model of the structure together with a robust reliability technique is needed, as also advocated for instance by Haldar and Mahadevan (2000). On the other hand, complex engineering problems with complicated boundary conditions usually are analysed with the finite element technique as presented in Smith and Griffiths (2004); Rajabalinejad et al. (2007a). The finite element method provides an implicit approximation to the limit state equation (LSE) that is far more accurate than other approaches. Therefore, if one wants to have the full advantage of the probabilistic approach one needs both an advanced model and a supporting reliability technique.

For the reliability analysis of engineering structures a variety of methods is known, of which Monte Carlo simulation is widely considered to be among the most robust and most generally applicable. The absence of systematic errors and the fact that its error analysis is well-understood are properties that many competing methods lack. A drawback is the often large number of runs needed, particularly in complex models, where each run may entail a finite element analysis or other time consuming procedure. Variance reduction methods may be applied to reduce simulation cost. This study describes methods to reduce the simulation cost even further, while retaining the accuracy of Monte Carlo, by taking into account widely-present monotonicity in limit state equations or other prior information.

This dissertation focuses on problems where a highly accurate estimate of the failure probability is required, but an explicit expression for the limit state equation is unavailable and the limit state equation can only be evaluated without loss

of accuracy via finite element analysis or some other time consuming process.

1

Introduction

1.1 Motivation

Probability and Statistics provide a wonderful tool to describe our lack of knowledge of the modeling and prediction. They courteously indicate our handicap in science. Besides, probabilistic assessments are more complete than commonly-used deterministic approaches in which we just rely on safety factors. At first, it may seem that the safety factor is good enough for the design process, but if we are interested in a very low probability of failure for some infrastructures[1], the safety factor lacks to help. In other words, given large consequences of failure, the reliability assessment becomes much more important. Moreover, the reliability assessment is a joint effort among different engineering fields, and it becomes more and more important to engineers, policy makers and people. In addition, it is interesting to see that how a multidisciplinary science is developed in different manners.

To facilitate the reliability assessment of complicated structures or models, two promising approaches are introduced in this dissertation: *Dynamic Bounds* (DB) and *Bayesian Monte Carlo* (BMC). DB dramatically reduces calculation efforts when there is monotonicity in a model with a limited number of influential variables. Bayesian Monte Carlo is a robust method which takes into account prior information.

The dynamic bounds method is applied to a part of the flood defence system in New Orleans. This case study is the 17th Street Flood Wall which failed during Hurricane Katrina. As a result, a complete process of the reliability assessment of a complicated structure is presented.

[1] For instance, we require $p_f \leq 1E-4/yr$ for flood defences along the Dutch sea coast, where p_f is the probability of failure.

1.2 Outline of the thesis

After the introduction, Chapter 2 reviews the reliability methods. In this chapter an overview of the most applicable ones is presented. Therefore, they are discussed in concept and compared in scope and power. Also references are provided for further details.

Chapter 3 describes *dynamic bounds* in reliability assessment. The concept of dynamic bounds is a simple concept which can facilitate the reliability analysis. It is based upon the fact that normally a stable structure remains stable when driving forces are reduced or the resistant parameters are increased. Apart from the fact that dynamic bounds speed up the whole process, they can be stored and used for the next series of simulations.

Chapter 4 gets along with Chapter 3 and suggests *Improved Dynamic Bounds*. It shows that the concept of dynamic bounds is like the concept of Monte Carlo which can be extended by taking the advantages of more properties. The suggested technique assumes the order of response of a limit state equation.

Chapter 5 suggests *Bayesian Monte Carlo* which is another robust approach towards facilitating the Monte Carlo process and speeding up the whole process. The concept is also simple; we try to get a judgement tool by the Bayesian methods to regulate the simulation process according to a required accuracy. It also makes possible to integrate the prior knowledge of the model as well as the prior information of Monte Carlo simulations which becomes gradually available during the calculations.

Chapter 6 presents a case study in which the probabilistic finite element is integrated with dynamic bounds. This case study is 17th Street Flood Wall which failed to protect New Orleans against Hurricane Katrina in 2005. The results are compared with the classical Monte Carlo method as an accurate base.

Chapter 7 contains conclusions, recommendations, and suggestions for future research projects.

2

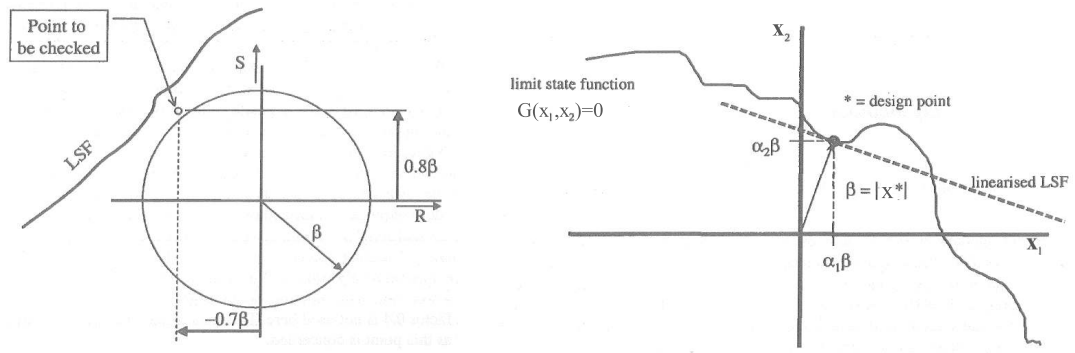
Probabilistic methods

2.1 Overview

Probabilistic techniques in engineering problems are needed because they provide a deeper understanding of failure mechanisms and occurrence probabilities than deterministic techniques. In addition, they draw our attention to the consequences of failure at an early stage in the design process. However, to achieve these advantages, a well-defined model of the structure together with a robust reliability technique is needed as advocated for instance by Haldar and Mahadevan (2000). On the other hand, complex engineering problems with complicated boundary conditions usually are analysed with finite element techniques (FE) presented for instance in Smith and Griffiths (2004); Rajabalinejad et al. (2007a). It provides an implicit approximation to the limit state equation (LSE) that is far more accurate than other approaches. Therefore, if one wants to have the full advantage of the probabilistic approach one needs both an advanced model and a supporting reliability technique. Known reliability techniques can be classified in three levels according to their performance.

Level I reliability methods compute only whether or not the reliability is sufficient, rather than computing the real probability of failure. Incorporating this method, Eurocode suggests that the limit state equation be checked in a standard normal space, in the point $(-0.7\beta,\ 0.8\beta)$ for resistance R and load S, respectively, as depicted in Figure 2.1(a) (see de Normalisation (1994)). With more than two variables, this principle is applied to the dominant variables as suggested by de Normalisation (1994).

The standard *level II* method is the so-called *First Order Reliability Method* (FORM), presented in Hasofer et al. (1973), in which the variables are assumed to be independent. The LSE is linearized in the *design point*, the point where the likelihood of failure is largest. In standardized space, this is the point on



(a) Level I (b) Level II

Figure 2.1: (a) Check point in a two dimensional standard normal space for a level I reliability method (from de Normalisation (1994)). (b) Design point and influence factors (α_1 and α_2) for the First Order Reliability Method Waarts (2000), in a two dimensional standard normal space.

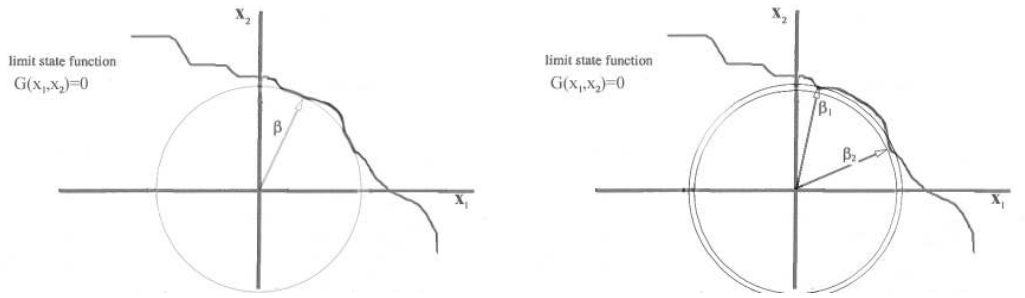
the LSE closest to the origin, see Figure 2.1(b). The failure probability is then computed by replacing the LSE with its linearization. The contribution of each of the variables to the failure can be estimated by the so-called α-factor, which represents how much that variable influences the location of the design point (see Figure 2.1(b)). According to Bjerager (1990), this method provides sufficiently accurate results when the limit state function is smooth and the number of variables is smaller than 50, and it is widely applied. However, in principle, FORM can only be used for Gaussian variables and the conversion of problems with other distribution types will generally introduce additional nonlinearity. In order to improve the accuracy of FORM in the design point, a quadratic approximation to the limit state equation could be used instead. The resulting method is sometimes called the *Second Order Reliability Method* (SORM) presented by Fiessler et al. (1979).
Level III reliability methods compute the probability of failure based on the exact probability density function and the exact limit state equation. The most reliable level III method is the *Monte Carlo* method (MC) as indicated for instance in Ouypornprasert (1988), which for this reason is often used as a base for comparison of other methods. The MC technique consists of sampling the relevant variables from their given distribution, subsequent checking whether failure occurs, and repeating this a number of times, say N. The observed proportion of failures $\hat{p}_\mathrm{f}$ of the simulation serves as an estimate for the true probability of failure p_f. The standard deviation of this estimate is $\sqrt{p_\mathrm{f}(1-p_\mathrm{f})/N}$, i.e., its accuracy increases as the square root of N. This is one of the great advantages of Monte Carlo: the obtained accuracy does not depend on the problem dimension. Even though this square-root-of-N behavior is immutable, reduction of the error may sometimes still be achieved, by application of variance reduction techniques. One of these is *impor-*

tance sampling (IS) and its application here would mean that, instead of sampling from the given distribution, one samples from a distribution which is centered at or near the design point. From the observed failures, by applying a weighing scheme, an unbiased estimate of p_f is be obtained that can be (much) more efficient than simple Monte Carlo. *Numerical integration* is another approach whose application to this reliability problem is straight-forward: one needs to integrate the joint probability density function (JPDF) of the variables over the failure domain as indicated by Ouypornprasert (1988). Typically, the approximation error is a low power of the grid size Δ, whereas the required number of evaluations of G is inversely proportional to Δ^n, where n is the problem dimension. This implies that the computational effort to attain a fixed level of accuracy increases very rapidly with the problem dimension. This remains true even if one choses an integration method tailored to the situation, for example, the directional integration method suggested by Deák (1980), based on transformation to polar coordinates. This seems similar to a Monte Carlo version called *directional sampling*, see Nie and Ellingwood (2000). Sometimes, no explicit expression of the LSE is available and methods like FORM cannot be applied. If it is possible to evaluate the limit state function G at any point in the parameter space the *response surface* method can be applied. This method, originally proposed by Box and Wilson for modeling of the response of a chemical process in Box and Wilson (1954), builds an approximation to the LSE from a collection of well-chosen points for which the limit state function is evaluated and then proceeds in a fashion similar to FORM and SORM.

2.2 Problem statement: high accuracy and complex models

The reliability methods which were briefly discussed present the main approaches in the reliability analysis of limit state equations. When models are complex and a high degree of accuracy is required, there are few viable methods.

The standard level II method, FORM, is widely applied and sometimes considered as reliable and robust addressed in Bjerager (1990). However, it has some problems and limitations, some of which are illustrated in Figure 2.2. In Figure 2.2(a) a number of points on the limit state equation attains the minimum distance to the origin, making the choice of design point ambiguous. When applying FORM, this tie would have to be broken, or else the procedure would not converge, and one of the points would be chosen for the linearization. Clearly, the influence factors are bound to be off as well. Another situation, which could appear in any area of civil engineering, is depicted in Figure 2.2(b), where there are two (or more) points with almost the same distance from the origin. In this case, selection of one of them will give a skewed impression of the α factors, the relative importance of the different variables in their contribution to the occurrence of failure will not be correctly assessed as advocated by Rajabalinejad et al.



(a) (b)

Figure 2.2: Limitations of the FORM: instability of calculations in (a), and incorrect estimates of influence factors, in both (a) and (b). The figures are depicted in a standard normal space.

(2007a). Furthermore, when FORM is coupled with a complex implicit model, its use of numerical derivatives introduces approximations into the analysis, with loss of accuracy as a consequence.

Some of the mentioned drawbacks may be overcome by using SORM or the response surface method. The essential drawback, however, is that an approximation to the limit state equation is made and this introduces inaccuracies that cannot be recovered from. Therefore, if accuracy is required, numerical integration and Monte Carlo are the only options. Both of them can result in almost exact solutions if it is possible to evaluate the limit state function for a sufficient number of points. However, when evaluating the function is time-consuming this is impractical because of the prohibitive computing effort. For numerical integration there is an additional problem for high-dimensional problems, as the required number of evaluations grows rapidly as a function of dimension and desired accuracy. This, in contrast with Monte Carlo, where the required computing effort grows quadratically with the desired accuracy, independent of the problem dimension. In practice, this means that numerical integration outperforms Monte Carlo (possibly with importance sampling) for low-dimensional problems and the latter is more efficient in high-dimensional ones.

This dissertation focuses on problems where a highly accurate estimate of the failure probability is required, but an explicit expression for the limit state equation is unavailable and the limit state function can only be evaluated without loss of accuracy via finite element analysis or some other time consuming process.

3

Dynamic Bounds

For the reliability analysis of engineering structures a variety of methods is known, of which Monte Carlo simulation is widely considered to be among the most robust and most generally applicable. The absence of systematic errors and the fact that its error analysis is well-understood are properties that many competing methods lack. A drawback is the often large number of runs needed, particularly in complex models, where each run may entail a finite element analysis or other time consuming procedure. Variance reduction methods may be applied to reduce simulation cost. This chapter describes a method to reduce the simulation cost even further, while retaining the accuracy of Monte Carlo, by taking into account widely-present monotonicity. For models exhibiting monotonic (decreasing or increasing) behavior, *dynamic bounds* are defined, which in a coupled Monte Carlo simulation are updated dynamically, resulting in a failure probability estimate, as well as a strict (non-probabilistic) upper and lower bound. Accurate results are obtained at a much lower cost than an equivalent ordinary Monte Carlo simulation. In a two-dimensional and a four-dimensional numerical example, the cost reduction factors are 130 and 9, respectively, at the 5% accuracy level. At higher accuracy levels, this factor increases, though this effect is expected to be smaller with increasing dimension.

3.1 Introduction

This chapter focuses on problems where a highly accurate estimate of the failure probability is required, but an explicit expression for the limit state equation is unavailable and the limit state equation can only be evaluated without loss of accuracy via finite element analysis or some other time consuming process. The main requirement, therefore, is to reduce the cost of calculations without reduction of accuracy. This can be achieved by a exploiting some properties common

to many engineering problems: monotonicity and the threshold behaviour. To illustrate these, consider that many engineering structures are designed with a balance between resistance forces and driving forces causing stresses. The ratio between the function of resistance and function of forces is called factor of safety, F_s. In this case, $F_s = 1$ may be interpreted as a threshold, and any point above this threshold is stable. If, for such a point, driving forces were to decrease of resistance forces were to increase, everything would remain stable ($F_s > 1$). In the Monte Carlo simulation this can be exploited since the stability or instability of some points may be decided by comparison with earlier results, thus avoiding evaluation of the limit state equation whenever possible. During the simulation *dynamic bounds* are constructed. Progressively more accurate approximations to the stable and unstable regions are obtained by generating points from the joint probability density function and, when necessary, evaluating the limit state equation. From these regions an upper and a lower bound to p_f may be computed, while a regular Monte Carlo estimate is obtained as well.

In the following section, the algorithm is described in detail and some of its properties are described.

3.2 Dynamic bounds

The mathematical formulation of the problem is the following. Given is a limit state equation $G(\vec{x})$, where $\vec{x} = (x_1, \ldots, x_n)$ represents the vector of relevant parameters, and n is the dimension of the problem (for purposes of illustration sometimes set equal to 2). The parameters are modeled as a random vector $\vec{X} = (X_1, \ldots, X_n)$, whose joint probability density function f is given. Since $G(\vec{x}) < 0$ corresponds to failure, the probability of failure is given by

$$p_\mathrm{f} = \mathrm{P}\Big(G(\vec{X}) < 0\Big) = \int \cdots \int_{\vec{x}:G(\vec{x})<0} f(\vec{x})\,\mathrm{d}\vec{x}.$$

In the simple Monte Carlo approach, one would take N independent replications $\vec{X}_1, \ldots, \vec{X}_N$ of $\vec{X}$ and set

$$\hat{p}_\mathrm{f} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}[G(\vec{X}_i) < 0],$$

where $\mathbf{1}[C]$ equals 1 if condition C is true and 0 otherwise. This procedure would take N evaluations of the limit state equation G. In many situations, however, G is increasing in some variables and decreasing in others, which can be exploited to reduce the number of times $G(\vec{x})$ is actually evaluated. First, it will be shown that there is no loss in generality if one assumes that G is *increasing* (in each variable).

3.2.1 Monotonicity

A function is called *increasing* with respect to a variable if increasing that variable causes the output to increase, regardless the values of the other variables. If the output decreases in this situation, one says that the function is *decreasing* with respect to the variable. A function is called *monotonic* if, in each of its variables, it is either increasing or decreasing. More formally, a limit state equation G is called monotonic if, for each i, and for each choice of x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n, the function h_i defined by

$$h_i(x) = G(x_1, x_2, \cdots, x_{i-1}, x, x_{i+1}, \cdots, x_n)$$

is either increasing, i.e., $x \leq y$ implies $h_i(x) \leq h_i(y)$, or decreasing, i.e., $x \leq y$ implies $h_i(x) \geq h_i(y)$. Note that this definition also includes functions that may not be *strictly* increasing.

Many engineering problems exhibit some form of monotonicity: a larger load on a structure means a smaller safety margin or possibly failure; strengthening the structure typically increases the safety margin. There are many examples of such monotonic behavior. Therefore, assuming a monotonic limit state equation does not limit the scope of application very much.

When the limit state equation $G(\vec{x})$ is monotonic with respect to *all* of its variables, it is possible to convert the problem stated at the beginning of this section to one with an *increasing* limit state equation, say, G_1, as follows. Without loss of generality, one may assume that $G(x_1, \ldots, x_n)$ is increasing in the first, say, k variables, and decreasing in the remaining $n-k$. Define G_1 by

$$G_1(x_1, \cdots, x_n) = G(x_1, \cdots, x_k, c_{k+1} - x_{k+1}, \cdots, c_n - x_n), \tag{3.1}$$

where $c_{k+1}, \ldots, c_n$ are constants, then G_1 is clearly increasing in all of its variables. Generally, one would also have to apply an appropriate transformation to the distribution of $\vec{X}$. However, in the case of independent normal variables X_1,..., X_n, if one chooses $c_i = \mu_i = \mathrm{E}[X_i]$ for $i = k+1, \ldots, n$, then $\mathrm{P}\Big(G(\vec{X}) < 0\Big) = \mathrm{P}\Big(G_1(\vec{X}) < 0\Big)$ and transformation is not even necessary. The explanation is that, for any normally distributed Y with expectation μ, the random variable $\mu - Y$ has the same distribution as Y.

From now on it will be assumed that one is dealing with a monotonically increasing limit state equation.

3.2.2 Thresholds

Thresholds divide the space of model parameters into several subsets with desired properties and make a logical division possible. A famous example is the already mentioned factor of safety, the ratio between resistance and driving forces, $F_s =$

resistance/force. According to this relation, $F_s = 1$ is defined as the threshold of the model and the structure is stable or unstable when F_s above or below one, respectively. For monotonic models the threshold concept is interesting from the point of view of stability: if a response of a monotonic model is above its threshold, it will remain stable by increasing its strength parameters and by decreasing the driving forces. For a multidimensional model, threshold points could be identified for any subset of the variables. In standardized space, starting from the origin, one could decrease the variables in the subset simultaneously and at the same rate, or alternatively, use some linesearch algorithm like bisection. The sequence of points will at first be stable and then cross the threshold and become unstable. The two points closest to the $G(\vec{x}) = 0$ boundary, on opposite sides are kept and called the *upper* and *lower threshold* point. In principle, this could be done for any subset of the variables, leading to a maximum of $2\,(2^n - 1)$ threshold points for an n dimensional model (with $2^n - 1$ subsets of the variables).

It is clear that in case of an implicit limit state equation, and without explicit knowledge of their location, serious computational effort may be required to determine the threshold points. On the other hand, a collection of threshold points may be a good starting point for the Monte Carlo algorithm described in the next section, in the sense that fewer evaluations of the limit state equation a needed than when starting randomly (as seems to be the alternative).

3.2.3 The Monte Carlo algorithm

To begin, let us consider the ideas of the Monte Carlo algorithm by looking at a 2-dimensional example. After that, the mathematical description is given, including computations. In Figure 3.1, a two-dimensional limit state equation $G(x_1, x_2) = 0$ is depicted, as well as the contours of the joint probability density function $f(x_1, x_2)$ of the two variables (X_1, X_2). The limit state equation $G(x_1, x_2)$ is assumed to be monotonically increasing in both variables. This means that $G(x_1, x_2) < 0$ for points below the LSE; we will call this the *unstable region* as these points correspond to failures. For points above the LSE, $G(x_1, x_2) > 0$, this is the *stable region*.

A first random point, $\vec{x}^{(1)} = (x_1^{(1)}, x_2^{(1)})$, is generated from the JPDF f. In Figure 3.1 it is depicted by a black square labeled 1, and upon evaluation it is found that $G(x_1^{(1)}, x_2^{(1)}) < 0$, hence it is a failure, in the unstable region. From the monotonicity of G it is inferred that $G(x_1, x_2) < 0$ for all points (x_1, x_2) in the quadrant to the left of and below $(x_1^{(1)}, x_2^{(1)})$. In the next step, the point $(x_1^{(2)}, x_2^{(2)})$ is generated from f, it turns out that $G(x_1^{(2)}, x_2^{(2)}) > 0$. So point 2 is in the stable region and all points in the right-upper quandrant from point 2 are stable as well.

This process continues; the result of a small number of iterations is shown in Figure 3.1. The shaded regions constitute approximations to the stable and unstable regions that can be used to obtain bounds on the probability of failure p_{f}.

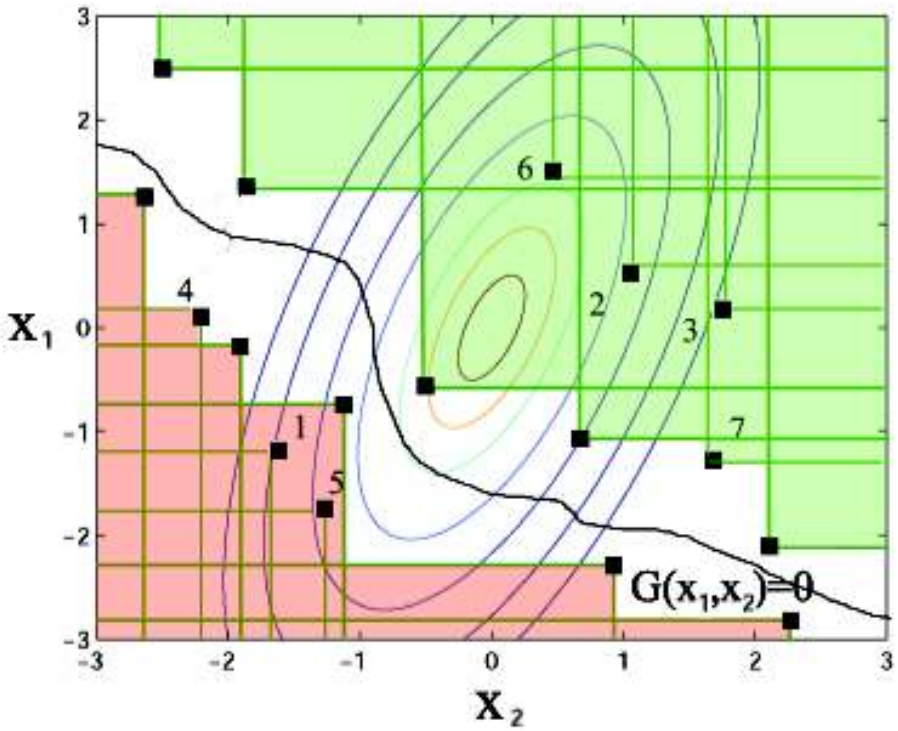


Figure 3.1: Illustration of the dynamic bounds algorithm in which a two dimensional joint probability density function is divided into the stable, unstable, and unqualified regions.

The reader should keep in mind that in practice the location of the LSE-curve is not known; it only gradually becomes visible as it is being sandwiched between the increasingly more accurate approximations to the stable and unstable regions. The corresponding dynamic upper and lower bounds on p_f become tighter as the number of generated points increases. Also note that for some generated points it is not necessary to evaluate the limit state equation, because from other points it can be determined whether the point is stable or unstable.

For the mathematical description we revert to dimension n. Define the stable set S and the unstable set U by

$$S = \{\vec{x} \,:\, G(\vec{x}) \geq 0\} \quad \text{and} \quad U = \{\vec{x} \,:\, G(\vec{x}) < 0\}. \tag{3.2}$$

If two points $\vec{x} = (x_1, \ldots, x_n)$ and $\vec{y} = (y_1, \ldots, y_n)$ satisfy the relationship $x_i \leq y_i$ for $i = 1, \ldots, n$, then we say that $\vec{x}$ is *less stable than* $\vec{y}$, or: $\vec{y}$ is *more stable than* $\vec{x}$. If $\vec{x} \in S$ then $\vec{y} \in S$ follows and a similar statement holds for U.

Consider the k-th iteration of the Monte Carlo process. A number of stable points, say, $\vec{s}^{(1)}$, ..., $\vec{s}^{(p)}$ and a number of unstable points $\vec{u}^{(1)}$, ..., $\vec{u}^{(q)}$ have been generated. The current approximation to the stable region S is the union of

the p orthants (generalizing the quadrants in Figure 3.1)

$$H_i = \{\vec{x} : \vec{x} \text{ is more stable than } \vec{s}^{(i)}\}, \quad i = 1, \ldots, p. \tag{3.3}$$

If $\vec{s}^{(i)}$ is more stable than $\vec{s}^{(j)}$, for some i and j, its orthant H_i is completely contained in H_j, and there would be no loss of information if $\vec{s}^{(i)}$ would be dropped from the list. Similarly, the current approximation to the unstable region U is the union of

$$L_i = \{\vec{x} : \vec{x} \text{ is less stable than } \vec{s}^{(i)}\}, \quad i = 1, \ldots, q. \tag{3.4}$$

From now on, it is assumed that only a minimal set of stable and unstable points is retained during the simulation and S_k and U_k are the corresponding approximations to S and U:

$$S_k = \cup_{i=1}^{p} H_i \quad \text{and} \quad U_k = \cup_{i=1}^{q} L_i. \tag{3.5}$$

Now, imagine the next random point $\vec{X}^{(k+1)}$ is generated from f. There are three possibilities. The first is: $\vec{X}^{(k+1)} \in S_k$, that is, the point is located in a region that is known to be part of the stable set. The second is: $\vec{X}^{(k+1)} \in U_k$, that is, the point is located in a region that is known to be part of the unstable set; the count of the number of failures should be incremented. The third is: $\vec{X}^{(k+1)} \notin S_k \cup U_k$, that is, the point is located in the unqualified region between S_k and U_k; $G(\vec{X}^{(k+1)})$ needs to be evaluated. If it is positive, $\vec{X}^{(k+1)}$ is added to the collection of known stable points and this collection is checked for its minimality, dropping any superfluous points. If it is negative, $\vec{X}^{(k+1)}$ is added to the collection of unstable points and a similar update is performed. Note that the numbers p and q vary during the simulation and in fact depend on the iteration number k.

Summarizing, the algorithm is as follows:

1. Determine S_0 and U_0. They could be empty sets or be determined from the threshold points, as described. Set $k = 0$, $n_{\mathrm{f}} = 0$.

2. Increase k by 1 and generate $\vec{X}^{(k)}$ from f. If $\vec{X}^{(k)} \in U_{k-1}$, add 1 to n_{f} and update U_{k-1} to obtain U_k. If $\vec{X}^{(k)} \notin S_{k-1} \cup U_{k-1}$, evaluate $G(\vec{X}^{(k)})$; if it is negative, add 1 to n_{f} and update U_{k-1} to obtain U_k; otherwise, update S_{k-1} to obtain S_k. Repeat until $k = N$.

3. $\hat{p}_{\mathrm{f}} = n_{\mathrm{f}}/N$ is a simple Monte Carlo estimate for p_{f}.

This estimate $\hat{p}_{\mathrm{f}}$ is as good as an ordinary Monte Carlo estimate based on N independent samples $\mathbf{1}[G(\vec{X}^{(1)}) < 0]$,...,$\mathbf{1}[G(\vec{X}^{(N)}) < 0]$, but requires evaluation of G in only a fraction of the samples. However, the simulation also yields an upper and a lower bound on p_{f}, as follows. Clearly, $U_N \subset U$ and $S_N \subset S$. So,

$$\hat{p}_{\mathrm{u}} := \mathrm{P}\Big(\vec{X} \in U_N\Big) \leq p_{\mathrm{f}} \quad \text{and} \quad \hat{p}_{\mathrm{s}} := \mathrm{P}\Big(\vec{X} \in S_N\Big) \leq 1 - p_{\mathrm{f}}, \tag{3.6}$$

where $\vec{X}$ is an independent draw from f. These imply the following bounds on p_{f}:

$$\hat{p}_{\mathrm{u}} \leq p_{\mathrm{f}} \leq 1 - \hat{p}_{\mathrm{s}}. \tag{3.7}$$

Remark. It may be hard to evaluate $\hat{p}_{\mathrm{u}}$ and $\hat{p}_{\mathrm{s}}$ in (3.7), because U_N and S_N are irregular sets. The inclusion-exclusion principle may enable the computation of two approximations that lead to looser bounds, as follows. From the representation of U_N as the union of L_i, $i = 1, \ldots, q$, one obtains

$$\mathrm{P}\Big(\vec{X} \in U_N\Big) = \mathrm{P}\Big(\vec{X} \in L_i, \text{for some } 1 \leq i \leq q\Big)\,, \tag{3.8}$$

and from the inclusion-exclusion principle it follows that this is greater than or equal to:

$$\sum_{i=1}^{q} \mathrm{P}\Big(\vec{X} \in L_i\Big) - \sum_{1 \leq i < j \leq q} \mathrm{P}\Big(\vec{X} \in L_i \cap L_j\Big)\,. \tag{3.9}$$

If the coordinates of $\vec{X} = (X_1, \ldots, X_n)$ are independent:

$$\mathrm{P}\Big(\vec{X} \text{ less stable than } \vec{u}\Big) = \prod_{i=1}^{n} \mathrm{P}(X_i \leq u_i) = \prod_{i=1}^{n} F_i(u_i), \tag{3.10}$$

where F_i denotes the distribution function of X_i. Since $L_i \cap L_j$ represents an orthant, just as the L_i, $\mathrm{P}\Big(\vec{X} \in L_i \cap L_j\Big)$ can be computed similarly. This shows that a lower bound for $\hat{p}_{\mathrm{u}}$ for p_{f} can be computed from the marginal distributions of $\vec{X}$.

Similarly, one obtains

$$\mathrm{P}\Big(\vec{X} \in S_N\Big) \geq \sum_{i=1}^{q} \mathrm{P}\Big(\vec{X} \in H_i\Big) - \sum_{1 \leq i < j \leq q} \mathrm{P}\Big(\vec{X} \in H_i \cap H_j\Big)\,, \tag{3.11}$$

where (in case of independence) the probabilities on the right-hand side can be computed from

$$\mathrm{P}\Big(\vec{X} \text{ more stable than } s\Big) = \prod_{i=1}^{n} \mathrm{P}(X_i \geq s_i) = \prod_{i=1}^{n} (1 - F_i(s_i)). \tag{3.12}$$

3.3 The efficiency of dynamic bounds

The estimate $\hat{p}_{\mathrm{f}}$, obtained from the DB simulation with N iterations is in fact an ordinary MC estimate, therefore with standard error $\sqrt{p_{\mathrm{f}}\,(1 - p_{\mathrm{f}})/N}$. The distinction is that, instead of N evaluations of the LSE, a much smaller number suffices. The exact number of evaluations N_{LF} is random and dependent on the (random) evolution of the simulation, so the efficiency can only be determined exactly afterwards. At the start, however, one would like to predict N_{LF}, which can be done to some extent. If parts of the stable and unstable regions are known (for example, after identifying threshold points) one can determine (estimate) the probability Δ that the next random $\vec{X}$ is in the unqualified region, i.e., the unknown part. If it

is in the unqualified region, the limit state equation is evaluated, and the known part of either the stable or the unstable region increases, which means for the next point Δ has decreased a bit. If the point is in a known region, Δ remains the same. For the whole simulation, this probability is Δ_k after k iterations, and it constitutes a (weakly) decreasing sequence. Formally, one could write:

$$\Delta_k = 1 - \hat{p}_{\mathrm{s},k} - \hat{p}_{\mathrm{u},k},$$

where $\hat{p}_{\mathrm{s},k}$ and $\hat{p}_{\mathrm{u},k}$ are determined as in (3.6), with N replaced by k. A conservative estimate for N_{LF} therefore is $N \cdot \Delta_0$, where Δ_0 is the initial Δ, from the initialization step (1) of the algorithm. The actual computation will most likely not exceed $\Delta_0 \cdot N$.

This result can also be used to estimate the cost of simulation to attain a required level of accuracy, by executing the procedure in two stages. For example, suppose one has a preset tolerance that the standard error should be less than 5%. This would mean

$$N = 400 \left(\frac{1}{p_{\mathrm{f}}} - 1 \right), \tag{3.13}$$

replicates in the Monte Carlo simulation (where, for actual computation, an initial estimate for p_{f} should be inserted). In a pilot stage, one runs $N_1 < N$ replicates. One may now estimate:

$$N_{LF} \approx 400\, \Delta_{N_1} \left(\frac{1}{\hat{p}_{\mathrm{f},N_1}} - 1 \right), \tag{3.14}$$

as the total number of LSE evaluations needed.

It can be argued that Δ_k will tend to decrease geometrically. This can also be used to predict required computation costs from the results of pilot stages. Suppose stage 1 is run with N_1 Monte Carlo replications, requiring n_1 LSE evaluations, and stage 2 up to $10 \cdot N_1$ replications, requiring n_2 evaluations. Contemplating stage 3 going up to $100 \cdot N_1$ replications, the geometric property predicts that $n_3 : n_2 \approx n_2 : n_1$, whence one would predict:

$$n_3 \approx \frac{n_2^2}{n_1}. \tag{3.15}$$

If the DB algorithm is combined with importance sampling, there will be an additional gain in efficiency. For the same accuracy a smaller N than in (3.13) will suffice. This will lead to a reduction of N_{LF} as well, though the Δ_k will behave differently. For example, when the sampling is performed from a distribution that is centered on or near the design point, then Δ_0 will be larger and the updating process will be different as well.

It is expected that the efficiency of the DB algorithm decreases as the number of variables increases, because it seems that the reduction of the unqualified

part of the space proceeds less rapidly, which can be understood intuitively by a geometrical consideration of the "known" orthants that accumulate in the cause of the simulation. In fact, with a large number of variables, the efficiency of this method may approach that of MC. However, nothing would be lost, if one had a MC simulation with only a marginally reduced computational effort.

3.4 Example: impact of water waves on coastal structures

In order to illustrate the efficiency of DB in comparison with the other reliability methods, a simple example of a limit state equation quantifying the effect of water waves on coastal structures is presented. The results, in this case, have been compared with the FORM, and several level III methods including Monte Carlo (MC), numerical integration (NI), and importance sampling (IS). DB coupled with IS is included in the comparison as well.

An important research topic in hydraulic engineering focuses on the impact of water waves on walls and other coastal structures. Breaking waves create velocities and pressures with magnitude much larger than those associated with the propagation under gravity of ordinary waves. They can generate pressures of up to 1000 kN/m^2, i.e., one hundred meters of water head! Although many coastal structures are damaged by breaking waves, very little is known about the impact mechanism. Insight into wave impact has been gained by investigating the role of entrained and trapped air in wave impacts. A simplified model of maximum pressure of ocean waves on the coastal structures is given by

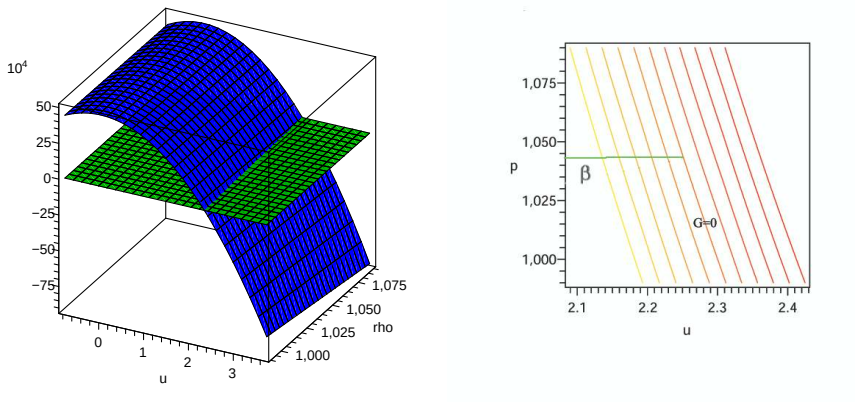
$$P_{max} = C \times \frac{\rho \times k \times u^2}{d}, \tag{3.16}$$

where ρ is density of water, k is the length of a hypothetical piston, d is the thickness of the air cushion, u is the horizontal velocity of the advancing wave, and C is a constant: 2.7 s^2/m. Based on this model, we are desire to find the probability that the (maximum) impact pressure exceeds $5 \times 10^5 N/m^2$. Before considering this four-dimensional model, we analyse a two dimensional simplification. Together, the results will then give some indication of dimension effects.

3.4.1 Two dimensional model

In a preliminary analysis, we fix two parameters at their expected values: $k = 3.5$ and $d = 0.1$. Rewriting the model a two-dimensional limit state equation is obtained:

$$G(\rho, u) = 500000 - 94.50 \times \rho \times u^2. \tag{3.17}$$



(a) 3D plot of LSE (b) Contour plot of LSE

Figure 3.2: (a) Visualization of the LSE (3.17) and the failure line, $G(\rho, u) = 0$. (b) Contours of the limit state equation, the β value and the design point.

Var.	μ	σ	PDF	upper threshold	lower threshold
ρ	1040	10	Normal	$\mu_\rho + 1.63\sigma_\rho = 1056.30$	$\mu_\rho + 1.64\sigma_\rho = 1056.40$
u	1.5	0.45	Normal	$\mu_u + 1.63\sigma_u = 2.2335$	$\mu_u + 1.64\sigma_u = 2.238$

Table 3.1: Variables in equation 3.17, its upper and lower threshold.

Figure 3.2(a) shows the limit state equation in 3D together with the failure plane and Figure 3.2(b) shows its contours. It is increasing in both ρ and u. One wishes to determine $p_f = P(G(\rho, u) < 0)$, where (ρ, u) are assumed to be independent normally distributed, their parameters are given in Table 3.2.

A FORM analysis of Equation 3.17 is performed, the results of which are given in in Table 3.2. The estimated probability of failure is 0.0466, which is very close to the exact value. The influence factors show that the failure is much more influenced by u than by ρ, as was expected. Level III reliability methods, also, are applied to the LSE(3.17), the results are in Table 3.3. For each method there are several rows, the first shows the number of evaluations of the LSE, the second the corresponding estimate for p_f, and the third shows the *relative standard*

Variable	μ	σ	α	Value in design point
ρ	1040	10	0.024	1040
u	1.5	0.45	1	2.255
β				1.68
P_f				0.0466

Table 3.2: FORM results of the LSE of equation 3.17.

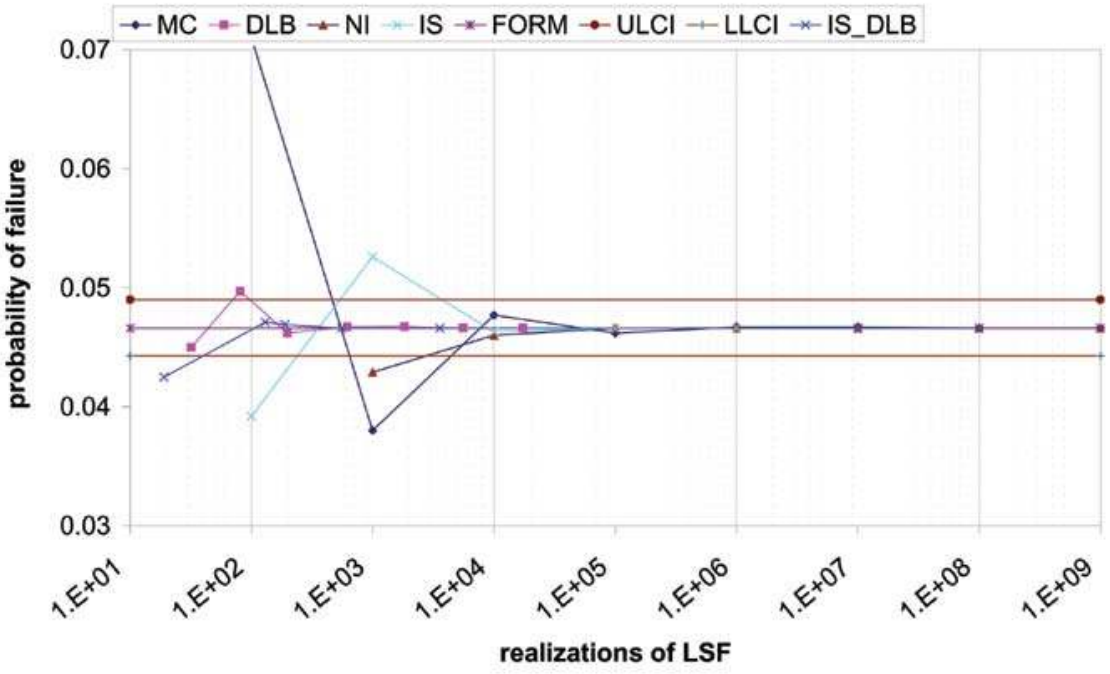


Figure 3.3: Comparison between the results of several reliability methods: dynamic bounds (DB), Monte Carlo (MC), numerical integration (NI), importance sampling (IS), and importance sampling coupled with dynamic bounds (IS-DB).

error, i.e., the standard error divided by p_f. The number of evaluations is varied, but the columns contain outcomes of approximately the same accuracy (except for the NI results, that are matched with MC in terms of number of evaluations of the LSE). For the DB method the last two rows present the lower and upper bound from equation (3.7) (the reader is reminded that these bounds are not confidence bounds, but hold with certainty). The required number of evaluations for the DB results show that comparatively few of them are needed to obtain accurate results. Looking at the 5% tolerance column (the second one), only 77 evaluations of the LSE were needed. As explained, this variance is the same as the equivalent Monte Carlo (MC) simulation, which was based on 10,000 realizations. Also note the number of LSE evaluations approximately triples when going to the next column, which corresponds to *ten* times as many Monte Carlo realizations. This means that the relative efficiency of DB with respect to Monte Carlo increases with the precision required. A numerical integration method has been applied on an equivalent mesh over the interval $[-5\sigma,\ +5\sigma]$. IS is applied centered at the design point of the FORM results. As shown in the last row of table, the combination of DB and IS provides robust approach. The convergence speeds of different approaches are presented in Figure 3.3.

	A comparison between DB and other level III methods						
DB	27	**77**	203	598	1828	5556	17286
p_f	0.0490	**0.0471**	0.0455	0.0468	0.0467	0.0466	0.0466
Rel. s.e.	13.93E-2	**4.5E-2**	1.45E-2	0.45E-2	0.14E-2	0.045E-2	0.014E-2
$p_f \geq \hat{p}_u =$	0.042	0.0451	0.0463	0.0466	0.0466	0.0466	0.0466
$p_f \leq 1 - \hat{p}_s =$	0.053	0.0490	0.0472	0.0469	0.0466	0.0466	0.0466
MC	1E+3	1E+4	1E+5	1E+6	1E+7	1E+8	1E+9
p_f	0.038	0.0477	0.0462	0.0467	0.0467	0.0466	0.0466
Rel. s.e.	15.9E-2	4.5E-2	1.44E-2	0.45E-2	0.14E-2	0.045E-2	0.014E-2
NI	1E+3	1E+4	1E+5	1E+6	1E+7	1E+8	1E+9
p_f	0.0429	0.0465	0.0466	0.0466	0.0466	0.0466	0.0466
IS	1E+2	1E+3	1E+4	1E+5	1E+6		
p_f	0.0412	0.0464	0.0467	0.0468	0.0466		
Rel. s.e.	15.8E-2	4.64E-2	1.42E-2	0.45E-2	0.14E-2		
IS-DB	12	55	136	271	959		
p_f	0.0425	0.0460	0.0469	0.0468	0.0467		
Rel. s.e.	16.4E-2	4.58E-2	1.44E-2	0.64E-2	0.17E-2		

Table 3.3: Comparison between level III methods: dynamic bounds (DB), Monte Carlo (MC), numerical integration (NI), important sampling (IS), and DB coupled with IS (IS-DB). Rows marked p_f contain the estimates for the probability of failure. *Rel. s.e.* is the relative standard error. The bottom two DB rows present $\hat{p}_u$ and $1 - \hat{p}_s$: certain (i.e., non-probabilistic) bounds on p_f. Bold numbers present the acceptable engineering level.

Variable	μ	σ	PDF	u_t (upper threshold)	l_t (lower threshold)
ρ	1040	10	Normal	$\mu_\rho + 0.98\sigma_\rho = 1049.80$	$\mu_\rho + 1.00\sigma_\rho = 1050$
k	3.5	0.7	Normal	$\mu_k + 0.98\sigma_k = 4.186$	$\mu_k + 1.00\sigma_k = 4.20$
u	1.5	0.45	Normal	$\mu_u + 0.98\sigma_u = 1.941$	$\mu_u + 1.00\sigma_u = 1.95$
d	0.1	0.01	Normal	$\mu_d - 0.98\sigma_d = 0.0902$	$\mu_d - 1.00\sigma_d = 0.09$

Table 3.4: Variables in equation 3.18, its upper and lower threshold.

Variable	μ	σ	α	Value in design point
ρ	1040	10	0.020	1040
k	3.5	0.7	0.317	3.894
u	1.5	0.45	0.90	2.103
d	0.1	0.01	-0.217	0.097
β				1.49
P_f				0.0681

Table 3.5: FORM results of the LSE of equation 3.16.

3.4.2 Four dimensional model

Now, we revert to the original model as in (3.16) from which we derive a four-dimensional limit state equation:

$$G(\rho, u, k, d) = 500000 - 2.7 \times \frac{\rho \times k \times u^2}{d}, \tag{3.18}$$

for which we wish to determine $p_\mathrm{f} = \mathrm{P}(G(\rho, u, k, d) < 0)$, where (ρ, u, k, d) are independently normally distributed with parameters given in Table 3.4. This LSE is increasing with respect to d and decreasing with respect to the others. In this case, the upper and lower thresholds can be calculated as presented in Table 3.4. A FORM analysis of Equation 3.18 is performed, the results of which are given in in Table 3.5. The estimated probability of failure is 0.0681, which is off by 7%. The influence factors and the design point are calculated. The latter is used as center of the importance sampling distribution.

Level III methods are applied to the four-dimensional LSE Equation 3.18, and the results are presented in Table 3.6. As noted, the FORM method does not give a very accurate answer in this example, which is an illustration of the theoretical limitation as discussed in Section 3.1. Still, the design point is useful as center for the shifted distribution used for the IS and IS-DB methods. Table 3.6 presents a comparison between DB and other methods. It again shows that the DB method provides the same order of accuracy as MC at a much lower cost. The cost ratio decreases with a factor of 2, when going from one column to the next.

	A comparison between DB and other level III methods						
DB	227	**1112**	5230	24987	128883		
p_f	0.0630	**0.0672**	0.0645	0.0637	0.0638		
Rel. s.e.	12.2E-2	**3.7E-2**	1.2E-2	0.38E-2	0.12E-2		
$p_f \geq \hat{p}_u =$	0.0218	0.0462	0.0489	0.0556	0.0593		
$p_f \leq 1 - \hat{p}_s =$	0.1719	0.1079	0.0843	0.0731	0.0687		
MC	1E+3	1E+4	1E+5	1E+6	1E+7	1E+8	1E+9
p_f	0.074	0.068	0.0628	0.0637	0.0638	0.0638	0.0638
Rel. s.e.	11.2E-2	3.7E-2	1.22E-2	0.38E-2	0.12E-2	0.045E-2	0.014E-2
IS	1E+2	1E+3	1E+4	1E+5	1E+6		
p_f	0.0600	0.0591	0.0631	0.0639	0.0637		
Rel. s.e.	14.8E-2	4.6E-2	1.4E-2	0.44E-2	0.14E-2		
IS-DB	68	427	2187	11025			
p_f	0.0630	0.0636	0.0637	0.0637			
Rel. s.e.	14.9E-2	4.7E-2	1.4E-2	0.44E-2			

Table 3.6: Comparison between level III methods: dynamic bounds (DB), Monte Carlo (MC), numerical integration (NI), important sampling (IS), and DB coupled with IS (IS-DB). Rows marked p_f contain the estimates for the probability of failure; *Rel. s.e.* is the relative standard error. The bottom two DB rows present $\hat{p}_u$ and $1 - \hat{p}_s$: certain (i.e., non-probabilistic) bounds on p_f. Bold numbers present the acceptable engineering level.

3.5 Conclusions

The presented dynamic bounds (DB) method is suggested for the reliability analysis of the engineering problems exhibiting monotonicity with a limited number of variables. It is fast and robust and intended for use with complicated limit state equations like finite elements, enabling a probabilistic approach even for these problems. Its main advantage over direct Monte Carlo simulation is that only a fraction of the limit state equation evaluations (finite element analyses) is needed, without loss of accuracy. By breaking up the simulation in two or more stages, initial estimates of the computing effort to attain a required level of accuracy can be updated at intermediate stages, resulting in good predictions of computation costs. Moreover, the method can be coupled with the importance sampling technique, further reducing the required calculations, speeding up the whole procedure. This is illustrated by the two examples based on a nonlinear limit state equation with two and four variables. In Table 3.3 one sees that DB requires 77 instead of 10 000 evaluations for 5% accuracy. Furthermore, every extra digit of accuracy takes approximately 8 times as many evaluations, instead of 100 times, as is the case for Monte Carlo. In Table 3.6 these numbers are: 1112 instead of 10 000, and 25 times instead of 100 times. This illustrates an anticipated dimension effect. However, there will always be some gain, even for much larger dimensional problems. The basic concept and approach would be the same. It seems, however, that these may be varied, perhaps leading to even faster algorithms, a subject that warrants further research.

4

Improved Dynamic Bounds

In the previous chapter the concept of dynamic bounds was described. Here, we address its improvement assuming the order of the response of a limit state equation (LSE) is known. The suggested technique is explained, and results are presented and compared with the classical Monte Carlo (MC) and Dynamic Bounds (DB). It is important to mention that we do not implement any uncertainty in the modeling, and what is done is equal to MC.

4.1 The stable and unstable bounds

The main concept of dynamic bounds is dividing the range of a limit state equation (LSE) into the stable, unstable, and unqualified region as presented in Figure 3.1 and 4.1. Therefore, it was possible to derive two sets of stable S and unstable U according to Equation 3.2. Here, the equivalent sets are introduced by Equation 4.1, where $G(\vec{x})$ is a monotonically increasing function[1] and S_b and U_b are respectively defined on the stable and unstable bounds. In other words, the difference is that these sets only include the points on the stable and unstable bounds and do not include the internal points. Therefore, if two points $\vec{x} = (x_1, \ldots, x_n)$ and $\vec{y} = (y_1, \ldots, y_n)$ which belong to S satisfy the relationship $y_i \geq x_i$ for $i = 1, \ldots, n$ in a monotonic increasing function, then we say that $\vec{x}$ *is less stable than* $\vec{y}$, or: $\vec{y}$ *is more stable than* $\vec{x}$. Then, S_b is any set such that if $\vec{x} \in S_b$ then there is not exists $\vec{y} \in S_b$ such that $\vec{y} \neq \vec{x}$ and $y_i \leq x_i$, $i = 1, \ldots, n$. If $\vec{x} \in U_b$ and $\vec{y} \in U_b$ then a similar statement holds.

[1] A similar statement is defined for a monotonic decreasing function. But it differs in such a way that in a monotonically decreasing function, the LSE gets less stable by increasing of its variables.

$$S_b = \{\vec{x} \ : \ G(\vec{x}) \geq 0 \ \mid \forall \vec{x}, \vec{y} \in S_b, \ \nexists \vec{y} \neq \vec{x}, \ y_i \leq x_i, \ i = 1, \ldots, n\},$$
$$U_b = \{\vec{x} \ : \ G(\vec{x}) < 0 \mid \forall \vec{x}, \vec{y} \in U_b, \ \nexists \vec{y} \neq \vec{x}, \ y_i \geq x_i, \ i = 1, \ldots, n\}. \quad (4.1)$$

In figure 4.1, a view of randomly generated values on the stable and unstable bounds are shown after a few realization of a two dimensional LSE. The values on the stable bounds are shown by $s_b^{(i)}$ which will be extended to $s_b^{(i)\prime}$, and the same concept holds for the values on the unstable bound $u_b^{(i)}$. This figure will be further referred and discussed in Section 4.3.

It was also described in Chapter 3 that given the stable and unstable sets[2], we are able to define a lower bound and upper bound for the probability of failure according to the following equation, derived in Equation 3.7.

$$\hat{p}_{\mathrm{u}} \leq p_{\mathrm{f}} \leq 1 - \hat{p}_{\mathrm{s}}.$$

As a result, the probability of hitting to the unqualified region, called as $p_\Delta$can be calculated by the following equation.

$$p_\Delta = 1 - p_{\mathrm{u}} - p_{\mathrm{s}}. \quad (4.2)$$

4.2 The most uncertain responses and transformation

The main purpose of this chapter is to explore the unqualified region, called as Δ, given some extra information over the LSE. This is in order to define the upper and lower bounds of the response of the LSE by assuming that the order of the response of a LSE and/or the value of its derivatives is known or can be calculated without having an explicit LSE. In fact, given the order of the LSE, we are going to build the most uncertain bounds and reduce the size of the unqualified region, Δ. In other words, we attempt to shrink the unqualified region by defining an upper and lower bound for the the p_Δ based on a given order of the LSE. Then, we extend the stable and unstable sets. This, however, does not implement any uncertainty in the results.

This concept is illustrated in Figure 4.1. In this figure, dynamic bounds for a two dimensional limit state equation $G(\vec{x})$ are schematically shown. Here, a two dimensional example is illustrated because of the fact that a 2D example is easier to illustrate comparing with the higher or lower dimensions.

A limit state equation (LSE) in its general form can be written as $G(\vec{x})$ where $\vec{x} = (x_1, \ldots, x_n)$ is a random vector in its general form assuming that the order of the relation between $G(\vec{x})$ and x_i is known, where $1 \leq i \leq n$. In other words,

[2]These sets become gradually available during the Monte Carlo simulations.

given the order of the h_i in Equation 4.3, it is a possibility to extend the dynamic bounds and shrink the unqualified space in its ith dimension.

$$h_i(x) = G(x_1, x_2, \cdots, x_{i-1}, x, x_{i+1}, \cdots, x_n) \tag{4.3}$$

From now on we turn to a one dimensional problem and assume that points A and B are two points on the lower (unstable) and upper (stable) bounds of the defined dynamic bounds in a one dimensional LSE, where $\mathrm{A} = (\mathrm{x_A}, \mathrm{G(x_A)})$ and $\mathrm{B} = (\mathrm{x_B}, \mathrm{G(x_B)})$. Therefore, $x_A \in U_b$ and $x_B \in S_b$ according to Equation 4.1. It concludes that $G(x_A) < 0$ and $G(x_B) \geq 0$. Accordingly, we can conclude for points A and B (see Equations 3.3 and 3.4) that:

$$G(\vec{x} : \vec{x} \text{ is less stable than } \vec{x}_\mathrm{A}) < 0,$$

$$G(\vec{x} : \vec{x} \text{ is more stable than } \vec{x}_\mathrm{B}) \geq 0.$$

4.2.1 Linear response

Now, having the unstable and stable bounds: points A and B, if we assume that the response of the limit state function is linear, we can directly draw a line from A to B and predict the behavior of the model in the unqualified region, so the line from A to the B in Figure 4.2(a) presents the LSE. As a result, any randomly generated number can be judged whether it is in stable/unstable area. Thereafter, we can proceed the simulation without any extra realization of the LSE. However, this usually does not happen, and we are usually confront with a higher order response of LSEs. In this case, the local coordinates need to be fixed on the bounds as shown in Figure 4.2(b); ξ and η present the local coordinates and will be used in the rest of this chapter. Therefore, $A = (\xi = 0, \eta = 0)$ and $B = (\xi = m, \eta = n)$, where m and n are known values and can be calculated.

4.2.2 Second order response and extreme conditions

Given the response of the model is a second order continuous and smooth, we are looking for the most uncertain response of the LSE. Equation 4.4 presents a general form of its response.

$$\eta = a\xi^2 + b\xi + c. \tag{4.4}$$

This function should fit in to the points A and B, therefore:

$$\begin{aligned} A &= (\xi = 0, \eta = 0) \implies c = 0 \\ B &= (\xi = m, \eta = n) \implies b = \frac{n - am^2}{m} \end{aligned}$$

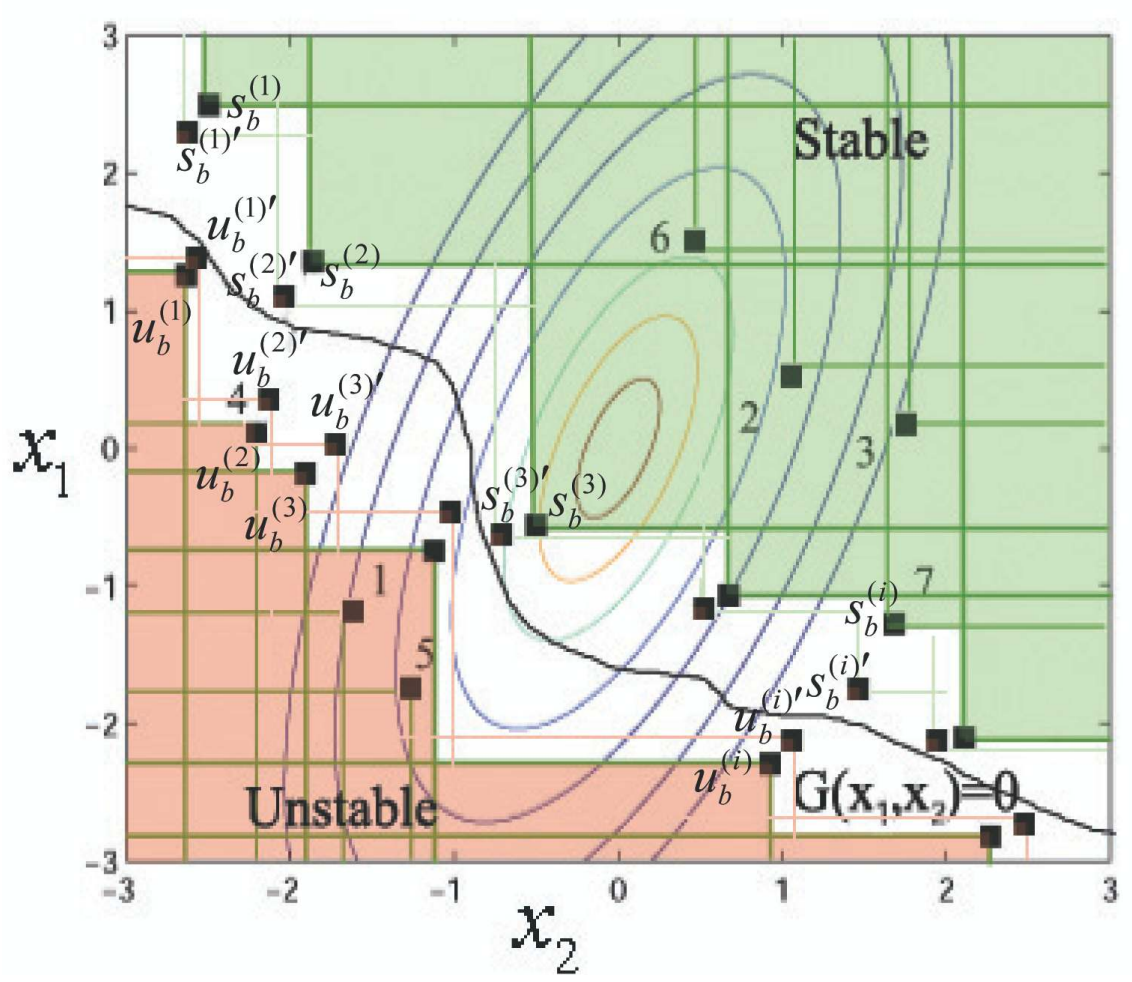
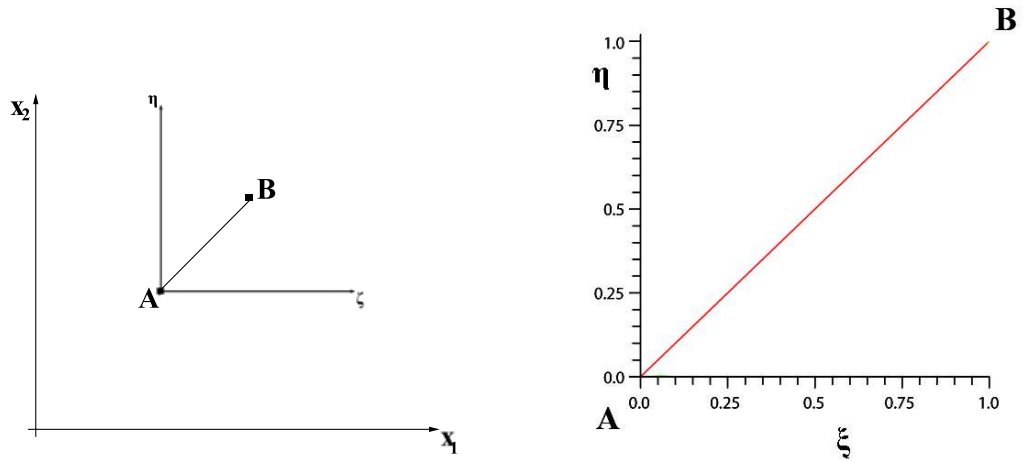


Figure 4.1: Distinction of the stable (safe), unstable (failure), and unqualified regions in a two dimensional space. Also The extended bounds are typically shown.



(a) Linear response of LSE where A is on the lower bound and B is on the upper bound

(b) Linearly transferred coordinates of the LSE from global to local (ξ, η). 1D

Figure 4.2: (a) A and B are two points located on the lower and upper limit bounds, respectively in a one dimensional space of $\vec{x} = (x_1)$. (b) Transferred coordinates of the LSE according to A and B.

Figure 4.3 presents a few different positions in which a second order polynomial passes through points A and B. There are an unlimited number of the second order polynomials which can be depicted. But we are in favor of most uncertain responses. The most uncertain response which can be defined by two bounds can be obtained by minimization or maximization of the integral of the area under the curve between points A and B with respect to each coordinate. The integration is presented by Equation 4.5.

$$I = \int_0^m \eta d\xi = \frac{1}{3}am^3 + \frac{1}{2}(\frac{n}{m} - am)m^2 \tag{4.5}$$

In order to get the maximum uncertainty, the integral should be minimized as presented in the following equation.

$$\frac{\partial I}{\partial m} = 0 \implies a = \frac{n}{m^2},\ b = 0,\ c = 0. \tag{4.6}$$

Equation 4.7 presents the most uncertain second order response of the model (or a lower bound of response surface) regarding its value in horizontal axis. This equation is depicted in Figure 4.3 (a).

$$\eta = g(\xi) = \frac{n}{m^2}\xi^2 \tag{4.7}$$

The same process needs to be done to get the upper bound of the uncertainty through a maximization process. In other terms, the integration or area under the curve between the points A and B should be maximized. The result is presented in Figure 4.3(b). Therefore, Figures 4.3(a,b) present two conditions in which the deviation of the response (uncertainty) is maximized in respect to the horizontal axis, ξ, assuming a second order polynomial response of the limit state function. As a result of these assumptions the lower and upper bounds can be extended to the unqualified area; this concept reduces the Monte Carlo simulations efforts and saves some costs for a time consuming process in the Monte Carlo simulations.

Remark. The integration needs to be also minimized and maximized in respect to the vertical axis η in which two other polynomial will be obtained. These polynomials are depicted in Figure 4.3(c,d). In conclusion, Figure 4.4 integrates all the positions presented in Figure 4.3 and leads to the fact that the response of a second order LSE remains inside of the enclosed region.

4.2.3 Third and higher order responses

Given a third order response of the LSE, which is usually advised in the interpolation process Triebel (1978), assumed continuous and smooth, the LSE in the unqualified region can be written in the following form:

$$\eta = a\xi^3 + b\xi^2 + c\xi + d. \tag{4.8}$$

This polynomial should pass through points A and B, where $A = (0,0)$ and $B = (m,n)$. This concludes that $d = 0$ and some relations between the other parameters. Then we need to minimize (and maximize) the integral in Equation 4.9.

$$I = \int_0^m \eta d\xi = \frac{1}{4}am^4 + \frac{1}{3}bm^3 + \frac{1}{2}cm^2 \tag{4.9}$$

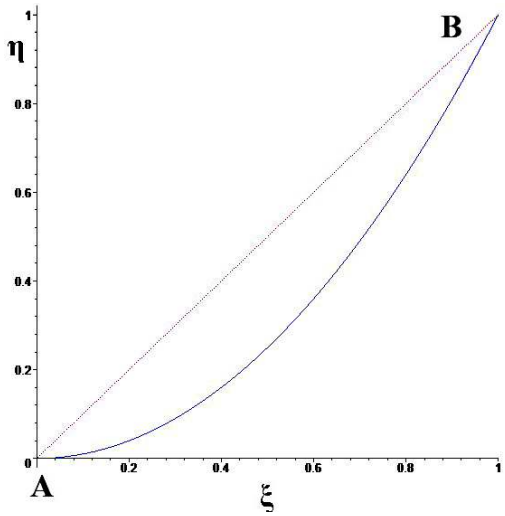
Besides, there is another assumption to get the equation of minimization solved. We assume that the first derivative of the polynomial (Equation 4.8) at the point A is zero to get the maximum distance from A. Then, the parameters of this polynomial are obtained as follows.

$$\frac{\partial I}{\partial m} = 0 \implies a = \frac{n}{m^3},\ b = 0,\ c = 0,\ d = 0. \tag{4.10}$$

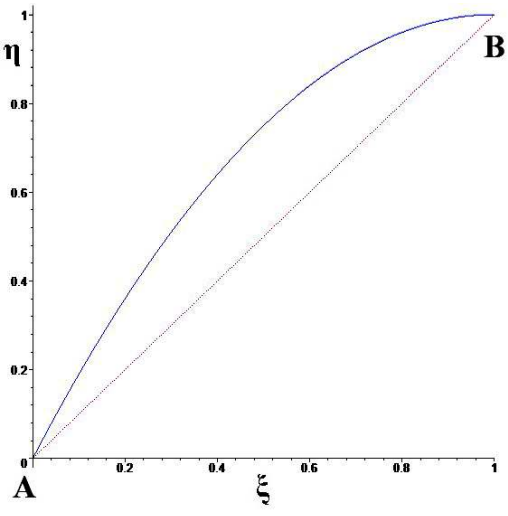
Therefore, Equation 4.11 presents the most uncertain third order response of the LSE regarding its value in horizontal axis.

$$\eta = g(\xi) = \frac{n}{m^3}\xi^3 \tag{4.11}$$

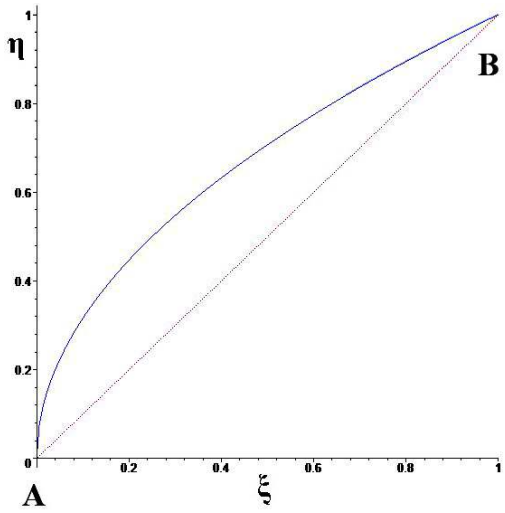
The integration also should be maximized in order to get the upper bound of the response of the LSE. Besides, the polynomial should be minimized and maximized



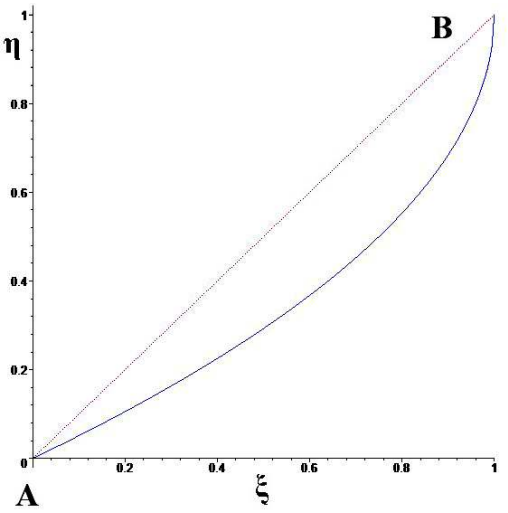
(a) 2nd order response of the LSE, minimized with respect to ξ



(b) 2nd order response of the LSE, maximized with respect to ξ

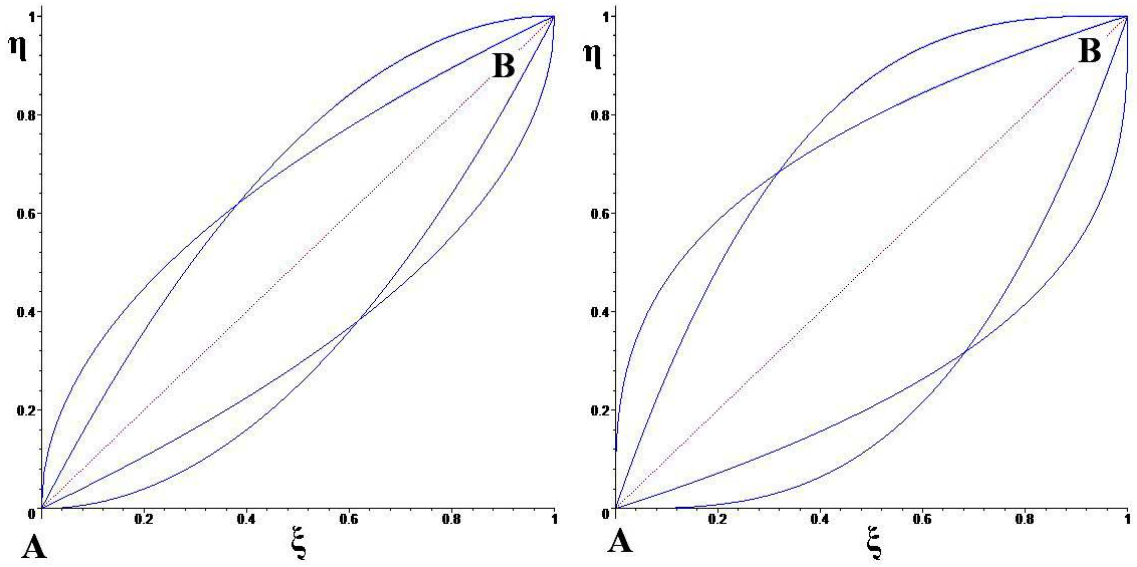


(c) 2nd order response of the LSE, minimized with respect to η



(d) 2nd order response of the LSE, maximized with respect to η

Figure 4.3: (a,b) Two bounds for a second order response of the limit state equation (LSE) for the most uncertain condition in respect to ξ. (c,d) Two bounds for a second order response of the LSE for the most uncertain condition in respect to η. ξ and η are local coordinates.



(a) Extreme responses of a 2nd order LSE remains inside the enclosed region.

(b) Extreme responses of a 3rd order LSE remains inside the enclosed region.

Figure 4.4: Extreme responses of (a) 2nd order LSE and (b) 3rd order LSE, given points Aand B.

regarding the vertical axis, η. Figure 4.4(b) presents a drawing of the LSEs for all the extreme responses. In other words, the response of a LSE given a third order continuous polynomial can not be out of the region, enclosed by the curves as shown in Figure 4.4(b).

It might be interesting to present the lower bound of the response for higher order of continuous and smooth polynomials. Figure 4.5 shows different bounds from a linear LSE towards the LSE with order of 150 in which the integration is minimized.

Information of the first derivative

Given some information over the derivatives at the start and end point of the unqualified region, points A and B, a higher order polynomial response of the LSE can be implemented with a smaller uncertainty. In fact, apart from the order of the polynomials, there are still some information from the neighbors which have not yet been considered. Since the dynamic bounds (DB) process is based upon the shrinkage of the unqualified region, there will be some neighboring points available. This fact encourages us to take the closest neighbor into account by the concept of derivatives. For demonstration, we assume a fourth order polynomial

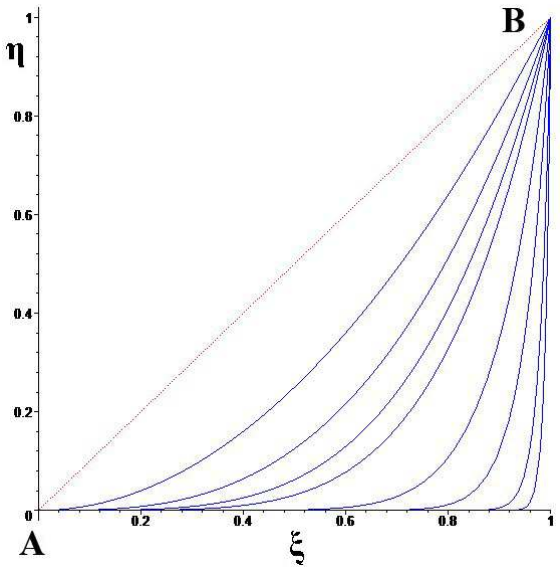


Figure 4.5: Lower bounds for the higher order of the response. Apart from the linear response, the order of polynomials are respectively equal to 2, 3, 4, 5, 10, 20, 50, 100 and 150.

in its general form as

$$\eta = a\xi^4 + b\xi^3 + c\xi^2 + d\xi + e \tag{4.12}$$

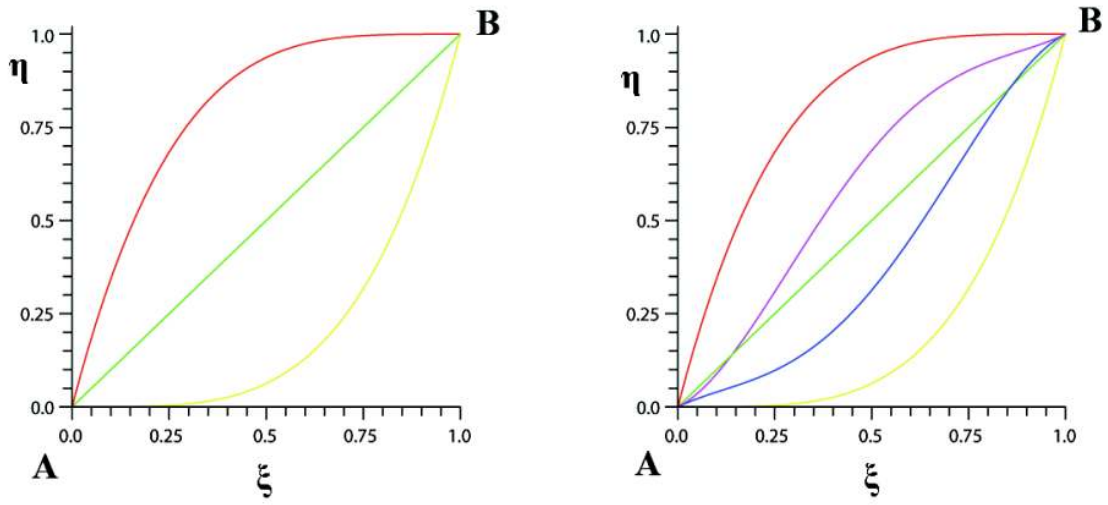
This polynomial should pass from points A and B, where $A = (0,0)$ and $B = (m,n)$. It concludes that $e = 0$, and another equation between the other parameters. Therefore, we aim to minimize Equation 4.13.

$$I = \int_0^m \eta d\xi = \frac{1}{5}am^5 + \frac{1}{4}bm^4 + \frac{1}{3}cm^3 + \frac{1}{2}dm^2 \tag{4.13}$$

Also we assume that the value of its first derivatives at points A and B are calculated and they are respectively equal to p and q. As a result of this minimization, the parameters of the polynomial are obtained according to the following equation.

$$\begin{aligned} \frac{\partial I}{\partial m} &= 0 \implies \\ a &= -\frac{mp - mq + 3n}{m^4},\ b = -\frac{-4n - 3mp + mq}{m^3},\ c = -3\frac{p}{m},\ d = p,\ e = 0 \end{aligned} \tag{4.14}$$

As a result, the lower bound of the response of the LSE is presented in Equation 4.15, and depicted in Figure 4.6. The result of maximization of Equation 4.12 in respect to the η is also included in this figure.



(a) Lower and upper bounds for the fourth order response

(b) Lower and upper bounds for the fourth order response including the derivatives information

Figure 4.6: Figure (a) presents two limit bounds for the responses of the limit state function assuming the response is fourth order. Figure (b) presents the same bounds assuming the information of the first derivatives in points A and B.

$$\eta = g(\xi) - \frac{(mp - mq + 3\,n)\,\xi^4}{m^4} - \frac{(-4\,n - 3\,mp + mq)\,\xi^3}{m^3} - 3\,\frac{p\xi^2}{m} + p\xi$$

To clarify the effect of applying information of the first derivatives, a comparison is provided in Figure 4.6. In this case, Figure 4.6(a) presents the fourth order upper and lower bounds for the response of the LSE (in respect to the ξ), and Figure 4.6(b) presents the bounds including the derivatives information. In other words, the value of derivatives can effectively reduce the enclosed region which is, in fact, the unqualified region.

4.3 Extended bounds

Having the boundaries of the response of a LSE, it is possible to extend the stable or unstable regions without a realization of the LSE. We call the the outcomes as the extended stable region and extended unstable region, respectively shown as S_b^e and U_b^e, where:

$$S_b^e \subset S \quad \text{and} \quad U_b^e \subset U \tag{4.15}$$

$$S_b \subseteq S_b^e \quad \text{and} \quad U_b \subseteq U_b^e \tag{4.16}$$

In other words, assuming a monotonic behavior and a specific order of the polynomial, it is possible to extend the bounds from both sides of stable and unstable regions. Given point A on the unstable bound (see Figure 4.2), we know that $G(\vec{x_A}) \leq 0$. Its value is also equal to a constant C, $C = G(\vec{x_A})$. Also we concluded from Section 4.2 that given the order of a monotone LSE, we can find its extreme responses in the form presented in Equation 4.17. $\vec{x_A} = (x_{1A}, \ldots, x_{nA})$ is a vector and every element of this vector presents the value of one of its coordinates. Therefore, i Equation 4.17 represents the dimension of the vector, while j represents the different extreme responses of the LSE which were addressed in Section 4.2. Therefore, n represents the dimension and m represents the number of different extreme responses; they are different with what was presented in Section 4.2.

$$\eta_{ij} = g_j(\xi_i), \quad \text{where} \quad i = 1, \ldots, n \quad \text{and} \quad j = 1, \ldots, m \tag{4.17}$$

The next step is finding the possible extension of the bounds shown as ξ_{ij}^0

$$|C| = g_j(\xi_{ij}^0), \quad i = 1, \ldots, n \quad \text{and} \quad j = 1, \ldots, m \tag{4.18}$$

Where

$$C = G(\vec{x}_A) \quad \text{and} \quad \vec{x_A} \in U \tag{4.19}$$

Then we can find the point, where the LSE, estimated to be zero from each curve as:

$$\xi_{ij}^0 = g_j^{-1}(\eta_{ij}) = g_j^{-1}(|C|) \quad i = 1, \ldots, n \quad \text{and} \quad j = 1, \ldots, m \tag{4.20}$$

And finally,

$$\xi_i^e = \min\{\xi_{ij}^0, \quad j = 1, \ldots, m\} \tag{4.21}$$

As a result, $x_{iA}{}' = x_{iA} + \xi_i^e$ presents the extended unstable bound. Therefore, without implementing any uncertainty the lower bound is extended as U^e. This process, can be also done for the stable bound in which the result will be in the form of $x_{Bi}{}' = x_{Bi} - \xi_i^{e\prime}$, where $x_B{}'$ suggests the extended bound of the stable region and $\xi^{e\prime}$ is obtained from a similar process of calculating ξ^e. However, if no information of derivation is implemented, we expect to have $\xi^{e\prime} = \xi^e$. In conclusion, $\vec{x_A} = (x_{1A}, \ldots, x_{nA})$ which represents a point on the unstable bound is transformed to $\vec{x_A}{}' = (x'_{1A}, \ldots, x'_{nA})$, and $\vec{x_B} = (x_{1B}, \ldots, x_{nB})$ which represents a point on the stable bound is transformed to $\vec{x_B}{}' = (x'_{1B}, \ldots, x'_{nB})$ without any realization of the LSE.

4.4 Monte Carlo Algorithm

Here we describe the algorithm of improved dynamic bounds (IDB) assuming the requirements of dynamic bounds (DB) are already fulfilled (reader may review

Section 3.2.3). In this algorithm we refer to a two dimensional problem because it is easier to illustrate and it has been already studied for the method of dynamic bounds.

To begin, let us consider the ideas of the Monte Carlo algorithm by looking at a 2-dimensional example. In Figure 4.1, a two-dimensional limit state equation $G(x_1, x_2) = 0$ is depicted, as well as the contours of the joint probability density function, $f(x_1, x_2)$, of the two variables (X_1, X_2). The limit state function $G(x_1, x_2)$ is assumed to be monotonically increasing in both variables. This means that $G(x_1, x_2) < 0$ for points below the LSE; we call this the *unstable region* as its internal points correspond to the failure. For points above the LSE, $G(x_1, x_2) > 0$, it is called the *stable region*.

A first random point, $\vec{x}^{(1)} = (x_1^{(1)}, x_2^{(1)})$, is generated from the JPDF f. In Figure 4.2, it is depicted by a black square labeled 1, and upon evaluation it is found that $G(x_1^{(1)}, x_2^{(1)}) < 0$, hence it leads to the failure and defines the unstable region. From the monotonicity of G it is inferred that $G(x_1, x_2) < 0$ for all points (x_1, x_2) in the quadrant to the left and below of $(x_1^{(1)}, x_2^{(1)})$. In the next step, the point $(x_1^{(2)}, x_2^{(2)})$ is generated from f, it turns out that $G(x_1^{(2)}, x_2^{(2)}) > 0$. So point 2 is in the stable region and all points the quadrant to the right and above of point 2 are stable as well.

This process continues; the result of a small number of iterations is shown in Figure 4.2. The shaded regions constitute approximations to the stable and unstable regions that can be used to obtain bounds on the probability of failure p_f.

These regions are limited within the stable and unstable bounds. Therefore, we keep the points of the bounds and leave the internal points. The result, however, is equivalent. Suppose two closest points on the unstable and stable bounds named respectively as $\vec{u_b}^{(i)}$ and $\vec{s_b}^{(j)}$, where in a two dimensional example are $\vec{u_b}^{(i)} = (u_{b1}^{(i)}, u_{b2}^{(i)})$ and $\vec{s_b}^{(j)} = (s_{b1}^{(j)}, s_{b2}^{(j)})$. Then we implement the technique presented in Section 4.3 to get these points (bounds) extended to $\vec{u_b}^{(i)\prime}$ and $\vec{s_b}^{(j)\prime}$. For illustration, we mention to the updating process of these points. As a result of transformation to a one dimension and adjusting the local coordinates, we will have $A = (0, 0)$ and $B = (c, d)$, where $c = s_{b1}^{(j)} - u_{b1}^{(i)}$ and $d = G(\vec{s}_{b1}^{(j)}) - G(\vec{u}_{b1}^{(i)})$. As a result the value of ξ_1^e can be calculated, and within the same process, the value of ξ_2^e can be calculated, where 1 and 2 refer to the dimensions. These values lead toward the extension of the stable and unstable bounds as schematically shown in Figure 4.2 by the points $\vec{u_b}^{(i)\prime}$ and $\vec{s_b}^{(j)\prime}$. Then this process needs to be done for the other boundary points[3].

[3]The reader should keep in mind that in practice the location of the LSE-curve is not known; it only gradually becomes visible as it is being sandwiched between the increasingly more accurate approximations to the stable and unstable regions. The corresponding dynamic upper and lower bounds on p_f become tighter as the number of generated points increases. Also note that for some generated points it is not necessary to evaluate the limit state function, because from other points

For the mathematical description we revert to dimension n. Define the stable bound S and the unstable bound U for a monotonically increasing $G(\vec{x})$ by

$$S_b = \{\vec{x} \ : \ G(\vec{x}) \geq 0 \ \mid \forall x, y \in S, \exists x_i < y_i, \ i = 1, \ldots, n\}. \tag{4.22}$$

$$U_b = \{\vec{x} \ : \ G(\vec{x}) < 0 \mid \forall x, y \in U, \exists x_i > y_i, \ i = 1, \ldots, n\}. \tag{4.23}$$

Equations 4.22 and 4.23 are equivalent to the definition of S and U presented in Equation 3.2, and the difference is that the sets addressed here are only include the value on the bounds.

Consider the k-th iteration of the Monte Carlo process. A number of points on the stable bound, say, $\vec{s_b}^{(1)}, \ldots, \vec{s_b}^{(p)}$ and a number of points on the unstable bound $\vec{u_b}^{(1)}, \ldots, \vec{u_b}^{(q)}$ have been stored. The current approximation to the stable region S is the union of the p orthants (generalizing the quadrants in Figure 3.1)

$$H_i = \{\vec{x} : \vec{x} \text{ is more stable than } \vec{s_b}^{(i)}\}, \quad i = 1, \ldots, p. \tag{4.24}$$

If $\vec{s}^{(i)}$ is more stable than $\vec{s_b}^{(j)}$, for some i and j, its orthant H_i is completely contained in H_j, and there would be no loss of information if $\vec{s}^{(i)}$ would be dropped from the list. Similarly, the current approximation to the unstable region U is the union of

$$L_i = \{\vec{x} : \vec{x} \text{ is less stable than } \vec{u_b}^{(i)}\}, \quad i = 1, \ldots, q. \tag{4.25}$$

From now on, it is assumed that only set of stable and unstable points located on the bounds are retained during the simulation and S_{bk} and U_{bk} are the corresponding approximations to S and U:

$$S_{bk} = \cup_{i=1}^{p} H_i \quad \text{and} \quad U_{bk} = \cup_{i=1}^{q} L_i. \tag{4.26}$$

Now, imagine the next random point $\vec{X}^{(k+1)}$ is generated from f. There are three possibilities. The first is: $\vec{X}^{(k+1)} \in S_{bk}$, that is, the point is located in a region that is known to be part of the stable set. The second is: $\vec{X}^{(k+1)} \in U_{bk}$, that is, the point is located in a region that is known to be part of the unstable set; the count of the number of failures should be incremented. The third is: $\vec{X}^{(k+1)} \notin S_{bk} \cup U_{bk}$, that is, the point is located in the unqualified region between S_{bk} and U_{bk}; $G(\vec{X}^{(k+1)})$ needs to be evaluated; but before the evaluation we extend the bounds and obtain the extended bounds, S^e_{bk} and U^e_{bk}; $\vec{X}^{(k+1)} \in S^e_{bk}$, that is, the point is located in a region that is known to be part of the stable set; $\vec{X}^{(k+1)} \in U^e_{bk}$, that is, the point is located in a region that is known to be part of the unstable set; the count of the number of failures should be incremented. The fourth is: $\vec{X}^{(k+1)} \notin S^e_{bk} \cup U^e_{bk}$, that is, the point is still located in the unqualified region between S^e_{bk} and U^e_{bk}; $G(\vec{X}^{(k+1)})$ needs to be evaluated; If it is positive, $\vec{X}^{(k+1)}$ is added to the collection of the points on the stable bound and this collection is checked for its minimality,

it can be determined whether the point is stable or unstable.

dropping any superfluous points. If it is negative, $\vec{X}^{(k+1)}$ is added to the collection of the points on the unstable bound and a similar update is performed. Note that the numbers p and q vary during the simulation and in fact depend on the iteration number k.

Summarizing, the algorithm is as follows:

1. Determine S_{b0} and U_{b0}. They could be empty sets or be determined from the threshold points, as described. Set $k = 0$, $n_f = 0$.

2. Increase k by 1 and generate $\vec{X}^{(k)}$ from f. If $\vec{X}^{(k)} \in U_{bk-1}$, add 1 to n_f and update U_{bk-1} to obtain U_k. If $\vec{X}^{(k)} \notin S_{bk-1} \cup U_{bk-1}$, calculate S^e_{bk-1} and U^e_{bk-1}; If $\vec{X}^{(k)} \in U^e_b$, add 1 to n_f and update U_{bk-1} by adding U^e_{bk-1} to obtain U_{bk}; else if $\vec{X}^{(k)} \in S^e_b$, update S_{bk-1} by S^e_{bk-1} to obtain S_{bk}. If $\vec{X}^{(k)} \notin S_{bk-1} \cup S^e_b \cup U_{bk-1} \cup U^e_b$, evaluate $G(\vec{X}^{(k)})$; if it is negative, add 1 to n_f and update U_{bk-1} to obtain U_{bk}; otherwise, update S_{bk-1} to obtain S_{bk}.

   Repeat until $k = N$.

3. $\hat{p}_f = n_f/N$ is a simple Monte Carlo estimate for p_f.

This estimate $\hat{p}_f$ is as good as an ordinary Monte Carlo estimate based on N independent samples $\mathbf{1}[G(\vec{X}^{(1)}) < 0]$,...,$\mathbf{1}[G(\vec{X}^{(N)}) < 0]$, but requires evaluation of G in only a fraction of the samples. However, the simulation also yields an upper and a lower bound on p_f, as follows. Clearly, $U^e_{bN} \subset U$ and $S^e_{bN} \subset S$. So,

$$\hat{p}^e_u := \mathrm{P}\left(\vec{X} \in U^e_{bN}\right) \le p_f \quad \text{and} \quad \hat{p}^e_s := \mathrm{P}\left(\vec{X} \in S^e_{bN}\right) \le 1 - p_f, \tag{4.27}$$

where $\vec{X}$ is an independent draw from f. These imply the following bounds on p_f:

$$\hat{p}^e_u \le p_f \le 1 - \hat{p}^e_s. \tag{4.28}$$

Therefore, the unqualified region has been shrank which can be presented as

$$p^s_\Delta = 1 - \hat{p}^e_u - \hat{p}^e_s. \tag{4.29}$$

Where p^s_Δ presents the shrank of unqualified region and this statement is valid.

$$p^s_\Delta \le p_\Delta \tag{4.30}$$

4.5 Numerical example

4.5.1 One dimensional model

Here, we refer to the example of the impact of water waves on coastal structures which was described in Section 3.4 and presented in Equation 3.16. A one

	A comparison between DB and the other level III methods						
IDB	N	4	**9**	15	16	18	22
	$\hat{p}_f$	0.04	**0.047**	0.0462	0.0466	0.0466	0.0466
	$V(\hat{p}_f)$	14.5E-2	**4.3E-2**	1.44E-2	0.45E-2	0.14E-2	0.045E-2
DB	N	9	**16**	22	27	33	36
	$\hat{p}_f$	0.045	**0.0497**	0.0462	0.0464	0.0467	0.0466
	$V(\hat{p}_f)$	14.5E-2	**4.3E-2**	1.44E-2	0.45E-2	0.14E-2	0.045E-2
MC	N	1E+3	**1E+4**	1E+5	1E+6	1E+7	1E+8
	$\hat{p}_f$	0.0487	**0.0478**	0.0462	0.0466	0.0466	0.0466
	$V(\hat{p}_f)$	14.5E-2	**4.3E-2**	1.44E-2	0.45E-2	0.14E-2	0.045E-2

Table 4.1: A comparison between different level III methods: Monte Carlo (MC), dynamic bounds (DB), Improved dynamic bounds (IDB).

dimensional limit state equation (LSE) of that example can be defined by Equation 4.31, where the velocity parameter is assumed to be normally distributed as $N(1.5, 0.45)$.

$$G(u) = 500000 - 98280 \times u^2. \tag{4.31}$$

It is clear that this limit state function is monotonically decreasing regarding its variable (u), and it is a second order function. This assumption may be obtained by experience or knowledge about an implicit limit state equation. Having this knowledge, we compare the results of improved dynamic bounds (IDB), dynamic bounds (DB), and Monte Carlo (MC) in Table 4.1. This shows that given more information over LSE, less calculation efforts needed. It is important to be reminded that the information of derivatives still has not been applied to this comparison. As a result, IDB is robust and can be used with a good efficiency and a high accuracy.

4.5.2 Two dimensional model

Here we refer to the two-dimensional limit state equation presented in Equation 3.17 as

$$G(\rho, u) = 500000 - 94.50 \times \rho \times u^2. \tag{4.32}$$

There are some extra information can be implemented into the reliability analysis. Given, the linear relation of the ρ and second order relation of u, the method of improved dynamic bounds can be implemented. The improved dynamic bounds and some other Level III reliability methods are applied to the LSE (Equation 4.32), the results are given in Table 4.2. For each method there are several rows, the first shows the number of evaluations of the LSE, the second the corresponding estimate for p_f, and the third shows the *relative standard error*, i.e., the standard error divided by p_f. The number of evaluations is varied, but the columns contain

	A comparison between DB and the other level III methods						
IDB	6	13	14	12	13		
$\hat{p}_f$	0.0400	0.0456	0.0457	0.0463	0.0466		
Rel. s.e.	0.1549	0.457E-1	3.44E-3				
$p_f \geq \hat{p}_u^e =$	0.0432	0.0465	0.0466	0.0466	0.0466		
$p_f \leq 1 - \hat{p}_s^e =$	0.0479	0.0467	0.0466	0.0466	0.0466		
DB	27	**77**	203	598	1828	5556	17286
p_f	0.0490	**0.0471**	0.0455	0.0468	0.0467	0.0466	0.0466
Rel. s.e.	13.93E-2	**4.5E-2**	1.45E-2	0.45E-2	0.14E-2	0.045E-2	0.014E-2
$p_f \geq \hat{p}_u =$	0.042	0.0451	0.0463	0.0466	0.0466	0.0466	0.0466
$p_f \leq 1 - \hat{p}_s =$	0.053	0.0490	0.0472	0.0469	0.0466	0.0466	0.0466
MC	1E+3	1E+4	1E+5	1E+6	1E+7	1E+8	1E+9
p_f	0.038	0.0477	0.0462	0.0467	0.0467	0.0466	0.0466
Rel. s.e.	15.9E-2	4.5E-2	1.44E-2	0.45E-2	0.14E-2	0.045E-2	0.014E-2
NI	1E+3	1E+4	1E+5	1E+6	1E+7	1E+8	1E+9
p_f	0.0429	0.0465	0.0466	0.0466	0.0466	0.0466	0.0466
IS	1E+2	1E+3	1E+4	1E+5	1E+6		
p_f	0.0412	0.0464	0.0467	0.0468	0.0466		
Rel. s.e.	15.8E-2	4.64E-2	1.42E-2	0.45E-2	0.14E-2		
IS-DB	12	55	136	271	959		
p_f	0.0425	0.0460	0.0469	0.0468	0.0467		
Rel. s.e.	16.4E-2	4.58E-2	1.44E-2	0.64E-2	0.17E-2		

Table 4.2: Comparison between level III methods: Improved dynamic bound (IDB), dynamic bounds (DB), Monte Carlo (MC), numerical integration (NI), important sampling (IS), and DB coupled with IS (IS-DB). Rows marked p_f contain the estimates for the probability of failure. *Rel. s.e.* is the relative standard error. The bottom two DB rows present $\hat{p}_u$ and $1 - \hat{p}_s$: certain (i.e., non-probabilistic) bounds on p_f. Bold numbers present the acceptable engineering level.

outcomes of approximately the same accuracy (except for the NI results, that are matched with MC in terms of number of evaluations of the LSE). For the IDB and DB methods the last two rows present the lower and upper bound from equation (3.7) (the reader is reminded that these bounds are not confidence bounds, but hold with certainty).

Table 4.2 shows that the number of Monte Carlo simulations has been reduced by the factor of 9 and 1E+4, which is considerable.

4.6 Conclusions

Dynamic bounds (DB) is a technique suggested for the reliability analysis of engineering problems which have monotonic behavior with a limited number of

influential variables. The DB method divides the range of a LSE into three parts of stable, unstable, and unqualified. An attempt was made in this chapter to shrink the size of the unqualified part, given the order of the LSE in respect to its variables. As a result, DB is improved by defining a lower bound for the response surface in the unqualified region of the LSE on the base of the minimization process. The improved dynamic bounds (IDB) presents more efficiency and accuracy without implementing uncertainty into the model. IDB still can be coupled with importance sampling technique, and more importantly the bounds can be stored for the next series of simulations.

5

Bayesian Monte Carlo

To reduce the cost of Monte Carlo (MC) simulations for time-consuming processes (like Finite Elements), Bayesian Monte Carlo (BMC) is introduced in this chapter. Applying this method, number of realizations in MC is reduced according to an accepted tolerance. Besides, there is a possibility of thinking about priors. In other words, different priors can be integrated into one model which can decrease calculation efforts in MC simulations. This study tries to speed up the Monte Carlo process by taking into account the information of neighboring points as a prior knowledge about the problem. As a matter of fact, this information provides a predictive tool for the next estimation; therefore, given a certain tolerance, the number of calculations are dramatically reduced. The general methodology and algorithm of the suggested approach are also presented in this chapter and further details are presented in Appendixes.

5.1 Overview

Monte Carlo technique, widely applied in engineering fields, considers each simulation independent of the previous simulations. As a result, a number of techniques have been developed to take into account the information of the previous simulations and (as a result) reduce the cost of simulations. These improvements are necessary for the simulation of a time-consuming process (like FE) indicated for instance by Rajabalinejad et al. (2007a). These methods and their advantages and disadvantages are briefly discussed in Chapter 2. Besides, the technique of dynamic bounds (DB) is introduced in Chapter 3, and its improvement is described in Chapter 4. This chapter presents a method to speed up the Monte Carlo process by considering the assumption that the value of every point (pixel) can be estimated by its neighboring pixels, given the position of the pixels and the best possible estimation. This assumption can be implemented into the model

by the Bayesian technique which is, in fact, based on the Bayesian interpolation technique presented by Bretthorst (1992). The Bayesian interpolation assigns a probability density function (PDF) to an arbitrary point, given an evenly spaced grid. This PDF, then, can be used as a judgement tool. Therefore, given a tolerance we will be able to judge about the accuracy of the prediction for the next simulation. The Bayesian Monte Carlo (BMC) still is based upon the interpolation, but it is fully adapted with the Monte Carlo method. As a result, the information of the neighborhoods are implemented into the model, and there is a possibility for implementing of more priors.

This chapter needs a background of Bayesian material as well as Monte Carlo methods. The concept of Bayesian approaches is very well described by Jaynes (2003). An instructive reference for applying the Bayes' Theorem into the practice is presented by Sivia (1996). It is assumed that the readers have background of the Monte Carlo method and further detailed of this method can be found in Melchers (1999); Hammersley and Handscomb (1964).

5.2 Introduction

To address the problem we assume that there is a signal U which is to be estimated at a number of discrete points. These discrete points will be called pixels, presented by u_i. These pixels present a vector of pixels $\mathbf{u} \equiv (u_0, \cdots, u_{v+1})$. Therefore, there are totally $v+2$ pixels. The first and last pixels, presented by u_0 and u_{v+1}, are called boundary pixels and are treated separately. As a result, v presents the number of interior pixels. The total number of observed data points is equal to n which are distributed in arbitrary (or random) locations among the pixels. Therefore, the maximum value of n is equal to $v+2$ when there is an observed data point for each pixel ($n \leq v+2$). The locations of the observed data points are collected in a vector $\mathbf{c}$, so this vector has n elements which are presented by c_i and $i = 1, 2, \cdots, n$. The vector of observed data points is called $\mathbf{d} \equiv (d_1, \cdots, d_n)$, and its elements are presented by d_i. Figure 5.1 presents an illustration of the internal and boundary pixels as well as data points. According to this figure, for instance, $\mathbf{c} \equiv (1, v-1, v+2)$.

The univariate probability density function (PDF) for an arbitrary pixel, given the data D and the informational context I, will be found by integrating out all pixels. In this case the sum rule is applied and the product is integrated all over the multivariate posterior pdf of all pixels of U except the required pixel u_j.

$$P(u_j|D,I) = \int P(U|D,I) \underbrace{\ldots du_i \ldots}_{i \neq j} \tag{5.1}$$

Also, according to the Bayes' rule we have:

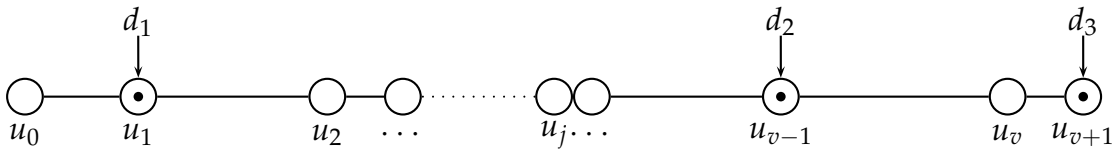


Figure 5.1: An illustration of the pixels which data points are assigned to.

$$P(U|D,I) = \frac{P(D|U,I)P(U|I)}{P(D|I)} \tag{5.2}$$

Where $P(D|I)$ is a normalization constant. Therefore, combination of Equations 5.1 and 5.2 produces the following equation.

$$P(u_j|D,I) \propto \int P(D|U,I)P(U|I) \underbrace{\ldots du_i \ldots}_{i \neq j} \tag{5.3}$$

This equation presents needs to define $P(D|U,I)$ and $P(U|I)$ which are respectively called likelihood function and the prior. The likelihood, or in this case more appropriate the PDF of the data (D) conditional on the pixels (U), is constructed by making the standard assumptions of noise. It is assumed that this noise has got a zero mean value and a specific standard deviation. As a result, it is important to define the prior on the base of the prior information we have.

5.3 The prior

There are some logical dependence among neighboring pixels and this expectation is translated in the following model, f, for an arbitrary pixel u_i. In this model, the value of u_i is estimated by its two neighbor points. Figure 5.2 clarifies this concept where two neighbor points of pixel i are shown, and their positions are shown by x. Given an arbitrary location in $[x_{i-1} \ , \ x_{i+1}]$, it is logical to assume that a closer neighbor has more influence on the estimate than the other neighbor.

Therefore, $\delta_{r,i}$ and $\delta_{l,i}$ are two relative weights which respectively present the influence of left and right neighbor points on the target pixel, and one gets $\delta_{l,i} + \delta_{r,i} = 1$. As a result, the value of the pixel u_i is estimated as

$$\hat{u}_i = f(u_{i-1}, u_{i+1}) = u_{i-1}.\delta_{r,i} + u_{i+1}.\delta_{l,i}, \tag{5.4}$$

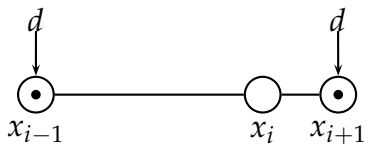


Figure 5.2: An illustration of the pixels which data points are assigned to.

where

$$\delta_{l,i} = \frac{x_i - x_{i-1}}{x_{i+1} - x_{i-1}}, \quad \text{and} \quad \delta_{r,i} = \frac{x_{i+1} - x_i}{x_{i+1} - x_{i-1}}.$$

Having the model defined, the error e_i also is implicitly defined by Equation 5.5.

$$e_i = u_i - \hat{u}_i = u_i - f(u_{i-1}, u_{i+1}) = u_i - (u_{i-1}.\delta_{r,i} + u_{i+1}.\delta_{l,i}). \tag{5.5}$$

The only thing we know about this error is that the error has a mean of zero (the error is either positive or negative) with some unknown variance ϕ_i^2. It means that we assume that the standard deviation of our error is relevant to the distance of neighbors, $\phi_i \propto (x_{i+1} - x_{i-1})$. As a matter of fact, a closer neighbor points to the pixel a smaller error for the estimation is expected. Using the principle of Maximum Entropy Jaynes (2003), we find the well known Gaussian probability density function of e_i presented in Equation 5.6. ϕ_i in this equation stands for the standard deviation of the pixel u_i.

$$P(e_i|\phi_i) = \frac{1}{\sqrt{2\pi}\phi_i} \exp[-\frac{1}{2\phi_i^2} e_i^2]. \tag{5.6}$$

Substituting Equation 5.5 into Equation 5.6 and making the appropriate change of variable from e_i to u_i, the PDF of the pixel u_i can be obtained by Equation 5.7.

$$\begin{aligned} &P(u_i|u_{i-1}, u_{i+1}, \phi_i, I) \\ &\quad = \frac{1}{\sqrt{2\pi}\phi_i} \exp[-\frac{1}{2\phi_i^2}[u_i - u_{i-1}.\delta_{r,i} - u_{i+1}.\delta_{l,i}]^2]. \end{aligned} \tag{5.7}$$

Assuming that there is no logical dependence between the errors $e_1, \ldots, e_v$, where v is the number of internal pixels, the multivariate PDF of all the errors is a product of the univariate PDFs. Then, by making the change of variable from e_i to u_i we find the following multivariate PDF for the pixels $u_1, \ldots, u_v$.

$$\begin{aligned} &P(u_1, \ldots, u_v|u_0, u_{v+1}, \phi_1, \ldots, \phi_v, I) \\ &\quad \propto \frac{1}{(2\pi)^{v/2}\phi_1 \ldots \phi_v} \exp[-\sum_{i=1}^{v} \frac{1}{2\phi_i^2}[u_i - u_{i-1}.\delta_{r,i} - u_{i+1}.\delta_{l,i}]^2]. \end{aligned} \tag{5.8}$$

To simplify our approach, we assume that the value of standard deviation or ϕ is in direct relation with the relative distance of pixels. In other words, we assume that the closer pixels we get, a smaller standard deviation is expected.

$$\begin{aligned} &\phi_i = \tfrac{\delta_i}{\Delta} \times \phi, \quad 1 \leq i \leq v, \\ &\Delta = \delta_0 + \delta_1 + \cdots + \delta_v + \delta_{v+1}, \\ &\delta_i = x_{i+1} - x_{i-1}. \end{aligned} \tag{5.9}$$

Also,

$$\begin{aligned} &\phi_0 = \tfrac{\delta_0}{\Delta} \times \phi, \quad \text{and} \quad \delta_0 = x_1 - x_0, \\ &\phi_{v+1} = \tfrac{\delta_{v+1}}{\Delta} \times \phi, \quad \text{and} \quad \delta_{v+1} = x_{v+1} - x_v. \end{aligned} \tag{5.10}$$

Remark. It is important to attend to the difference between δ_i and $\delta_{l,i}$ or $\delta_{r,i}$. δ_i refers to the absolute distance while the $\delta_{l,i}$ and $\delta_{r,i}$ refer to a relative weight. We substitute Equation 5.9 to 5.8 and we get the following equation:

$$\begin{aligned} &P(u_1, \ldots, u_v | u_0, u_{v+1}, \phi, I) \\ &\quad \propto \tfrac{1}{(2\pi)^{v/2} A \phi^v} \exp[-\tfrac{\Delta^2}{2\phi^2} \textstyle\sum_{i=1}^{v} \tfrac{1}{\delta_i^2} [u_i - u_{i-1}.\delta_{r,i} - u_{i+1}.\delta_{l,i}]^2], \\ &\text{where} \quad A = \tfrac{\delta_1 \ldots \delta_v}{(\delta_1 + \cdots + \delta_v)^v}. \end{aligned} \tag{5.11}$$

The boundary pixels are treated separately. In fact, these two pixels are assigned to the first and last position and presented as $u_0 = v_1$ and $u_{v+1} = v_v$. As a result of using the principle of Maximum Entropy, the PDF of the boundary pixel u_0 is obtained in Equation 5.12. And a similar equation can be established for the other boundary pixel u_{v+1}.

$$\begin{aligned} &P(u_0 | u_1, \phi_0, I) \\ &\quad = \tfrac{1}{\sqrt{2\pi}\phi_0} \exp[-\tfrac{1}{2\phi_0^2} [u_0 - u_1]^2] \\ &\quad = \tfrac{\Delta}{\sqrt{2\pi}\phi\delta_0} \exp[-\tfrac{\Delta^2}{2(\phi\delta_0)^2} [u_1 - u_0]^2] \\ &\quad = \tfrac{\Delta}{\sqrt{2\pi}\phi\delta_0} \exp[-\tfrac{\Delta^2}{2\phi^2} [\tfrac{u_1 - u_0}{\delta_0}]^2]. \end{aligned} \tag{5.12}$$

And

$$\begin{aligned} &P(u_{v+1} | u_v, \phi_{v+1}, I) \\ &\quad = \tfrac{1}{\sqrt{2\pi}\phi_{v+1}} \exp[-\tfrac{1}{2\phi_{v+1}^2} [u_{v+1} - u_v]^2] \\ &\quad = \tfrac{\Delta}{\sqrt{2\pi}\phi\delta_{v+1}} \exp[-\tfrac{\Delta^2}{2(\phi\delta_{v+1})^2} [u_{v+1} - u_v]^2] \\ &\quad = \tfrac{\Delta}{\sqrt{2\pi}\phi\delta_{v+1}} \exp[-\tfrac{\Delta^2}{2\phi^2} [\tfrac{u_{v+1} - u_v}{\delta_{v+1}}]^2]. \end{aligned} \tag{5.13}$$

To combine these priors, the product rule can be used to combining Equations 5.8, 5.12, and 5.13 by applying of Bayes' Theorem,

$$\begin{aligned} &P(u_0, u_1, \ldots, u_{v+1} | \phi, I) \\ &\quad \propto P(u_0 | u_1, \phi, I) P(u_{v+1} | u_v, \phi, I) P(u_1, \ldots, u_v | u_0, u_{v+1}, \phi, I). \end{aligned} \tag{5.14}$$

The next equation is obtained which is written in a matrix form, where **u** is vector of pixel positions, and **R** is presented by Equation 5.16.

$$P(u_0, u_1, \ldots, u_{v+1}|\phi, I) \propto \frac{1}{(2\pi)^{(v+2)/2} A' \phi^{v+2}} \exp[-\frac{1}{2\phi^2} \sum_{k=0}^{v+1} \sum_{l=0}^{v+1} R_{kl} u_k u_l], \tag{5.15}$$

where

$$A' = \frac{\delta_0.\delta_1 \ldots \delta_v.\delta_{v+1}}{(\delta_1 + \cdots + \delta_v)^{v+2}}.$$

We have derived the above equation which provides the joint probability density function (JPDF) for the pixels $u_0, \ldots, u_{v+1}$ by implementing the prior based upon the weighted neighbors presented in Equation 5.4.

Remark. If $\phi = 0$, we get to the conclusion that our model (Equation 5.4) holds exactly. So setting $\phi = 0$ produces an extremely informative prior which determines the values of the pixels. On the other hand, if $\phi \to$ inf then the prior relaxes to an extremely uninformative distribution which lets the values of the pixels totally free. So in a sense ϕ 'regulates' the freedom allowed to the pixels $u_0, \ldots, u_{v+1}$.

5.4 The likelihood

Apart from our model and prior, we also have n non-overlapping data points, $n \leq v + 2$. These data points can be assigned arbitrarily to any pixel u_i where $i \in \mathbf{c}$, where vector **c** described in Section 5.2. The values of **c** correspond with the location of the observed data regarding the pixel numbers; **c** has N elements. The error of the model at the location of any observed data point is defined as:

$$e_i = u_i - d_i, \qquad \text{and} \qquad i \in \mathbf{c}. \tag{5.16}$$

Assuming that this error has a mean of zero (the error is either positive or negative) with some unknown variance σ_i^2 and using the principle of Maximum Entropy we find that this error (Equation 5.16) has the following probability density function:

$$P(e_i|\sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp[-\frac{1}{2\sigma_i^2} e_i^2]. \tag{5.17}$$

Substituting 5.16 into 5.17 and making a change of variable from the error e_i to the data d_i, the likelihood function can be obtained according to Equation 5.18.

$$P(d_i|u_i, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp[-\frac{1}{2\sigma_i^2}(d_i - u_i)^2], \qquad \text{and} \qquad i \in \mathbf{c}. \tag{5.18}$$

$$
\mathbf{R} \equiv \Delta^2 \times \begin{pmatrix}
\frac{1}{\delta_0^2}+\frac{\delta_{r,1}^2}{\delta_1^2} & \frac{-1}{\delta_0^2}-\frac{\delta_{r,1}}{\delta_1^2} & \frac{\delta_{l,1}.\delta_{r,1}}{\delta_1^2} & 0 & 0 & \cdots & 0 \\
\frac{-1}{\delta_0^2}-\frac{\delta_{r,1}}{\delta_1^2} & \frac{1}{\delta_1^2}+\frac{1}{\delta_0^2}+\frac{\delta_{r,2}^2}{\delta_2^2} & \frac{-\delta_{l,1}}{\delta_1^2}-\frac{\delta_{r,2}}{\delta_2^2} & \frac{\delta_{l,2}.\delta_{r,2}}{\delta_2^2} & 0 & \ddots & \vdots \\
\frac{\delta_{l,1}.\delta_{r,1}}{\delta_1^2} & \frac{-\delta_{l,1}}{\delta_1^2}-\frac{\delta_{r,2}}{\delta_2^2} & \frac{1}{\delta_2^2}+\frac{\delta_{l,1}^2}{\delta_1^2}+\frac{\delta_{r,3}^2}{\delta_3^2} & \frac{-\delta_{l,2}}{\delta_2^2}-\frac{\delta_{r,3}}{\delta_3^2} & \frac{\delta_{l,3}.\delta_{r,3}}{\delta_3^2} & \ddots & \vdots \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & \frac{\delta_{l,v-2}.\delta_{r,v-2}}{\delta_{v-2}^2} & \frac{-\delta_{l,v-2}}{\delta_{v-2}^2}-\frac{\delta_{r,v-1}}{\delta_{v-1}^2} & \frac{1}{\delta_{v-1}^2}+\frac{\delta_{r,v}^2}{\delta_v^2}+\frac{\delta_{l,v-2}^2}{\delta_{v-2}^2} & \frac{-\delta_{l,v-1}}{\delta_{v-1}^2}-\frac{\delta_{r,v}}{\delta_v^2} & \frac{-\delta_{l,v-1}}{\delta_{v-1}^2}-\frac{\delta_{r,v}}{\delta_v^2} \\
\vdots & \ddots & 0 & \frac{\delta_{l,v-1}.\delta_{r,v-1}}{\delta_v^2} & \frac{-\delta_{l,v-1}}{\delta_{v-1}^2}-\frac{\delta_{r,v}}{\delta_v^2} & \frac{1}{\delta_v^2}+\frac{1}{\delta_{v+1}^2}+\frac{\delta_{l,v-1}^2}{\delta_{v-1}^2} & \frac{\delta_{l,v}.\delta_{r,v}}{\delta_v^2} \\
0 & \cdots & 0 & 0 & \frac{\delta_{l,v}.\delta_{r,v}}{\delta_v^2} & \frac{-1}{\delta_{v+1}^2}-\frac{\delta_{l,v}}{\delta_v^2} & \frac{1}{\delta_{v+1}^2}+\frac{\delta_{l,v}^2}{\delta_v^2}
\end{pmatrix}',
$$

where, $\Delta = (\delta_0 + \delta_1 + \cdots + \delta_v + \delta_{v+1})$.

Again by assuming logical independence between the errors and making the appropriate substitutions and changes of variables, the following likelihood function can be obtained.

$$P(d_1,\ldots,d_n|u_1,\ldots,u_n,\sigma_1,\ldots,\sigma_n) \propto \frac{1}{(2\pi)^{n/2}\sigma_1\ldots\sigma_n}\exp[-\sum_{i=0,i\in\mathbf{c}}^{v+1}\frac{1}{2\sigma_i^2}(d_i-u_i)^2]. \tag{5.19}$$

In Equation 5.19, σ_i presents the standard deviation of the error. Assuming that the error of the likelihood is proportional to the error of the prior, it will conclude that

$$\sigma_i = \frac{\delta_i}{\Delta}\times\sigma, \quad 1\leq i\leq v, \tag{5.20}$$

where, $\Delta=\delta_0+\delta_1+\cdots+\delta_v+\delta_{v+1}$, and $\delta_i = x_{i+1}-x_{i-1}$. Therefore, Equation 5.19 can be written as

$$P(d_1,\ldots,d_n|u_1,\ldots,u_n,\sigma) \propto \frac{1}{(2\pi)^{n/2}B\sigma^n}\exp[-\frac{\Delta^2}{2\sigma^2}\sum_{i=0}^{v+1}\frac{(d_i-zu_i)^2}{\delta_i^2}], \tag{5.21}$$

where z is

$$z_i=\begin{cases}1 & \text{if } d_i\neq 0,\\ 0 & elsewhere.\end{cases} \tag{5.22}$$

and B is

$$B=\frac{\delta_1\ldots\delta_n}{(\delta_1+\cdots+\delta_v)^n}.$$

5.5 The posterior

The posterior can be obtained by using the product rule,

$$P(U|D,\sigma,\phi,I)\propto P(U|\phi,I)P(D|\sigma,U). \tag{5.23}$$

As a result, the multiplication of the prior, Equation 5.15, and the likelihood, Equation 5.21, leads to a function which is proportional to the posterior PDF of all the pixels.

$$P(u_0,\ldots,u_{v+1}|D,\sigma,\phi) \propto \frac{1}{(2\pi)^{\frac{n+v+2}{2}}A'B\phi^{v+2}\sigma^n}\exp\left[-\frac{1}{2\phi^2}\sum_{k=0}^{v+1}\sum_{l=0}^{v+1}R_{kl}u_ku_l-\frac{\Delta^2}{2\sigma^2}\sum_{i=0}^{v+1}\frac{(d_i-zu_i)^2}{\delta_i^2}\right]. \tag{5.24}$$

To get the PDF of a target pixel, u_j, Equation 5.24 needs to be integrated over all pixels except the target pixel which results

$$P(u_j|D,\sigma,\epsilon,I)\propto \frac{1}{(2\pi)^{\frac{n+v+2}{2}}A'B\phi^{v+2}\sigma^n} \times\int\exp\left[-\frac{1}{2\phi^2}\sum_{k=0}^{v+1}\sum_{l=0}^{v+1}R_{kl}u_ku_l-\frac{\Delta^2}{2\sigma^2}\sum_{i=0}^{v+1}\frac{(d_i-z_iu_i)^2}{\delta_i^2}\right]\underbrace{du_0\ldots du_{v+1}}_{\text{except}du_j}. \tag{5.25}$$

5.5.1 Regularizer, ϵ

The model error regarding the prior information was described with a parameter ϕ, and the error of the data was described by σ. A parameter ϵ, called the regularizer, can be defined as

$$\epsilon = \frac{\sigma}{\phi}. \tag{5.26}$$

There are two extremes for the value of the regularizer. If the regularizer approaches to a large value, $\epsilon \to \inf$, then the error of the prior information is much smaller than the error of the data. However, if its value approaches to a very small value, $\epsilon \to 0$, the error of the data is much smaller than the error of the model; therefore, the data is more accurate and will dominate the analysis.

As a result of Equation 5.26, for the transformation holds that

$$d\epsilon = \frac{\sigma}{\phi^2} d\phi. \tag{5.27}$$

Taking the advantage of the definition of an Equation 5.25 turns to the following form:

$$\begin{aligned} &P(u_j|D,\sigma,\epsilon,I) \\ &\propto \frac{\epsilon^{v+2}}{(2\pi)^{\frac{n+v+2}{2}} A' B \sigma^{n+v+2}} \\ &\times \int \exp\left[-\frac{\epsilon^2}{2\sigma^2}\sum_{k=0}^{v+1}\sum_{l=0}^{v+1} R_{kl}u_k u_l - \frac{\Delta^2}{2\sigma^2}\sum_{i=0}^{v+1}\frac{(d_i - z u_i)^2}{\delta_i^2}\right] \underbrace{du_0 \dots du_{v+1}}_{\text{except} du_j}. \end{aligned} \tag{5.28}$$

Where constant parameters like π, A' and B can be dropped. Then, Equation 5.28 can be rewritten as

$$\begin{aligned} &P(u_j|D,\sigma,\epsilon,I) \\ &\propto \frac{\epsilon^{v+2}}{\sigma^{n+v+2}} \int \exp\left[-\frac{1}{2\sigma^2}[N\bar{d}^2 + \sum_{k=0}^{v+1}\sum_{l=0}^{v+1} g_{kl}u_k u_l - 2\Delta^2 \sum_{i=0}^{v+1}\frac{d_i z u_i}{\delta_i^2}]\right] \underbrace{du_0 \dots du_{v+1}}_{\text{except} du_j}. \end{aligned} \tag{5.29}$$

where $\bar{d}^2$ is the mean-square data value as

$$\bar{d}^2 \equiv \frac{1}{N}\sum_{i=1}^{N}(\frac{\Delta}{\delta_i})^2 d_i^2. \tag{5.30}$$

and the g_{kl}, called as the integration matrix, is

$$g_{kl} \equiv \epsilon^2 R_{kl} + S_{kl}. \tag{5.31}$$

where the S_{kl} is a diagonal matrix with entry 1 for pixels which data points are assigned to them and 0 everywhere else.

$$S_{kl} = \begin{cases} \frac{\Delta^2}{\delta_k^2} & \text{if } k = l \text{ and data point is assigned to the pixel } u_k, \\ 0 & \text{elsewhere.} \end{cases}$$

For example, if $\mathbf{d} = [d_1, 0, 0, d_2, 0, d_3, 0]^T$ then the S_{kl} is

$$S_{kl} = \begin{pmatrix} \frac{\Delta^2}{\delta_1^2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\Delta^2}{\delta_4^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\Delta^2}{\delta_6^2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since there is no integration over u_j it behaves like a constant. It can, as a result, be separated from the other variables as

$$\begin{aligned} & P(u_j|D,\sigma,\epsilon,I) \\ & \propto \frac{\epsilon^{v+2}}{\sigma^{n+v+2}} \int \exp\left[-\frac{N\bar{d}^2 - 2d_j u_j (\frac{\Delta}{\delta_j})^2 + g_{jj}u_j^2}{2\sigma^2}\right] \\ & \times \exp\left[\sum_{\substack{k=0 \\ k\neq j}}^{v+1}\sum_{\substack{l=0 \\ l\neq j}}^{v+1} g_{kl}u_k u_l - 2\sum_{\substack{i=0 \\ i\neq j}}^{v+1}(d_i z(\tfrac{\Delta}{\delta_i})^2 - g_{jj}u_j)u_i\right] \underbrace{du_0 \ldots du_{v+1}}_{except\ du_j}. \end{aligned} \tag{5.32}$$

The integral may be done by the following change of variables:

$$A_k = \sqrt{\lambda'_k}\sum_{\substack{i=0 \\ i\neq j}}^{v+1} u_i e_{ki} \qquad (k \neq j), \tag{5.33}$$

$$u_k = \sum_{\substack{i=0 \\ i\neq j}}^{v+1} \frac{A_i e_{ik}}{\sqrt{\lambda'_i}} \qquad (k \neq j), \tag{5.34}$$

and λ'_i is the ith eigenvalue of the jth cofactor[1] of the interaction matrix (the g matrix), presented in Equation 5.31, and e_{ik} is the kth component of the ith eigenvalue. Also the integration variables have the property that

$$\sum_{\substack{k=0 \\ k\neq j}}^{v+1} g_{kl}e_{ik} = \lambda'_i e_{il} \tag{5.35}$$

$$\sum_{\substack{k=0 \\ k\neq j}}^{v+1} e_{lk}e_{ik} = \delta_{li} \tag{5.36}$$

[1] The cofactor of a matrix of rank $v+2$ is a matrix of rank $v+1$, and it is formed by deleting the jth row and column of the matrix.

Where δ_{li} is the Kronecker delta function

$$\frac{dA_0 \cdots dA_{j-1} dA_{j-1} \cdots dA_{v+1}}{\sqrt{\lambda'_0 \cdots \lambda'_{j-1} \lambda'_{j+1} \cdots \lambda'_{v+1}}} = du_0 \cdots du_{j-1} du_{j+1} \cdots du_{v+1} \tag{5.37}$$

Making the change of variables and introducing a new quantity $h_l(u_j)$

$$h_l(u_j) = \frac{1}{\sqrt{\lambda'_l}} \sum_{\substack{i=0 \\ i \neq j}}^{v} [d_i z (\frac{\Delta}{\delta_i})^2 - g_{ij} u_j] e_{li} \quad (l \neq j) \tag{5.38}$$

one obtains

$$\begin{aligned} & P(u_j|D, \sigma, \epsilon, I) \\ & \propto \frac{\epsilon^{v+2}}{\sigma^{n+v+2}} \exp\left[-\frac{N\bar{d}^2 - 2d_j u_j (\frac{\Delta}{\delta_j})^2 z + g_{jj} u_j^2 - h(u_j) \cdot h(u_j)}{2\sigma^2} \right] \\ & \times \int \exp\left[-\frac{1}{2\sigma^2} \sum_{\substack{i=0 \\ i \neq j}}^{v+1} (A_i - h_i(u_j))^2 \right]. \end{aligned} \tag{5.39}$$

where

$$h_l(u_j) \cdot h_l(u_j) = \sum_{\substack{i=0 \\ i \neq j}}^{v+1} h_l(u_j)^2. \tag{5.40}$$

Evaluating the $v+1$ integrals gives a factor of σ^{v+1}, and one obtains

$$P(u_j|D, \sigma, \epsilon, I) \propto \frac{\epsilon^{v+2}}{\sigma^{n+1}} \exp\left[-\frac{N\bar{d}^2 - 2d_j u_j (\frac{\Delta}{\delta_j})^2 z + g_{jj} u_j^2 - h(u_j) \cdot h(u_j)}{2\sigma^2} \right]. \tag{5.41}$$

Which is, in fact, equal to

$$P(u_j|D, \sigma, \epsilon, I) \propto \exp\left[\frac{2d_j u_j (\frac{\Delta}{\delta_j})^2 z - g_{jj} u_j^2 + h(u_j) \cdot h(u_j)}{2\sigma^2} \right]. \tag{5.42}$$

5.6 Eliminating σ

Given the fact that most of the times neither value of σ nor ϵ are known, they are treated as nuisance parameters. They need to be integrated out in order to get a posterior which is not dependent on them

$$P(u_j|D, \epsilon, I) = \int P(u_j, \sigma|D, \epsilon, I)\, d\sigma. \tag{5.43}$$

The integrand may be factored to obtain

$$\begin{aligned} P(u_j,\sigma|D,\epsilon,I) &= P(u_j,\sigma|\epsilon,I)\, P(D|u_j,\epsilon,I) \\ &= P(u_j|I)\, P(\sigma|I)\, P(D|u_j,\sigma,\epsilon,I) \\ &= P(\sigma|I)\, P(u_j|D,\sigma,\epsilon,I)\,. \end{aligned} \tag{5.44}$$

Assuming that the prior probability, $P(u_j,\sigma|\epsilon,I)$, is independent of the ϵ, and that the $P(u_j,\sigma|I) = P(u_j|I)\cdot P(\sigma|I)$ presents Equation 5.43 in the form of

$$P(u_j|D,\epsilon,I) = \int P(\sigma|I)\, P(u_j|D,\sigma,\epsilon,I)\; d\sigma. \tag{5.45}$$

$P(u_j|D,\epsilon,I)$ is the prior probability for the variance, and the second part of the equation above is proportional to Equation 5.25. Having no assumption over the σ, a Jeffreys prior[2] $\frac{1}{\sigma}$ is assigned to obtain

$$\begin{aligned} P(u_j|D,\epsilon,I) &= \int_0^\infty d\sigma P(u_j|D,\epsilon,I) \\ &= \int_0^\infty d\sigma\sigma^{-(N+1)} \exp\left[-\frac{N\bar{d}^2 - 2d_j u_j(\frac{\Delta}{\delta_j})^2 z + g_{jj}u_j^2 - h(u_j)\cdot h(u_j)}{2\sigma^2}\right]. \end{aligned} \tag{5.46}$$

This leads to

$$P(u_j|D,\epsilon,I) = \left[1 - \frac{2d_j u_j(\frac{\Delta}{\delta_j})^2 z - g_{jj}u_j^2 + h(u_j)\cdot h(u_j)}{N\bar{d}^2}\right]^{-\frac{N}{2}}. \tag{5.47}$$

5.6.1 Estimation of the regularizer ϵ

We would like to integrate out the parameter ϵ and get a closed form solution for it. However, the formula is very nonlinear. Therefore, an attempt is made to present the best estimation of this parameter. The effect of the value of ϵ is discussed in Section 5.5.1, so it is important to assign a reasonable value to it. A good estimate of the ϵ, in fact, leads to a reasonable result. Otherwise, the effect of prior or likelihood on the posterior is neglected. Using the sum rule leads to

$$P(\epsilon|D,I) = \int du_0 \cdots du_{v+1} d\sigma\, P(\epsilon,\sigma,u_0,\cdots,u_{v+1}|D,I). \tag{5.48}$$

The integrand can be factored and written as

$$P(\epsilon|D,I) = \int du_0 \cdots du_{v+1} d\sigma\, P(\epsilon,\sigma|I) P(u_0,\cdots,u_{v+1}|D,I,\epsilon,\sigma). \tag{5.49}$$

[2] It is an uninformative prior.

$P(\epsilon,\sigma|I)$ is the joint probability density function for ϵ and σ. Besides,

$$P(u_0,\cdots,u_{v+1}|D,I,\epsilon,\sigma)$$

can be written as the multiplication of the prior and likelihood, respectively presented in Equation 5.15 and 5.21.

$$\begin{aligned} P(\epsilon|D,I) = &\int du_0\cdots du_{v+1}d\sigma P(\epsilon|I)P(\sigma|I) \\ &\times P(u_0,\cdots,u_{v+1}|I,\epsilon,\sigma)P(D|u_0,\cdots,u_{v+1},I,\epsilon,\sigma). \end{aligned} \tag{5.50}$$

Substitution of the prior and likelihood in to the Equation 5.50, and working out the formula in the same process that was done for the posterior leads to

$$\begin{aligned} P(\epsilon|D,I) \propto &\int du_0\cdots du_{v+1}[\lambda_0\cdots\lambda_{v1}]^{\frac{1}{2}}\sigma^{-(v+N+3)}\epsilon^{v+1} \\ &\times \exp\left[-\frac{\epsilon^2}{2\sigma^2}\sum_{k=0}^{v+1}\sum_{l=0}^{v+1}R_{kl}u_ku_l\right] \\ &\times \exp\left[-\frac{\Delta^2}{2\sigma^2}\sum_{k=0}^{v+1}\frac{[d_i-zu_i]^2}{\delta_i^2}\right]. \end{aligned} \tag{5.51}$$

The ϵ is implemented into R matrix. Therefore, its eigenvalues can not be considered as a constant, and they need to be included in the equation. Equation 5.51 is extended in the same process that Equation 5.25 was extended to obtain

$$\begin{aligned} P(\epsilon|D,I) \propto &\int du_0\cdots du_{v+1}[\lambda_0\cdots\lambda_{v1}]^{\frac{1}{2}}\sigma^{-(v+N+3)}\epsilon^{v+1} \\ &\times \exp\left[-\frac{1}{2\sigma^2}\left[N\bar{d}^2-2\sum_{i=0}^{v+1}d_iu_iz(\frac{\Delta^2}{\delta_i^2})+\sum_{k=0}^{v+1}\sum_{l=0}^{v+1}g_{kl}u_ku_l\right]\right]. \end{aligned} \tag{5.52}$$

After integrating of integrals, the posterior probability for ϵ which is independent of the pixel values is given by

$$\begin{aligned} P(\epsilon|D,I) = &\int d\sigma\left(\frac{\lambda_0\cdots\lambda_{v1}}{\lambda'_0\cdots\lambda'_{v1}}\right)^{\frac{1}{2}}\sigma^{-(v+N+3)}\epsilon^{v+1} \\ &\times \exp\left[-\frac{N\bar{d}^2-h(\epsilon)\cdot h(\epsilon)}{2\sigma^2}\right]. \end{aligned} \tag{5.53}$$

Where

$$h_l(\epsilon) \equiv \frac{1}{\sqrt{\lambda'_l}}\sum_{\substack{i=0\\ i\neq j}}^{v+1}d_i(\frac{\Delta}{\delta_i})^2ze_{li}, \tag{5.54}$$

$$h_l(\epsilon)\cdot h_l(\epsilon) \equiv \sum_{i=0}^{v+1}h_i(\epsilon)^2, \tag{5.55}$$

$\{\lambda_0\cdots\lambda_{v+1}\}$ are the eigenvalues of the R_{kl} matrix, which was introduced in Equation 5.16 and $\{\lambda'_0\cdots\lambda'_{v+1}\}$ and e_{li} are the eignevalues and eigenvectors of the interaction matrix, g_{kl}, introduced in Equation 5.31. As a result of the same process which was done on Equation 5.47, one obtains

$$P(\epsilon|D,I) \propto \left(\frac{\lambda_0\cdots\lambda_{v1}}{\lambda'_0\cdots\lambda'_{v1}}\right)^{\frac{1}{2}}\epsilon^{v+1}\left[1-\frac{h(\epsilon)\cdot h(\epsilon)}{N\bar{d}^2}\right]^{-\frac{N}{2}}. \tag{5.56}$$

5.7 Algorithm

Here the algorithm is described for the Bayesian Monte Carlo (BMC). To begin, let us consider the ideas of the Monte Carlo algorithm by looking at a one dimensional example. Consider a one-dimensional limit state equation $G(x) = 0$ with the probability density function $f_X(x)$ of the variable X.

The first and second random points, $\vec{x}^{(1)} = x^{(1)}$ and $\vec{x}^{(2)} = x^{(2)}$, are generated from the PDF f_X. Without any loss of the generality, we assume that $x^{(1)} < x^{(2)}$. This comes from the sorting requirement. Then, we assign pixels u_1 and u_2 to their positions. The limit state equation (LSE) needs to be evaluated for these pixels (positions); the calculated value of these pixels are called data points named as d_1 and d_2. Then, the third random point, named as $x^{(3)}$, is generated. This newly generated random number is called a target pixel[3].

This point, or the target pixel, may have three positions which are $x^{(3)} < x^{(1)}$, $x^{(1)} < x^{(3)} < x^{(2)}$ or $x^{(3)} > x^{(2)}$. Based upon the position of the target pixel among the other pixels, we rename all the pixels and data points. Again without any loss of the generality we assume the target pixel is located between u_1 and u_2. Therefore the target pixel is now named as u_2, while u_1 is still u_1 and the previous u_2 (which was related to $x^{(2)}$) is named as u_3. Therefore, the related data points will be respectively named as d_1 and d_3. Then, according to the presented formula in Equation 5.56 we need to estimate the best value of the regularizer based on our information over the model and data. Given the value of the regularizer, ϵ, we can estimate the PDF of the the target pixel, $u_j = u_2$ by Equation 5.42. In this stage of the process, a Cauchy type of the PDF is expected in which the mean value of the target pixel can be estimated but the second moment is infinity. Assuming that the estimated accuracy for the target pixel is not enough, we proceed to further simulations in the same process.

The reader should keep in mind that in practice the location of the LSE-curve is not known; it only gradually becomes visible as it is interpolate or extrapolate based upon the prior and data information and obtains the increasingly more accurate approximations to the value of the target pixel. The corresponding probability density function (PDF) becomes tighter as the number of generated points increases. Also note that for some generated points it is not necessary to evaluate the limit state function, because from their PDF it can be determined whether the point is stable or unstable.

Summarizing, the algorithm is as follows:

1. In the Monte Carlo process, a random number is generated from the probability density function (PDF) of the variable X, f_X, and according to its value a pixel is defined, this pixel is called the target pixel.

2. If there are neighbors for the target pixel, we estimate the value of the target

[3]The target pixel is presented as u_j in this chapter.

pixel; otherwise, we back to the previous step to get more neighbors which data point has been assigned to them.

3. Given the neighbor pixels and the assumed model, the PDF of the current pixel (we assume u_j) is calculated by Equation 5.58.

4. Having the PDF of u_j and according to the accepted tolerance criterion, it is decided whether there is a need to calculate the limit state equation for the jth point or the accuracy is enough. In this case, if the confidence interval is narrow enough, we estimate its value and proceed our simulation by coming back to the first step. Otherwise, we calculate the value of limit state equation (LSE) for this point.

   Repeat until the Monte Carlo requirements are obtained.

5. $\hat{p}_f = n_f/N$ is a simple Monte Carlo estimate for p_f.

This estimate $\hat{p}_f$ is as good as an ordinary Monte Carlo estimate based on N independent samples $\mathbf{1}[G(\vec{X}^{(1)}) < 0],\ldots,\mathbf{1}[G(\vec{X}^{(N)}) < 0]$, but requires evaluation of G in only a fraction of the samples.

5.8 Numerical example

Here, we refer again to the break water model presented in Section 3.4 and Equation 3.16. The one dimensional limit state equation (LSE) of this model can be defined by Equation 5.57, where the velocity parameter is assumed to be normally distributed as $N(1.5, 0.45)$.

$$G(u) = 5 - 0.98280 \times u^2. \tag{5.57}$$

Without of loosing generality of the problem, we assume that there are two data points are generated from the normal distribution function of $N(1.5, 0.45)$ where the mean value of the PDF is 1.5 and its standard deviation is 0.45. These data points are $u_1 = 0.15$ and $u_3 = 2.85$. The values of limit state equation for these two points are $G(u_1) = 4.978$ and $G(u_3) = -2.983$. We assume that the randomly generated pixel or the target pixel, u_j is in $u_2 = 2.25$. This situation, is graphically presented in Figure 5.3.

Before we continue the simulation process, we would like to present the probable values of u_j with the suggested model. Therefore, we need to use Equation 5.58 in order to get the required PDF. Nevertheless, this equation contains σ and ϕ. As a matter of fact, σ can be integrated out of the equation, but we need to estimate a value for the ϕ. In this case, we define $\epsilon = \frac{1}{\phi}$ which is called regularizer. Then we can get the PDF of our regularizer to find its optimal value which leads to the most narrow PDF. The most probable value of ϵ is estimated to be

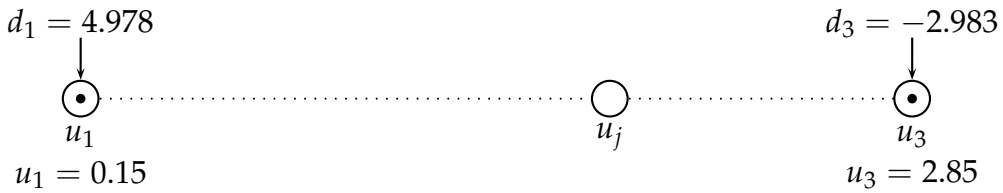


Figure 5.3: An illustration of the pixels which data points are assigned to.

2.6 as suggested by Rajabalinejad et al. (2008a) and we use this value during the rest of this work. As a result, Equation 5.58 presents the PDF of the target pixel, $u_j = 2.25$ given two data points: d_1 and d_3 as shown in Rajabalinejad et al. (2008a).

$$P(u_j|d_1,d_3) \propto \frac{0.3126\,10^9}{0.5897\,10^{10}+0.5265\,10^{10}\,u_j+0.1339\,10^{10}\,u_j{}^2}$$

The PDF of u_j given d_1 and d_3 is depicted in Figure 5.4. This figure is, in fact, a plot of Equation 5.58. The mean value or the most probable value of this PDF is -1.97. Besides, the 95% confidence interval by assuming a symmetrical distribution leads to the confidence interval of [-11.28, 7.35]. This interval is obtained by solving the equation which defines the integration of a symmetrical area around mean value should be equal to 0.95. As presented in the figure, it is a wide PDF and its tails are more informative than the Gaussian.

5.8.1 Comparison between linear interpolation and Bayesian interpolation

To clarify the advantage of the Bayesian Monte Carlo comparing with the linear interpolation, we consider two data points. Given just two data points in a linear interpolation, there is no other way than assuming a linear relationship which leads toward the value of -1.21 for the target pixel while there is no estimation of the uncertainty. But the Bayesian Monte Carlo technique enables us to get a criterion for the uncertainty of the estimated value of each pixel as depicted in Figure 5.4. Now, the distinction between two methods is obvious; a judgment tool is obtained by the Bayesian Monte Carlo.

This comparison is illustrated in Figure 5.5 for the LSE presented in Section 5.8. In this figure there are two data points called **A** and **B**. These two points are the only information which provide point **e** using a linear interpolation for the target pixel (u=2.25), where **e**=-1.21. This is not close to real value of the limit

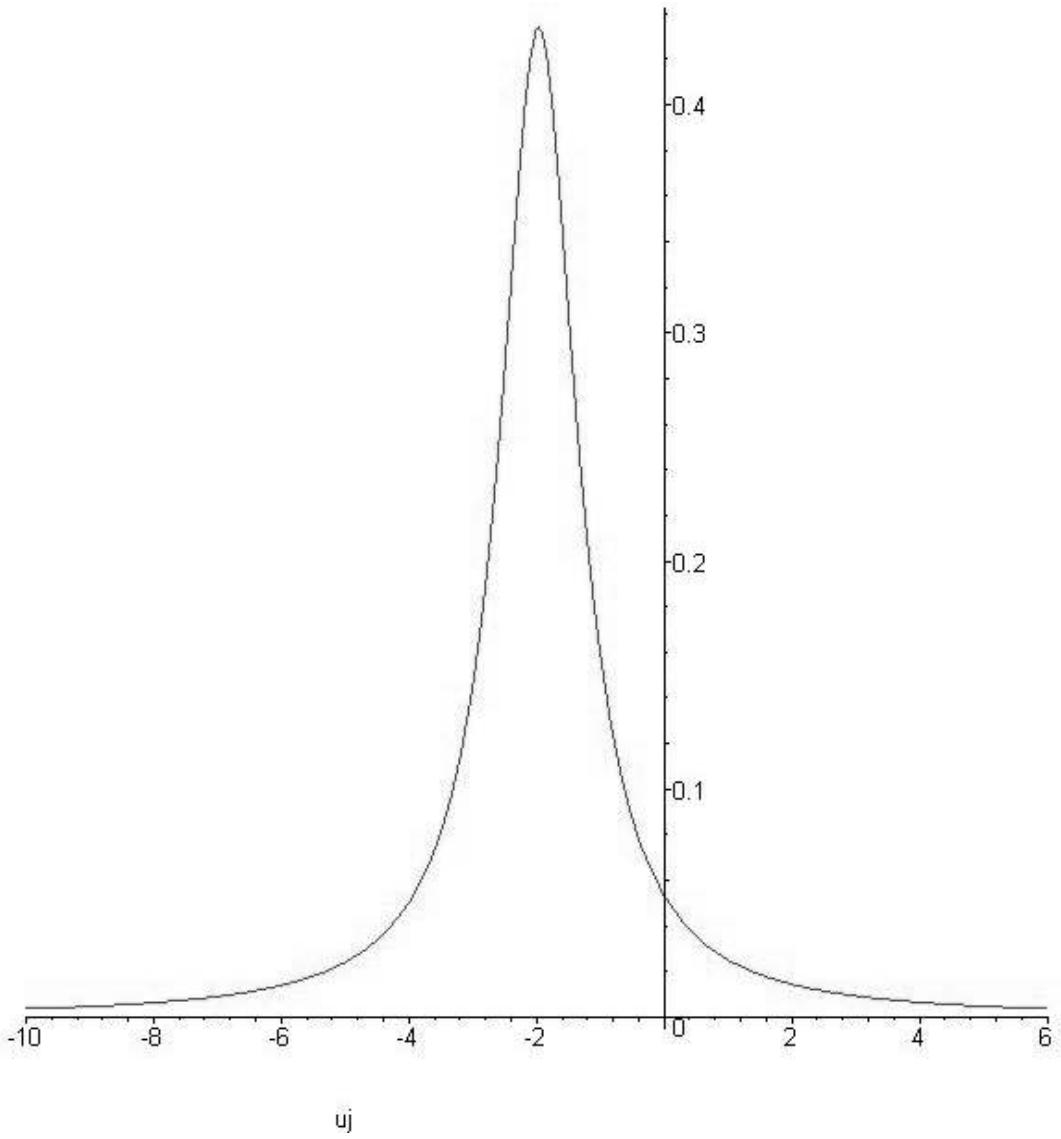


Figure 5.4: This figure presents the probability density function (PDF) of the target pixel (u_j) where $u = 2.25$ given 2 measured data points: d_1 and d_3.

state function **g**=0.0246. Nevertheless, there is no information over the certainty of the estimated point **e** from the interpolation. On the other hand, point **f** is the mean value of the PDF calculated by the Bayesian technique (**f**=-1.97). The uncertainty is shown by its PDF. Having a look at the presented PDF, a rather wide PDF can be seen; and both of positive and negative values are expected for this pixel.

The difference of the results of linear and Bayesian interpolation at this case is because of the value of the regularizer(ϵ). The effect of epsilon (or ϕ which is inversely related to it) was previously described. In fact, we can have two extreme situations when we consider two extreme values for Φ. These extreme values are 0 and infinity. In the first case we just stick to our data values and in the second case we just consider our model assumption and leave the data information. Therefore, the difference between **e** and **f** should be related to the value of regularizer.

5.8.2 Change of the PDF in a certain pixel

From now on we start the Monte Carlo simulation by generating random numbers from N(1.5,0.45). However, before we run the limit state equation for each random number which is assigned to a new target pixel u_j, we check if it is necessary or its value can be assigned without any realization of the LSE regarding the accepted tolerance. To investigate the change of the accuracy during the simulation, we monitor a certain random pixel to clarify the change of its PDF during the calculations. This helps to clearly understand the concept of Bayesian Monte Carlo. The pixel which is correlated to the $u = 1.575$ is considered as a target pixel and is monitored during the simulation in order to observe the change of its PDF given different number of data points. This, however, is not fully in correspondence with the suggested algorithm in Section 5.7, and it is presented for the clarification purpose. The value of the regularizer is also considered to be $\epsilon = 2.6$ as it is suggested by Rajabalinejad et al. (2008a).

P(u_j|*4* data points)

After *4* realizations of the LSE (or getting *4* random data points) which are assigned to their location, the calculated PDF of the target pixel, u_j, for the location $u = 1.575$ is calculated as

$$P(u_j|d_1,d_2,d_3,d_5) \propto \frac{1.010906880\times10^{122}}{\left(2.187557533\times10^{60}-1.266735392\times10^{60}\ uj+4.165036151\times10^{59}\ uj^2\right)^2}, \tag{5.58}$$

where the random pixels and their correlated data points are presented in Table 5.1.

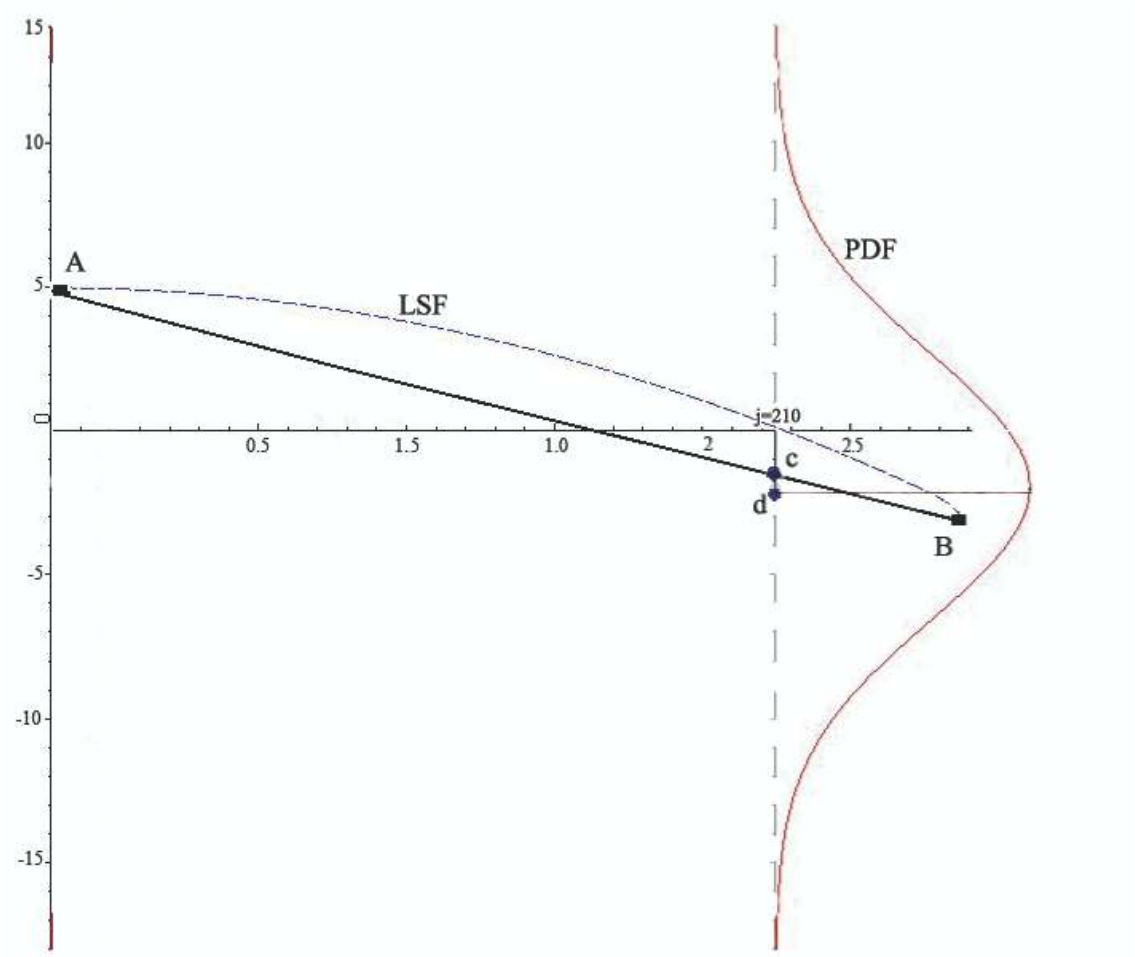


Figure 5.5: A comparison between the linear interpolation and the Bayesian Monte Carlo for the target pixel ($u = 2.25$), given 2 measured data points: d_1 and d_3. The exact value of the function (Equation 5.57) is depicted by dashed line.

pixel positions		data points	
u_1	0.1500	d_1	4.9779
u_2	0.7505	d_2	4.4465
u_3	1.5564	d_3	2.6193
$u_j = u_4$	1.5750		
u_5	2.2500	d_5	0.0246

Table 5.1: Given *4* random data points, the PDF of $u_j = u_4$ is presented in Equation 5.58.

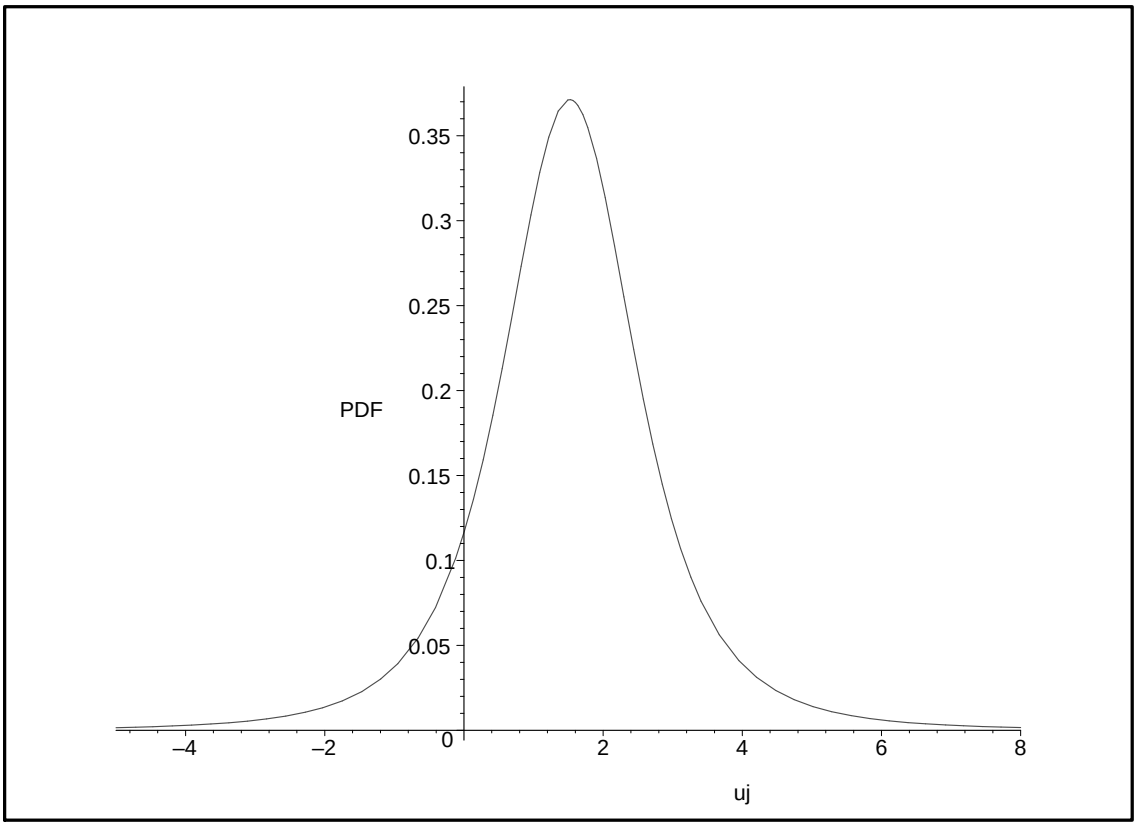


Figure 5.6: This figure presents the probability density function (PDF) of the target pixel u_j given *4* measured data points given in random locations.

Equation 5.58 is also depicted in Figure 5.6. The mean value of this PDF is 1.5207 , and the 95% accuracy by assuming a symmetrical distribution leads to the interval of [-1.68 4.72].

P(u_j|*10* data points)

Since we are not satisfied with the calculated accuracy by *4* data points, we continue to generate more random data points. Figure 5.7 presents the PDF of the target pixel u_j having 10 calculated data points. The randomly generated pixels and the correlated data points are presented in Table 5.2. The PDF is presented in Equation 5.59. The mean value of this PDF is 2.5637, and the 95% accuracy by assuming a symmetrical distribution leads to the values in the interval of [2.413, 2.713]. This shows that by implementing more data points, a narrower PDF is obtained. Since this interval is small enough, we might assume that we have got

pixel positions		data points	
u_1	0.1500	d_1	4.9779
u_2	0.7505	d_2	4.4465
u_3	0.9841	d_3	4.0482
u_4	1.4831	d_4	2.8383
u_5	1.5564	d_5	2.6193
$u_j = u_6$	1.5750		
u_7	1.5786	d_7	2.5509
u_8	1.6473	d_8	2.3331
u_9	2.0351	d_9	0.9295
u_{10}	2.0359	d_{10}	0.9264
u_{11}	2.2500	d_{11}	0.0246

Table 5.2: Given *10* random data points, the PDF of $u_j = u_6$ is presented in Equation 5.59.

the enough accuracy.

$$P(u_j|d_1,\ldots,d_5,d_7,\ldots,d_{11}) \propto \frac{2.517788410\times10^{302}}{\left(2.540061290\times10^{60}-1.970352850\times10^{60}\,uj+3.842724004\times10^{59}\,uj^2\right)^5}. \tag{5.59}$$

where the random pixels and correlated data points are

P(u_j|more data points)

In the simulation process, we may continue calculations until we obtain the required accuracy. In the process of Bayesian Monte Carlo, given more data points is equal to a higher accuracy. This is shown in Table 5.3 in which the calculated mean value and the 95% confidence interval are shown given different number of data points. This table shows that from 20 data points, the mean value with two digits accuracy does not change. The 95% confidence interval, however, gets to the length of 0.015 if 200 data points are given.

The mean values and confidence intervals are obtained from Equation 5.42 and in the same form as presented in Equation 5.58 and Equation 5.59. The calculated PDFs given different data points are

$$P(u_j|20 \text{ data points}) \propto \frac{1.796652634\times10^{618}}{\left(6.063151530\times10^{60}-4.697890134\times10^{60}\,uj+9.163607020\times10^{59}\,uj^2\right)^{10}}. \tag{5.60}$$

$$P(u_j|50 \text{ data points}) \propto \frac{6.259047446\times10^{1546}}{\left(7.086530901\times10^{59}-5.525452842\times10^{59}\,uj+1.078255115\times10^{59}\,uj^2\right)^{25}}. \tag{5.61}$$

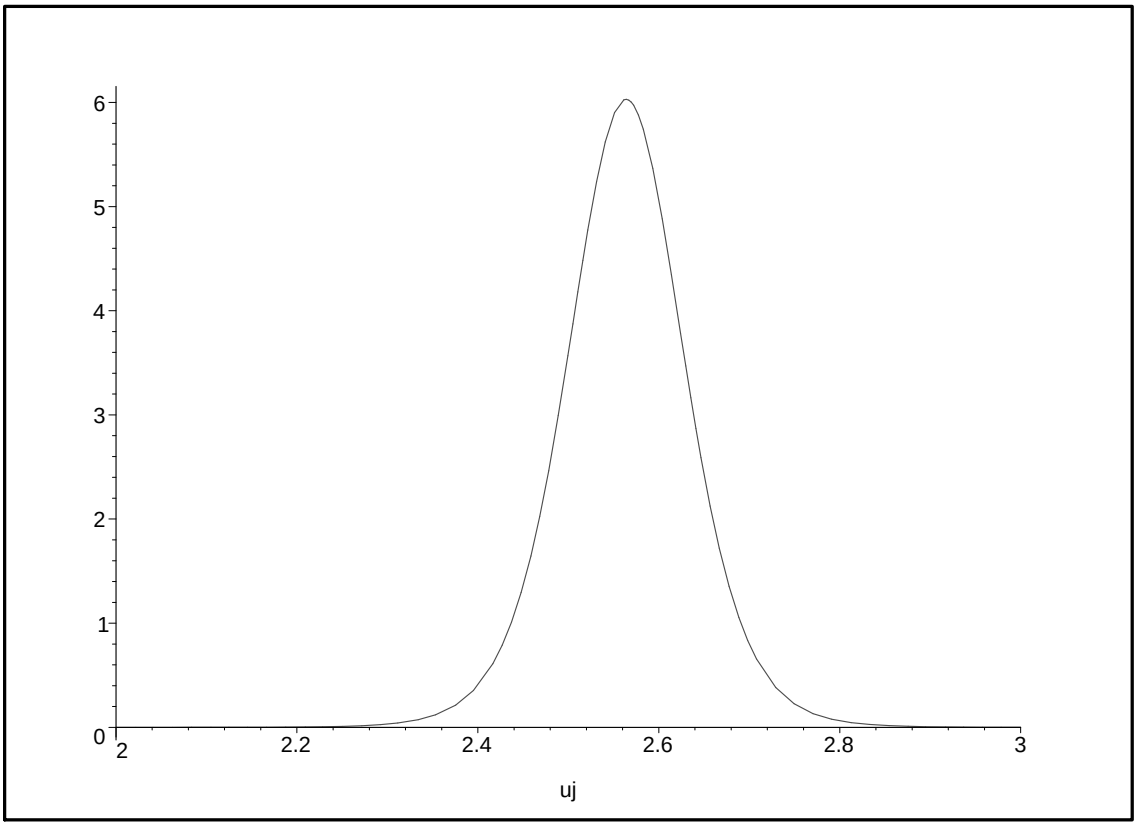


Figure 5.7: This figure presents the probability density function (PDF) of the u_j given 10 random data points.

Data points	mean value	95% interval	
20	2.5633	2.4583	2.6683
50	2.5622	2.5375	2.5869
80	2.5621	2.5420	2.5822
120	2.5620	2.5473	2.5767
160	2.5620	2.5495	2.5745
200	2.5620	2.5541	2.5699

Table 5.3: A comparison between the calculated accuracy and 95% confidence interval for a certain target pixel (u=1.575 in Equation 5.57), given different number of data points.

$$P(u_j|80 \text{ data points}) \propto \frac{8.432112281\times10^{2487}}{\left(8.584336472\times10^{59}-6.692751222\times10^{59}\,uj+1.306095943\times10^{59}\,uj^2\right)^{40}}. \tag{5.62}$$

$$P(u_j|120 \text{ data points}) \propto \frac{7.302223585\times10^{3801}}{\left(2.111872960\times10^{60}-1.647021045\times10^{60}\,uj+3.214289667\times10^{59}\,uj^2\right)^{60}}. \tag{5.63}$$

$$P(u_j|160 \text{ data points}) \propto \frac{1.401727467\times10^{5058}}{\left(1.111650541\times10^{60}-8.670298560\times10^{59}\,uj+1.692092777\times10^{59}\,uj^2\right)^{80}}. \tag{5.64}$$

$$P(u_j|200 \text{ data points}) \propto \frac{3.917096557\times10^{6335}}{\left(9.942829468\times10^{59}-7.758149247\times10^{59}\,uj+1.514077898\times10^{59}\,uj^2\right)^{100}}. \tag{5.65}$$

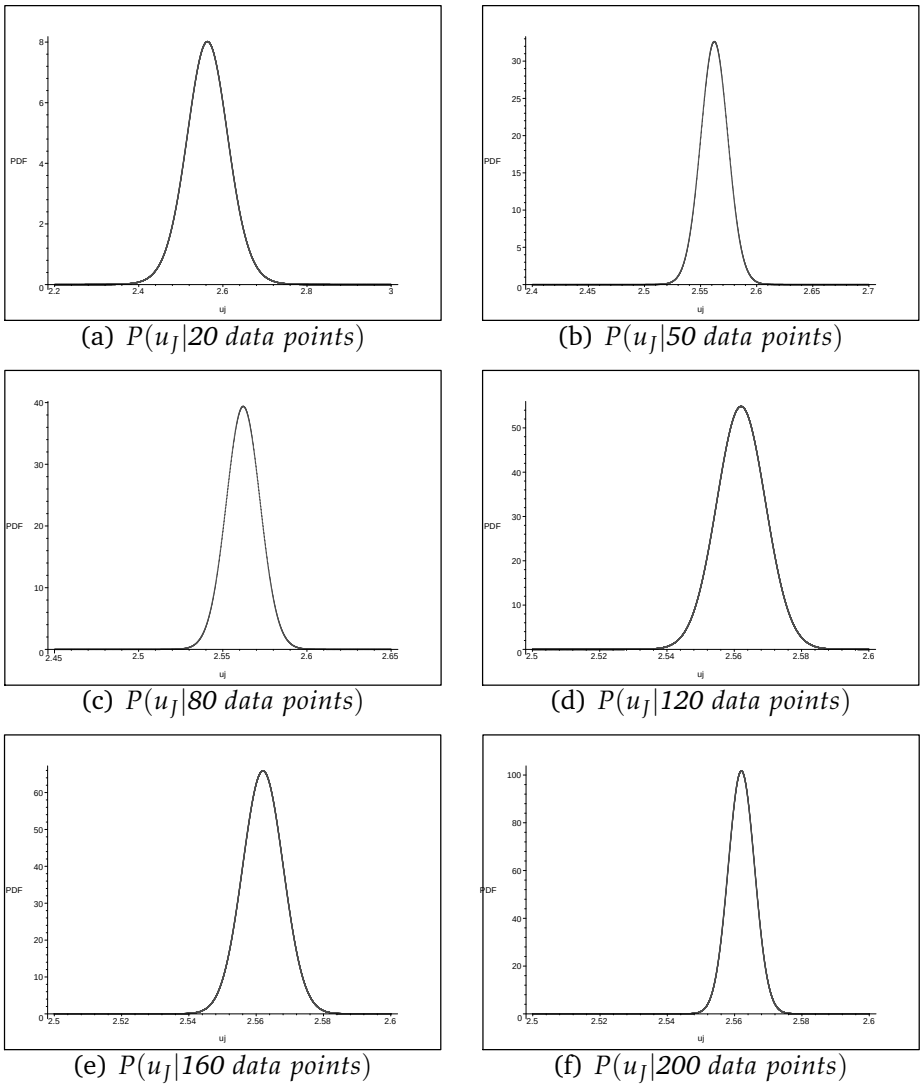
Figure 5.8 presents the probability density functions for different data points as presented in Equations 5.60 to 5.65.

5.9 Bayesian Monte Carlo

In fact, the number of simulations in the Monte Carlo technique depends on several factors, and the tolerance is the most important one. In order to get a required accuracy, random data points are implemented into the model. As a result, a higher number of random pixels lead to a higher accuracy. In other words, to get more precise results, higher number of data points should be implemented into the model.

It is useful to compare the calculated PDFs for a certain pixel during the simulation process as was obtained in Section 5.8.2. In this case, Figure 5.9 provides a comparison among the different PDFs with a fixed scale. Figure 5.9(a) presents the PDF of the target pixel when there are just *4* data points presented. Figure 5.9(b) presents the PDF of the same target pixel when there are 10 data points randomly generated and assigned to the related pixels. Figure 5.9(c) again presents the PDF of the same pixel where the information of 20 pixels are implemented. As a result of further data points, a narrower PDF is obtained till Figure 5.9(h) in which 200 data points are implemented. It is obvious that as far as the simulation is in progress and more data points are implemented, more precise estimation of a target pixel is possible.

Therefore, the presented example clarifies the whole process of the Bayesian Monte Carlo technique. In this study, an attempt is made to get the prior information of the model incorporated to the current level of the analysis. This is a step forward in Monte Carlo simulations. In fact, the **R** in Equation 5.15 provides a link between the information of each pixel and its neighborhood. In other words, information of each point passes through this link and effect the others. Besides, this approach provides a nice tool to get the other priors incorporated to

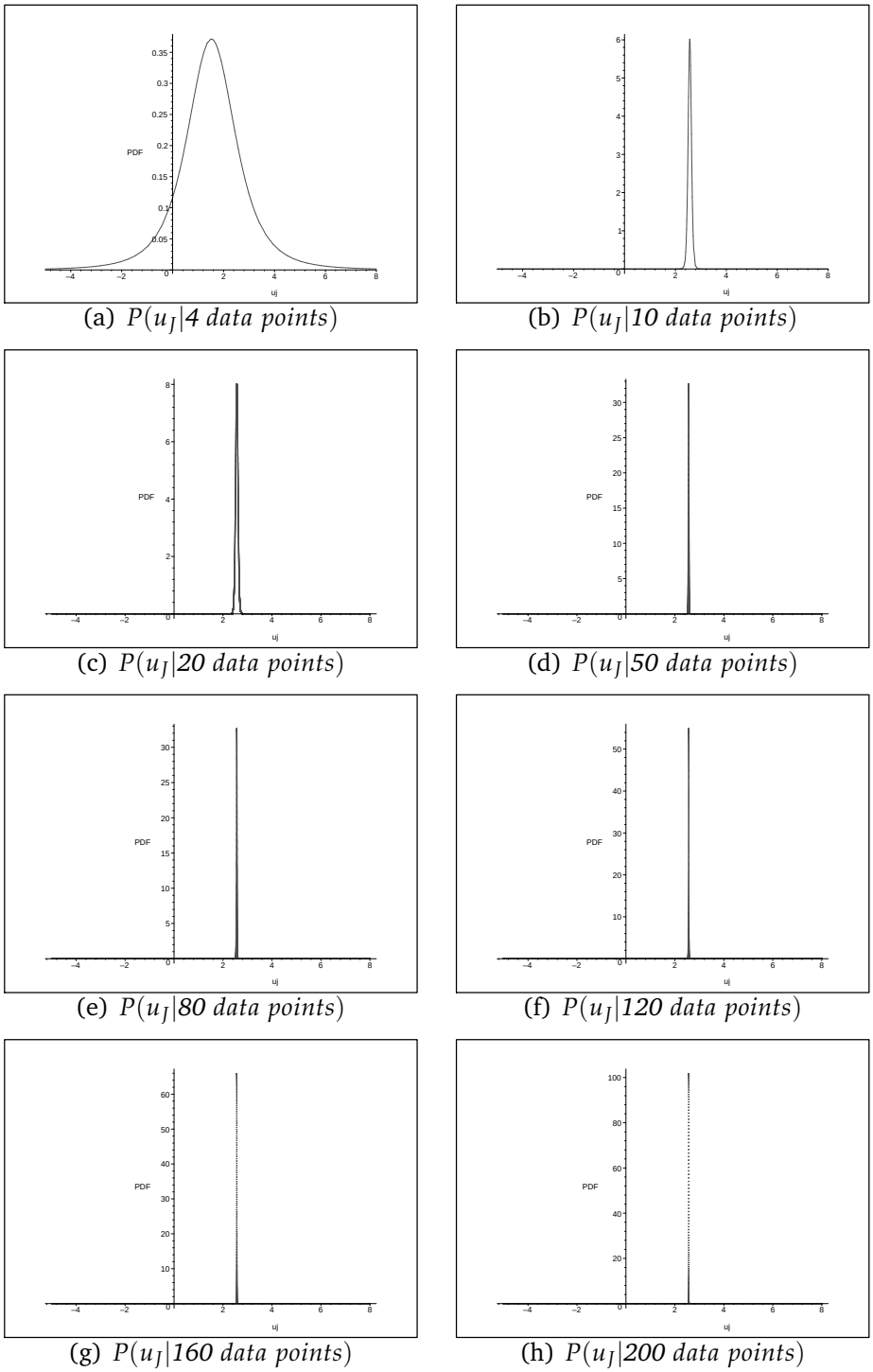


(a) $P(u_J|20\ data\ points)$ (b) $P(u_J|50\ data\ points)$

(c) $P(u_J|80\ data\ points)$ (d) $P(u_J|120\ data\ points)$

(e) $P(u_J|160\ data\ points)$ (f) $P(u_J|200\ data\ points)$

Figure 5.8: This figure shows the probability density function which is assigned to the target pixel (u_j) given different number of data points. In Figure(a) Just the information of 20 data points are considered while in figure b, c, d, e and f the information of 20, 50, 80, 120, 160 and 200 data points are considered, respectively.



(a) $P(u_J|4\ data\ points)$ (b) $P(u_J|10\ data\ points)$

(c) $P(u_J|20\ data\ points)$ (d) $P(u_J|50\ data\ points)$

(e) $P(u_J|80\ data\ points)$ (f) $P(u_J|120\ data\ points)$

(g) $P(u_J|160\ data\ points)$ (h) $P(u_J|200\ data\ points)$

Figure 5.9: This figure shows the probability density function which is assigned to the pixel $u_j = 1.575$ given different number of data points. In Figure(a) Just the information of 4 data points are considered while in figure b, c, d, e, f, g and h, the information of 10, 20, 50, 80, 120, 160 and 200 data points are considered, respectively.

the Monte Carlo simulations. For instance, the dynamic bounds method 3 assumes some other prior information which can be implemented in this approach.

5.9.1 The reduced number of simulations

The reduced number of simulations using the Bayesian Monte Carlo technique depends on a few parameters. It is more importantly the tolerance which dictates the minimum number of the required data points. However, it is important in the Monte carlo process to have a judgment if the model fails or not fails. If we just consider the failure or not failure for a certain target pixel, presented in the example, just *10* realizations are enough given the fact that 95% interval lies in the positive axis as presented in Section 5.8.2. If we consider further parameters, however, we need more data points. For instance, for getting a two fixed digits of the mean value, the required data points for a random target pixel is just *20* realizations of the limit state equation according to Table 5.3. If the minimum length of 95% interval should be less than 0.05, then still we need to take into account *50* data points as presented in Table 5.3. This interval still can be reduced to the length of 0.015 given *200* data points. Therefore, the real number of the realization of the LSE is depended on the tolerance and the selected approach for the reliability analysis.

5.10 The Matrix form

The matrix form of the prior and likelihood are presented in the appendixes. Appendix *A* presents the derivation of the prior in the matrix form. Appendix *B* shows the same process for the likelihood. The outcome of an approach in the matrix form provides an easier formulation of the whole process.

5.11 Conclusion

The suggested procedure can speed up the Monte Carlo simulations integrated with finite elements or the other highly complicated and time consuming processes. The proposed method also provides a tool for implementing the informative priors of a model. The extension of this work with a higher dimensionality of a limit state equation is recommended for future research.

6

A case study: 17th Street Flood Wall, New Orleans

Probabilistic approaches provide a better understanding of the failure mechanisms and assessment of the occurrence probabilities in engineering problems. Besides, it encourages us to take into account the consequences of failure in design processes. Nevertheless, to get its full advantages a well defined model of the structure is necessary. Probabilistic finite elements offer the complementary part, but the outcome is a time consuming process. To improve its efficiency the method of Dynamic Bounds (DB) can be applied. This method speeds up the simulation process through the storage of dynamic bounds, which are continually updated. This means that the produced bounds can be kept and used for next series of simulations. However, this method is efficient when the number of variables are limited which mainly happens in geotechnical structures and flood defences. The DB method is applied to the 17th Street flood wall which was a part of the failed flood defence of New Orleans in hurricane Katrina. The variation of soil parameters are assumed to estimate the reliability of this structure; this assumption will infect the calculated probability of failure.

6.1 Introduction

Flood defences, which are mainly geotechnical structures, protect people from floods in vulnerable areas and their failures usually result in numerous casualties as well as a large economical damage. Besides, global warming and increasing of the normal sea levels, leading to higher storm surges, increase the risk of flooding for many cities and with that the importance of flood defences. Therefore, their role in safety is quite obvious.

For instance, failure of the flood defences in New Orleans during Hurricane Ka-

trina caused a lot of casualties and economical damage in New Orleans. This city was protected by a system of levees, flood walls, and barriers. However, as Link et al. (2006) states 'the system did not perform as a system: the hurricane protection in New Orleans and Southeast Louisiana was a system in name only'. In fact, the flood protection systems are an example of a series system, and it is important that all of its components and elements are safe enough. In other words, the reliability of each element of a flood system should be accurately assessed. Therefore, an accurate safety assessment of the flood defences has a high priority, and it is investigated by many research projects like as www.floodsite.net (2008). In this case, two main issues should be addressed to get the optimal results: an accurate model and a suitable reliability method. Probabilistic finite elements (see Section 6.4) provide a tool by which we obtain a better understanding of the behavior of flood defences and a broader spectrum of its physical behavior based on the engineering properties of the components, systems, and parameter variations. However, this approach is not cheap, and many analyses should be done to find different aspects of model behavior and the role of different variables. Here, we concentrate on the Level III reliability methods which were reviewed in Chapter 2. These methods, especially Monte Carlo family, Level III, are more capable of dealing with the complicated problems. This family includes importance sampling (IS), explained in Melchers (1999), directional sampling (DS), presented in Nie and Ellingwood (2000) and the extended methods addressed by Waarts (2000). Each of these method has its advantages and disadvantages when it is coupled with finite elements as was discussed in Chapter 2. Here we address a method which can be coupled with Monte Carlo and then integrated with IS, DS or the other methods in the Monte Carlo family.

In other words, an attempt is made in this study to overcome the addressed problem by applying the method of Dynamic Bounds in a practical case. This method accelerates the Monte Carlo process, benefiting from the property of monotonicity of the problem regarding its variables. 17th Street Flood Wall is considered as a case study to show the application of this method, and the results are compared with the Monte Carlo method.

6.2 Importance of the flood defences in the Netherlands

The Wiki definition of flood-defences in the Netherlands is fair enough to be quoted: "The Netherlands has been struggling against floods since the first people settled there. Over 60% of the country lies beneath mean sea-level. Countless people have lost their homes and their lives to floods from the sea or the rivers that could not be held by the flood-defences. The importance of the protection has led the Dutch to dedicate a Department solely to the protection against floods. Furthermore, local water-boards are an extra layer of government specially dedi-

Figure 6.1: Aerial Photograph of the 17th Street Canal Breach in New Orleans, USACE-e (2006).

cated to water management and protection against floods. This has resulted in a very high level of flood-protection. Flood-protection remains a continuous point of interest due to the vulnerability of the Dutch economy with regard to flooding" (Wikipedia (2008)). Besides, there are similarities between the Netherlands and New Orleans from the safety point of view. In fact, the main parts of the both lands are below sea level which are being protected by levees, dikes, barriers, and other flood defences. Moreover, the main industrial areas and the most densely populated region of the Netherlands lie below normal sea level as is shown in Figure 6.2. It shows that the capital, Amsterdam, and some of the biggest cities like Rotterdam and The Hague are below sea level. Therefore, a careful management of the flood defences is vital in the Netherlands as well as doing research about the assessment of the current situation of flood defences which motivates us to do research about the different aspects of failure of the flood protection system in New Orleans .

6.3 The flood wall at the 17th Street Canal, New Orleans

The 17th Street Flood Wall, the I-wall on the east side of the 17th Street Canal in New Orleans, was breached in hurricane Katrina when the water level was around

Figure 6.2: The vulnerable areas in the Netherlands against flooding according to www.kang.nl (2007).

8.0 feet as is shown in Figure 6.1. The failure of the flood wall at the 17th Street Canal was one of the flood wall failures that occurred by Katrina. Flood walls are constructed over a sheet pile penetrated into the dike as presented in Figure 6.3. This picture shows the flood wall including a concrete cap and concrete wall (I wall) located over a sheet pile penetrated into the levee, and soil materials. The materials of levee from above are two layers of clay, a thick layer of peat (March), and then a layer of mixed clay and clay laid over a thick sand layer. There is, in addition, a thin layer of sensitive clay located between March and the intermix zone. The normal water level on the left hand side is at the level of +1 foot (0.3 meter), which can be accumulated behind the flood wall up to the +11 feet (3.3 m). The source of geometry and material properties for this research project was the published materials by the Independent Levee Investigation Team (Team (2006)) and Interagency Performance Evaluation Task Force (Link et al. (2006); USACE-c (2006); USACE-e (2006)). Here, an attempt is made to show up a broader spectrum of possible behaviors of the typical I-wall structure. Moreover, we try to understand its full performance limits and to present new approaches to create adaptive designs based on physical behavior of engineering components, systems, and parameter variations.

6.3.1 Failure scenarios

A fault tree is used to separate the failure scenarios of a complicated system. Since the main function of the designed flood wall at the 17th Street Canal was protecting the city, its failure to do this job is the top event (TE) presented in the fault tree in Figure 6.4. In this figure, there are two main intermediate events showing the importance of a reliability analysis in two modes: expected and extreme conditions. In other words, a flood defence system should be stable with the expected loads and able to tolerate the overtopping or overflowing for the expected time. This means that a good design should consider the resiliency. The main intermediate events for the expected conditions in Figure 6.4 are sliding, piping, and failure of the concrete wall. Sliding was the main mechanism for the failure of the 17th street flood wall concluded in USACE-g (2006). Therefore, it is the sliding that forms the main focus of attention here. Nevertheless, the other failure modes are shortly discussed. However, a fault tree has some disadvantages which may lead designers to the wrong supposition that the failure modes can be separated while is not always true in a physical system. This problem is addressed in Section 6.3.2.

Sliding

The stability analysis of slopes is not an easy task because the failure mode and shape are depended on the geometry, boundary conditions, soil parameters, and their variations. In fact, an evaluation of some variables such as the soil stratifi-

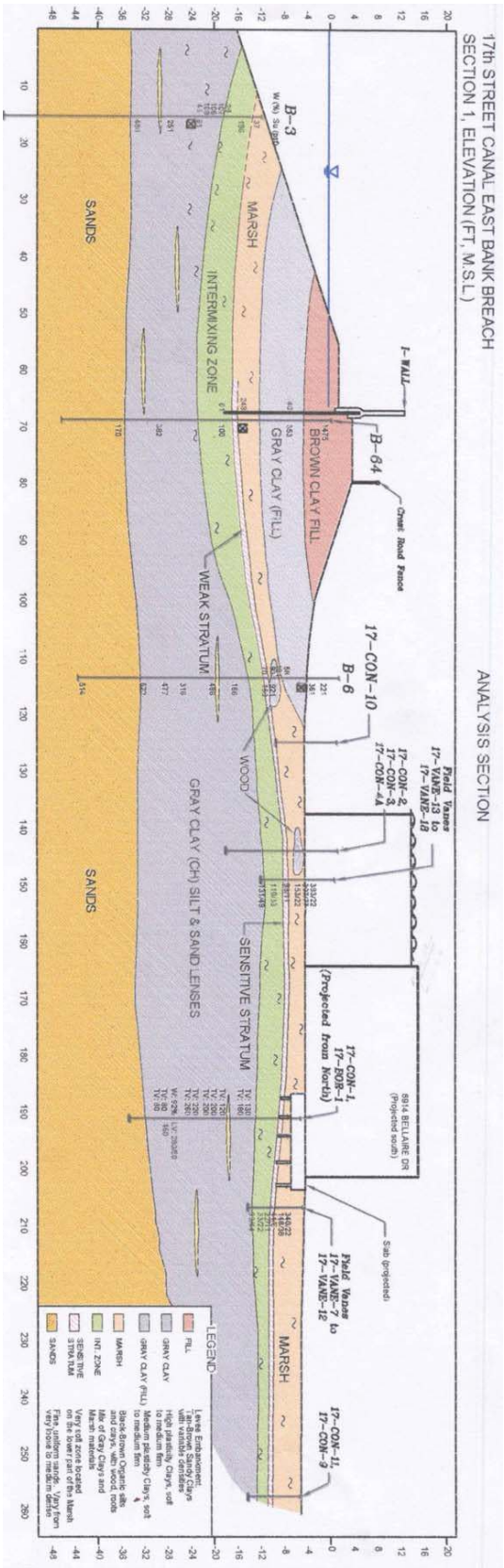


Figure 6.3: The cross section of 17th street flood wall and its foundation according to the Team (2006).

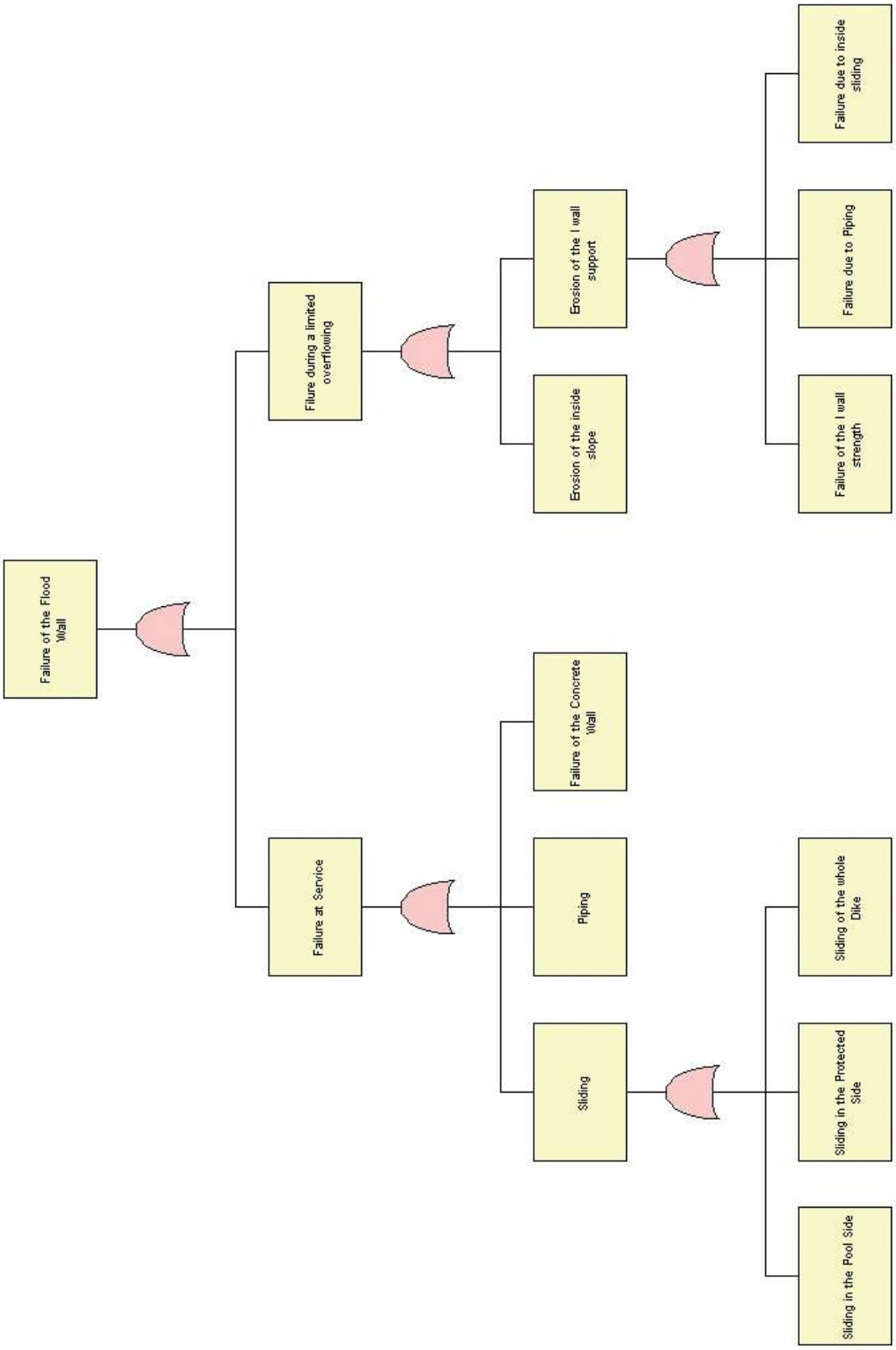
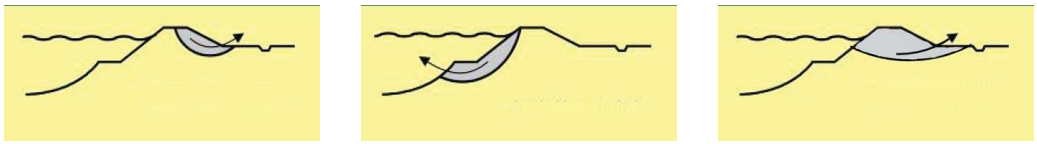


Figure 6.4: The fault tree of most expected failure mechanisms for the 17th street flood wall, New Orleans.



(a) Partial sliding (b) Partial sliding (c) Sliding of whole system

Figure 6.5: The sliding for a typical system of dikes can be expected in three categories; (a) partial sliding toward the landside, (b) partial sliding toward the pool-side, and (c) sliding of the whole dike or levee.

cation and its in-place shear strength may prove difficulty of the problem. Meanwhile, water seepage through the slope and the choice of a potential slip surface adds to the complexity of the problem. In general, sliding of a slope may occur partially toward the landside, pool-side or the whole system; these three modes of failure are presented in Figure 6.5.

Different methods are used to analyze the stability of slopes and assess the safety factor, defined as a ratio of resistant forces over the driving forces. Those methods can be divided into two categories; analytical methods and finite element methods. The analytical methods are based on the physical modeling of soil behavior via an imposed failure shape which can be circular or none circular. Bishop's famous model, which considers a circular failure shape, is one of the most widely applied methods in stability of slopes advocated for instance by BAKER and GARBER (1978). Finite elements models are the second category of analysis; they are more accurate and more reliable methods which do not assume any failure shape in advance, see for instance Griffiths and Lane (1999).

For sliding, the finite element analysis provides several advantages compared to the analytical approach. The first advantage is that there is no need to define a failure shape from the beginning. In fact, it finds the most probable failure shape according to the maximum deformation of the nodes. Besides, there is no need to consider separated failure modes; finite elements integrate all the failure modes and presents a more realistic approach.

Piping

Piping plays a big role in the failure of embankment dams, dikes, and other flood defences. The failure of the 17th Street Flood Wall, however, was not led by piping or seepage concluded by Team (2006). Therefore, in this study the probability of failure by seepage is not discussed in detail. However, there are explicit (analytical) limit state functions which can be approximately used to evaluate the probability of piping, which might be a case for further research indicated in Allsop (2007).

It is important to point it out that the piping itself can not be included in a simple finite element model. However, the sand boil phenomenon can be modeled. In fact, finite elements will assign the safety factor related to sand boil if its probability of failure is higher than sliding.

Overtopping and overflowing

In a design process of every flood defences, a probability of overtopping and overflowing should be taken into account. Overtopping is defined as a condition where waves pass over the flood defences. Also, overflowing is the situation where water continuously flows over the flood defences.

In the design phase or reliability assessment of a flood defence, the expected time of overflowing or overtopping can be estimated. Then, the flood defence can be resiliently designed taking into account the condition of the system under overloading. In other words, an infrastructure like a flood defence should sustain the overtopping or overflowing conditions to provide us with a resilient flood defence.

Resilience of the 17th street could have been considered. It could have been designed such that it had been remained stable for the condition of erosion of the flood wall support, overtopping, or overflowing. Meanwhile, the structural system should be able to tolerate more moment as a result of decreasing the support. These aspects are considered to the right of Figure 6.4.

Failure of the concrete wall

Failure of the concrete wall is considered as a failure mode in the fault tree analysis of the 17th street flood wall. Its safety factor can easily be estimated in a linear relation between the maximally resistant moment of flood wall minus the maximum moment produced in the cross section of the flood wall. Also, this failure mode can easily be taken into account by a finite elements model and can be integrated into one model. Having established that failure of the concrete wall was not the case at the 17th Street Canal, the properties of concrete wall are kept constant during our probabilistic analysis.

6.3.2 An integrated model of failure mechanisms

A big advantage of the finite elements model is its ability to integrate different failure modes into one accurate model. Therefore, the interaction of some of the failure mechanisms are taken into account. For instance, sometimes the sand boil or heave near the toe encourage the slope sliding process. These types of interaction can be taken into account by the finite element method. In other words, finite elements can consider the failure of a slope considering different condition which may cause failure. For instance, whether the water elevation

causes failure, or the internal water pressure of soil causes failure, or the slope of embankment is not enough, they are not separated from each other. This means that all the possible variables play their role in an integrated failure model, and the prediction of failure is not artificially separated.

6.3.3 Loads and resistance

For the 17th Street Flood Wall, the water level behind the flood wall produced the driving force. Therefore, by rising the water behind the flood wall, its probability of failure increases. Therefore, the probability of failure, p_f, is estimated for the normal water level and five other higher levels: MSL +4, +6, +8, +10, and +12 feet (1.2, 1.8, 2.4,3, and 3.6 meter). The last water level is higher than the elevation of the flood wall, but it is considered to simply evaluate the resiliency of this structure under the horizontal water pressure when overflowing occurs.

The main aspects of resistant variables at the 17th Street Flood Wall are soil parameters. There is a total of ten soil parameters which vary both in horizontal and vertical directions as presented in Table 6.1. The geotechnical data are selected according to the previous research projects conducted by USACE-e (2006) and Team (2006). Besides, a variety of information about the design of the structure is available on the Ipet website. All the parameters are assumed to be normally distributed, and the related coefficients of variation of soil layers are assumed. The summary of soil parameters and their variation are presented in Table 6.1. The soil layers are numbered according to the numbers presented in Figure 6.7, and more discussion on the model parameters is provided in Section 6.5.1.

6.4 Probabilistic finite elements

Probabilistic Finite Elements (PFE), applicable in different engineering fields, has advantages and some disadvantages. It is a powerful tool with accurate results which enables us to take into account the variation of the input parameters, but it is a time consuming process.

The finite element method is a powerful tool which presents accurate models based on the dominant differential equation of a phenomenon. This technique is one the most accurate and applicable methods which is used in different engineering fields and all over the world. But probabilistic finite elements even promises more. It provides a better understanding of the model as well as the contribution of every random variable in the final and desired output (see Section 6.5.5). More importantly, probabilistic finite elements provide a more accurate and reliable foundation for the reliability estimation considering variation of inputs. Figure 6.6 presents a variety of the possibilities which may be integrated in an accurate model. This figure shows that variation of the geometry, loads,

	Soil	Model	Behavior	Dis.	Param.	CV*
1	Brown Clay	MC	Undrained	Normal	C$^{\diamond}$	0.2
2	Gray Clay	MC	Undrained	Normal	C	0.2
3	Marsh Under Levee	MC	Undrained	Normal	C	0.3
4	Marsh Free Field	MC	Undrained	Normal	C	0.3
5	Sensitive Layer-Under Levee	MC	Undrained	Normal	C	0.3
6	Sensitive Layer-Free Field	MC	Undrained	Normal	C	0.3
7	Intermix Zone	SSM	Undrained	Normal	ϕ°	0.3
8	Gray Clay Horizontal	MC	Undrained	Normal	C	0.3
9	Gray Clay Vertical	MC	Undrained	Normal	C	0.3
10	Sand	MC	Drained	Normal	ϕ	0.3

Table 6.1: The variation of soil parameters considered in the probabilistic finite element analysis. * CV is the coefficient of Variation; the normal variation of soil parameters are considered. $^{\diamond}$ The cohesion of soil in the Mohr-columb soil model behavior. $^{\circ}$ The friction angle in the soil model.

material properties, environmental conditions, and dimension can be take into account in a probabilistic way to get an estimation for the reliability or the probability of failure. Accurate reliability estimation is necessary for risk estimation, risk management, planning management, and resiliency estimation. These issues have been addressed in a paper which presents the advantages of an accurate local estimation of the probability of the failure, called as micro scale risk assessment by Rajabalinejad (2008). Figure 6.6 also addresses the important role of the engineering judgment both in pre and post analysis. In other words, engineering judgment is necessary for considering a realistic variation of inputs and verification of the outputs. This supervision is necessary and more important than the ordinary finite element analysis. The reason for a higher importance is that there is a possibility of getting a randomly generated input data producing results far from the reality which are not acceptable by the engineering experience. In that case, some constraints might be applied to help getting more realistic output. These constraints, therefore, indirectly modify the probability density function (PDF) of the input variables in addition to the output results. Besides, we need to take into account the resiliency in the design process of an important structure. Resiliency is an step further than the reliability in the way toward a safer society. As a result, it is important to take into account the resiliency of the flood defences into account. Apart from the advantages provided by the probabilistic finite elements, we need to address its disadvantages. The main disadvantage of a PFE model is that it is a time consuming process. This matter has been addressed in Chapter 2. As a conclusion, the PFE does a great job in the modeling and analysis, and it is

Type	Type of element	Model	Total no.
Soil	15-node triangle	MC/SSM	289
Plate	5-node line	Elastoplastic	14
Interface	5-node line	Gap Elements	19

Table 6.2: Numbers and type of elements used in the 17th street canal model. MC stands for Mohr-Coulomb, and SSM stands for Soft Soil Model.

also beneficial for us to apply this powerful tool to our flood defences. However, it is expensive and we need to reduce its cost with some techniques which are discussed in Chapters 2, 3, 4, and 5. Here we implement the probabilistic finite elements into the case study of 17th Street Flood Wall. In this study the variation of soil parameters are considered, and the average variation of layers are assumed. Since the input variation of the parameters were assumed due to a lack of information, we implemented a constraint into the analysis which the model should remain stable under its own weight and without interfering of driving forces. In other words, if this assumption is not fulfilled, the values of input variables are rejected. Therefore, the intention of this chapter is presenting a method to calculate the safety of a structure, and we acknowledge that the outcomes may be infected by the assumption of coefficient of variation (CV) presented in Table 6.1.

6.5 Failure simulations

6.5.1 Model

A finite element program, Plaxis, is used to analyze the behavior of the 17th Street Flood Wall. The flood wall is modeled as depicted in Figure 6.7, according to the geometry used by Team (2006). This is the basic model which is used in the simulations. Table 6.2 presents type and number of elements used in the modeling. Triangular elements are used to model the soil, plate elements to model the flood wall and sheet-pile, and interface elements to model the separation of sheet pile (steel) and the surrounding soil. Interface elements are assumed to be fully impermeable.

Mohr-Columb (MC) and soft soil model (SSM) are used to model the soil behavior. It is preferred to mainly use MC criteria and reduce the calculation time, given the fact that the more advanced models like SSM require more calculation time. Besides, the MC Model gives good results for failure prediction.

Figure 6.8 shows a typical[1] deformation of the flood wall and its foundation when the water level is MSL +8 feet (2.4 m). The $C/tan(\phi)$ reduction technique is used to plot the most probable scenario. As is shown in the figure, the flood wall

[1]The analysis of the model with the mean value of its variables,(μ_{X_i}).

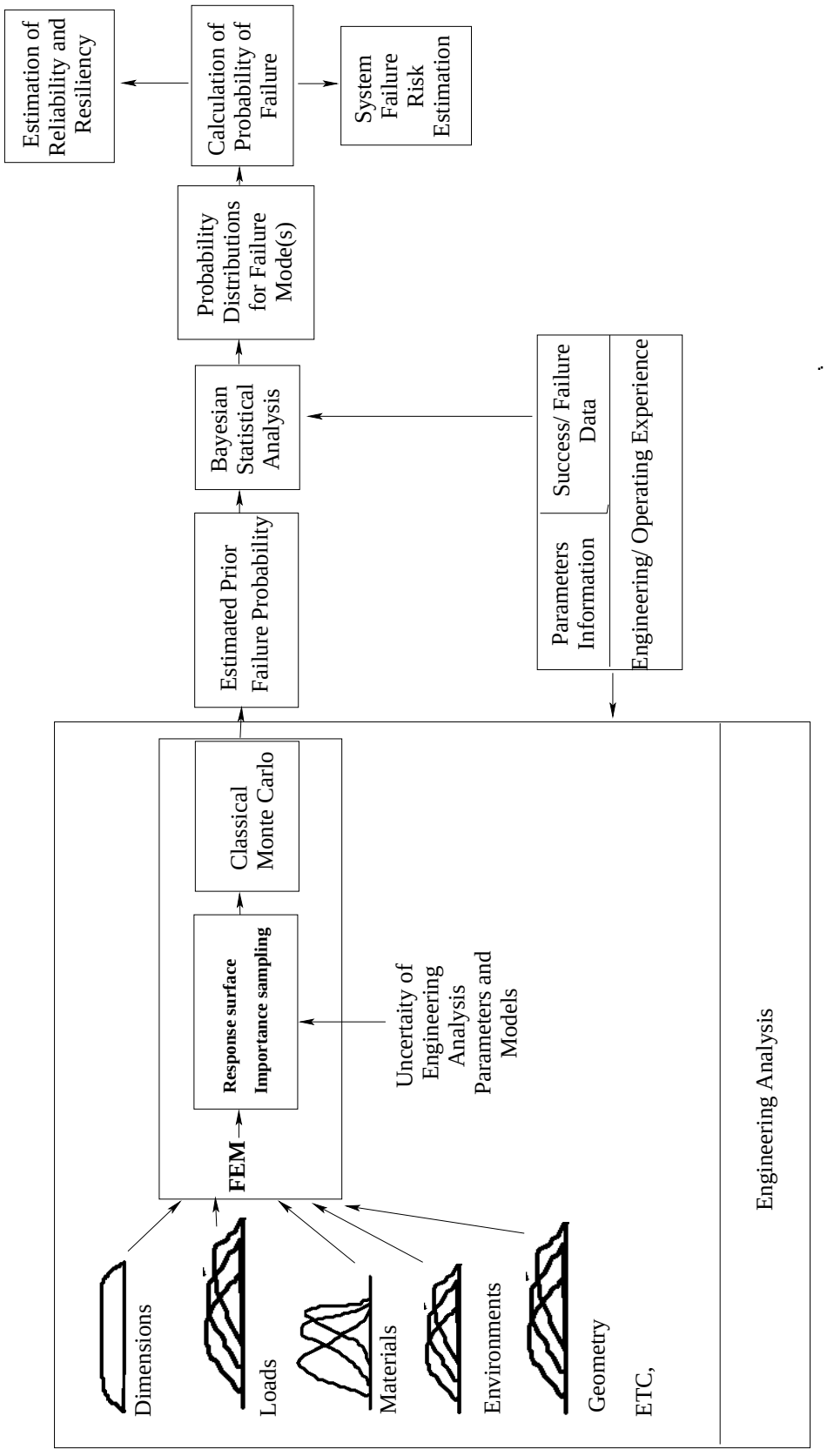


Figure 6.6: The general scheme of the probabilistic finite elements; only the variation of materials has been considered in this research.

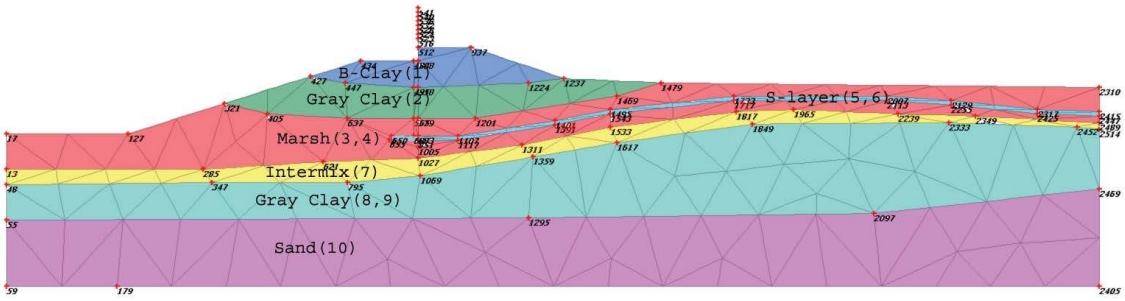


Figure 6.7: Plot of the finite element model for 17th street flood wall and its foundation, modeled with Plaxis. This plot is based on the cross section of the flood wall at 17th street canal and its foundation presented by Team (2006).

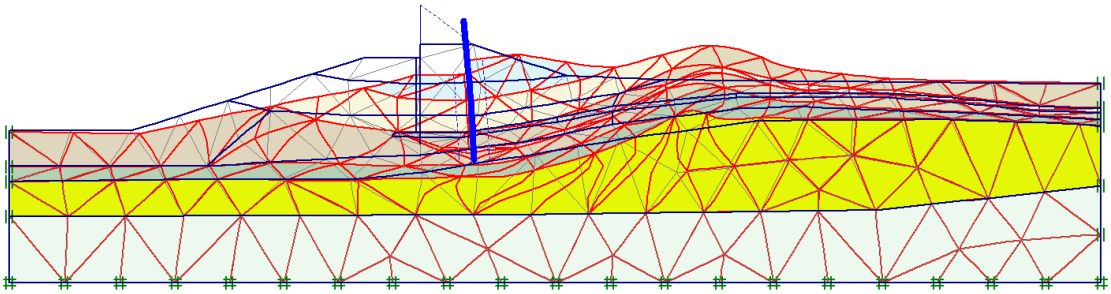


Figure 6.8: The plot of typical deformation of the 17th Street Flood Wall modeled by Plaxis. The water level of MSL +8 ft (2.4 m) and the mean value of the variables are selected in this analysis.

remains almost vertical when it is pushed and moved by the water pressure. This deformation was also observed during its collapse reported by USACE-e (2006); Team (2006).

The model behavior is compared with the results of Team (2006), as presented in Figure 6.9. The arrow in this figure shows the final conclusion about the behavior of the 17th street flood wall. The stars (*) in this figure are the calculated safety factors by mean value of resistant parameters introduced in Table 6.1. This figure shows a good correspondence between the results.

6.5.2 Monte Carlo process

For the probabilistic analysis, a code is developed to interactively work with Plaxis. This code feeds Plaxis with the desired probability density function of input variables, invokes it to calculate different adapted phases, and finally gathers the outputs. In other words, a finite element program is confined inside a probabilistic

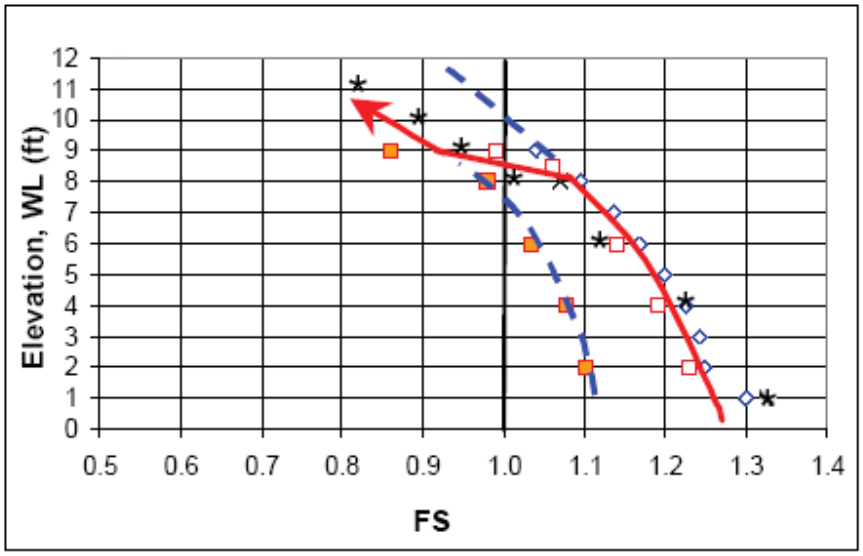


Figure 6.9: Calculated Safety Factors for the Plaxis models of 17th Street Canal breach. The arrow shows the final conclusion about the behavior of the 17th Street Flood Wall and transition between modes as suggested by Team (2006). The star notation (*) is added to present the correspondence of the model used in this study with the research project of Team (2006).

loop to provide a broader spectrum of physical behavior and better understanding of engineering components and the effects of their variations.

A mobile failure mode is presented in the analysis; the failure shape varies with the variations of parameters. This is possible by using the $C/tan(\phi)$ reduction technique, which reduces all the resistant parameters with a growing ratio in a step-by-step procedure. In fact, by every variation of soil parameters a new problem is defined and solved. For instance, by variation of the clay parameters in the foundation, the shape, curvature, and depth of sliding are adapted.

On the other hand, the variation of input parameters provides a wide range of different combinations of variables under the probability density function (PDF). Given independent variables, the probability of getting several variables at the same time under the tail of their PDF is very small. One of the main sources of this uncertainty is the uncertainty of distribution of input variables. But, it can also be recovered by constraints which can define how acceptable the results are. We here use a constraint to adjust our model, and this constraint is defined that the model is stable before increasing water level. In other words, if a certain combination of the input variables provide a condition that the system cannot sustain its own weight, it is not accepted. Therefore, we apply this concept to our probabilistic finite elements to improve our prior results and update it to a more accurate level.

6.5.3 Safety factor

Safety factor is defined as a ratio between resistance and driving forces: $F_s = \frac{Resistance}{Stress}$. Therefore, $F_s = 1$ is a threshold which clarifies the stability of a model or its failure for values greater or smaller than 1, respectively. In finite elements, however, we can investigate the stability or collapse of a model on the base of the maximum deformation. To calculate the safety factor the resistant parameters, mainly C and ϕ, are step by step reduced till failure of the model. Then, the ratio for the reduction is considered as a safety factor. In another approach, when the model is not stable itself, the ratio of loading is considered as the equivalent safety factor. In fact, when we start to step by step load the model up until its collapse, the percentage of loading can be considered as an indication of the safety factor.

6.5.4 Variation of safety factors

Safety factors are calculated as a result of the simulations of the 17th Street Flood Wall's behavior under different combinations of input variables, using the $C/tan(\phi)$ reduction technique. This safety factor includes the process of slope sliding, heave, or failure of the flood wall or sheet pile in the finite elements analysis. The calculated histograms are presented in Figure 6.10. These histograms present the distribution of the safety factors for different water levels. It is shown that the histogram of safety factors moves toward smaller values when the water

	N	W.L. (ft)	$\hat{p}_f$%	n_f	$N \geq$	$V(\hat{p}_f)$
1	1218	+1 (0.3 m)	*	5	97000	> 0.05
2	1218	+4 (1.2 m)	*	246	2000	> 0.05
3	1218	+6 (1.8 m)	*	254	1600	> 0.05
4	1218	+8 (2.4 m)	43.6	687	520	0.032
5	1218	+10 (3.0 m)	82.9	1010	100	0.013
6	1218	+12 (3.6 m)	95.6	1165	20	0.006

Table 6.3: The probability of failure estimated in different water levels for the 17th Street Flood Wall, New Orleans. $n_f$is the number of failures. * More calculations are needed to estimate p_f.

level increases. It can also be concluded that the response of a system is almost normal when this system has a high reliability or when it certainly fails.

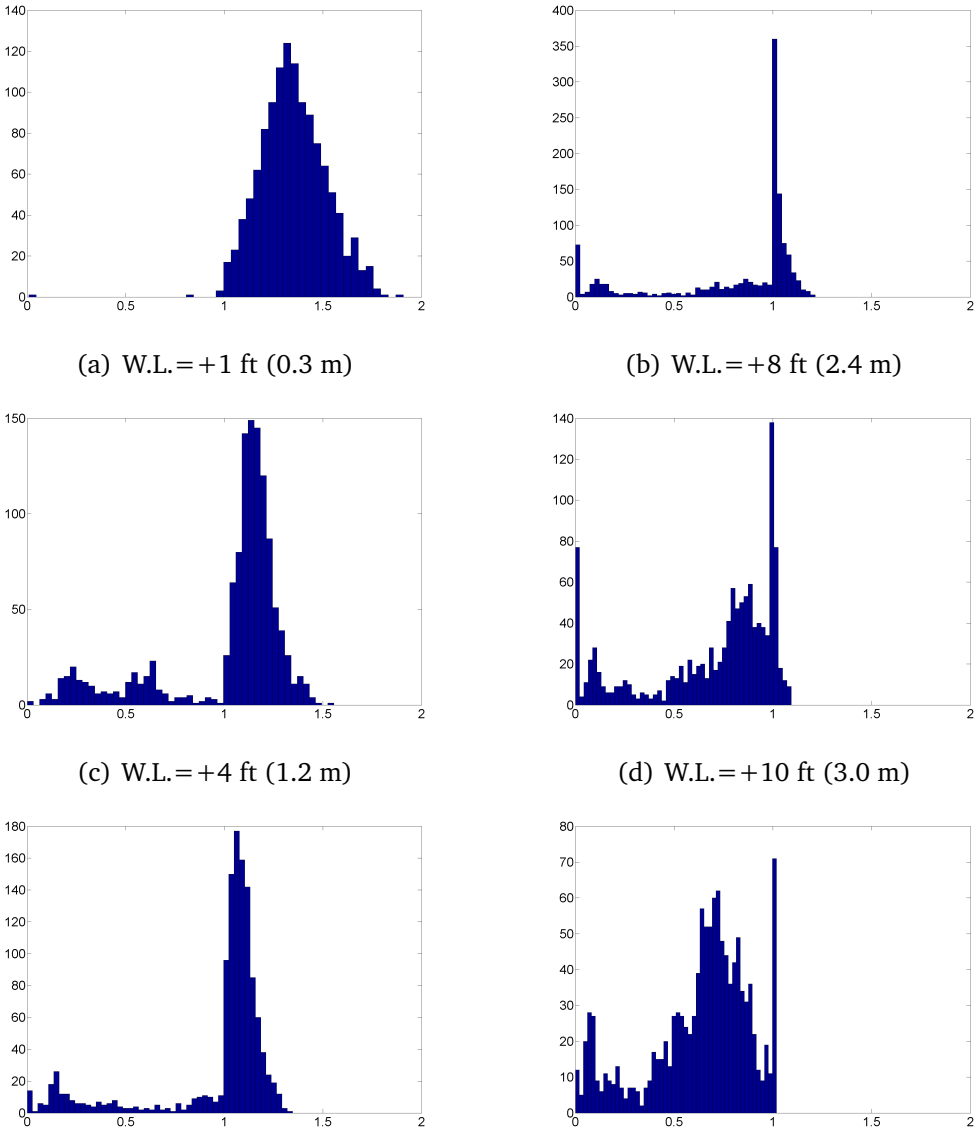
6.5.5 Probability of failure

Having the calculated safety factors, the cumulative distribution functions (CDF) of safety factors are plotted in Figure 6.11 for different water levels behind the flood wall. According to these figures, the integration of area less than 1 presents the probability of failure. Therefore, by increasing the water level the area under the CDF gets bigger as well as the CDF shifts toward the smaller values. Another way to calculate the probability of failure which is used here is dividing the number of failures over the total number of simulations. Failure probabilities are calculated for different water levels and presented in Table 6.3. Please note that the number of calculations (N) in order to obtain an accurate result can be estimated by Equation 6.1. In many engineering works, we accept the coefficient of variation of the estimate p_f less than 0.05: $V(\hat{p}_f) \leq 0.05$. Where it is defined as $\frac{\sigma[X]}{E[X]}$. This value is also called the coefficient of variation. Therefore, we can get the required number of simulation presented in Table 6.3.

$$N \geq \frac{1}{V(\hat{p}_f)^2} \times (\frac{1}{\hat{p}_f} - 1). \tag{6.1}$$

The probability of failures based on classical Monte Carlo simulations is presented in Table 6.3.

Figure 6.12 presents the calculated safety factors in a normal plot. It can be said that for two extreme conditions, where the safety factors are barely below one and where the safety factors are barely above one, the response is close to the normal distribution. This fact can also be seen in Figure 6.10 and 6.11. In fact, its main reason is the difference between two processes of calculating safety factors when they are above or below one.



(a) W.L. = +1 ft (0.3 m)

(b) W.L. = +8 ft (2.4 m)

(c) W.L. = +4 ft (1.2 m)

(d) W.L. = +10 ft (3.0 m)

(e) W.L. = +6 ft (1.8 m)

(f) W.L. = +12 ft (3.6 m)

Figure 6.10: Histograms of the calculated safety factors for different water levels for the 17th Street Flood Wall, New Orleans.

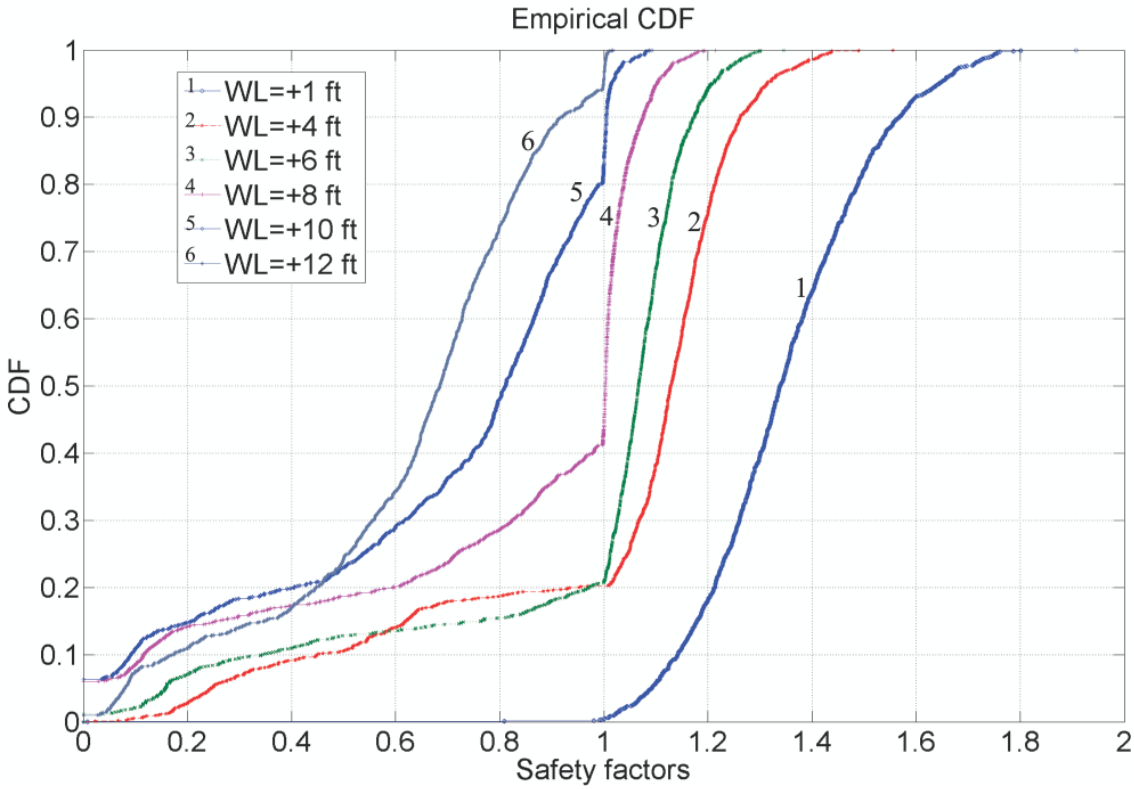


Figure 6.11: The plot of cumulative distribution functions (CDF) versus the safety factor (F_s) for different water levels; they show the probability of failure of the 17th street flood wall, New Orleans.

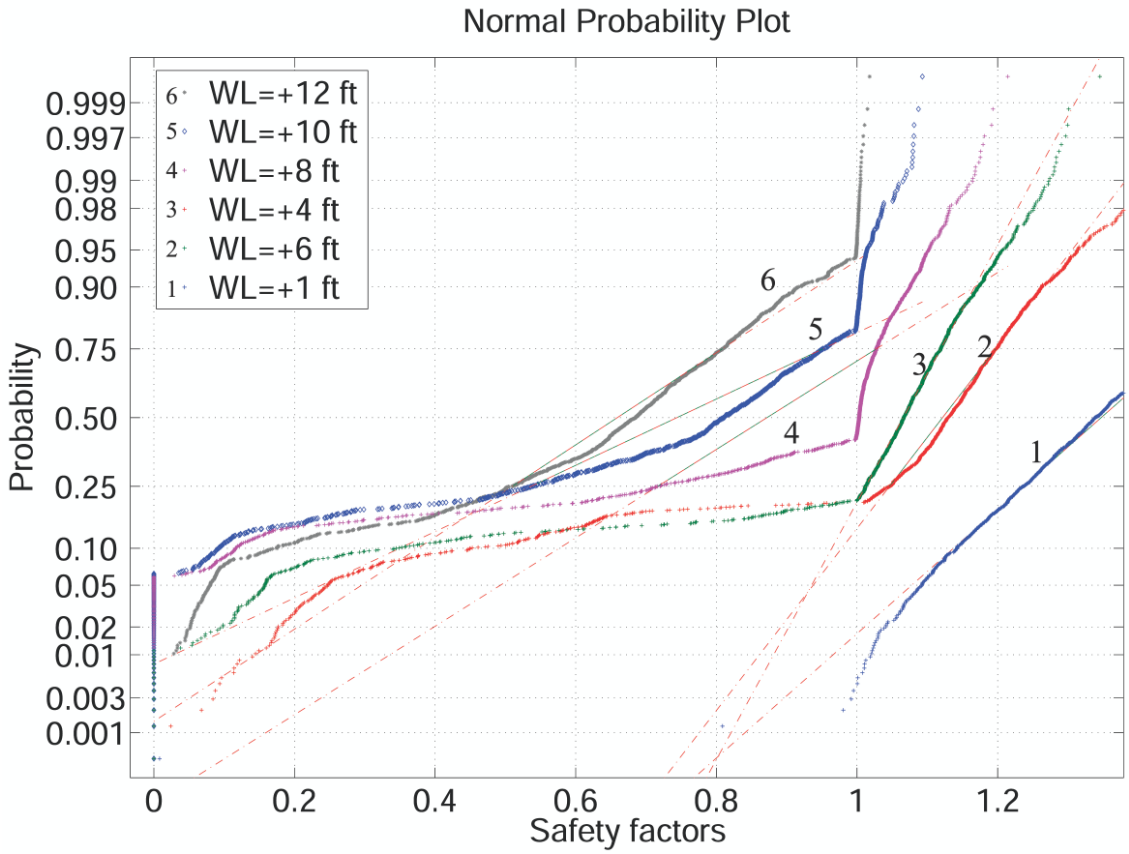


Figure 6.12: The plot of probability distribution functions (PDF) versus the safety factors (F_s) for different water levels; they show the probability of failure of the 17th street flood wall, New Orleans.

6.5.6 Estimation methods for contribution to the failure

The contribution of every variable, X_i, in level III reliability methods can be established according to different tools described in this section. On the basis of these formula, the linear or nonlinear relation of each variable, X_i, regarding the limit state function, Z, can be calculated. We call X_i a base variable and Z a predicted variable. For instance, in the case of the flood wall, base variables are the soil parameters and Z (or the predicted variable) is safety factor. Therefore, estimation of the correlation between basic and predicted variables define their contribution into the failure. In other words, having a higher correlation between a basic variable (X_i) and the predicted variable (Z), a bigger contribution of that variable into the failure is expected.

Product moment correlation

The product moment correlation or Pearson product moment correlation defines a linear relation between two variables of X_i (base variable) and Z (predicted variable) by Equation 6.2. It can take values between -1 and 1; these two boundary limits present a completely linear relation between variables when $Z = aX_i \mp b$, where a, b are two constants.

$$\rho(X_i, Z) = \frac{Cov(X_i, Z)}{\sigma_{X_i}.\sigma_Z} \tag{6.2}$$

$$Cov(X_i, Z) = E(X_i Z) - E(X_i)E(Z). \tag{6.3}$$

Correlation ratio and linearity index

The correlation ratio of the predicted variable Z and base variable X_i is a square product moment correlation between G and a function $f(x)$ which maximizes this correlation as presented in Equation 6.4.

$$CR(X_i, Z) = \max_f \rho^2(X_i, Z) \tag{6.4}$$

On the other hand, Equation 6.4 is maximized if $f(X_i) = E(Z|X_i)$ presented in Kurowicka and Cooke (2006), therefore:

$$CR(X_i, Z) = \rho^2(X_i, E(Z|X_i)) = \frac{Var(E(Z|X_i))}{Var(Z)} \tag{6.5}$$

Equation 6.5 presents a ratio of the variance of the conditional expectation of Z given X_i and the variance of Z. Since the square of the product moment correlation is less than or equal to $CR(X_i, Z)$, Equation 6.6 can measure the nonlinearity of $E(Z|X_i)$. Therefore, the bigger the difference, the higher the nonlinear relation are expected.

$$\rho^2(X_i, E(Z|X_i)) - \rho^2(X_i, Z). \tag{6.6}$$

Rank correlation

Spearman rank correlation is a good measurement for two variables which are nonlinearly related and they have monotone relationship. It is suggested that the rank correlation is the best option to present the relation between parameters in the monotonic problems. Spearman rank correlation is defined by the following Equation:

$$\rho_r(X, Z) = \frac{C_x + C_Z - \sum_{i=1}^{n} d_i^2}{2\sqrt{C_x C_Z}}, \tag{6.7}$$

where

$$C_x = \frac{n^3 - n}{12} - \sum_{tx} \frac{t_x^3 - t_x}{12},$$
$$C_Z = \frac{n^3 - n}{12} - \sum_{t_z} \frac{t_z^3 - t_z}{12},$$
$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} [R(x_i) - R(z_i)]^2.$$

Index t_x and t_y stand for the number of observations of X and Z with the same rank and $R(x_i), R(z_i)$ stand for the rank ordered X and Z variables presented in William et al. (1992).

6.5.7 Contribution of variables to failure

Equation 6.2, which is usually used in engineering applications, is based on the linear correlation between basic variable (X_i) and the predicted variable (Z). Since, we cannot be concerned about the linear relation, all three methods explained in Section 6.5.5 are applied to the flood wall case study. Further discussion is necessary to make our decision to select the best tool. Table 6.4 presents the calculated product moment correlation (ρ), correlation ratio (CR), and rank correlation (ρ_r) for three water levels. This table is ranked according to the absolute value of product rank correlation for MSL +8 ft (2.4 m), as it is common in engineering works.

It is clear in Table 6.4 that the Marsh layer and Gray Clay layer (Layer numbers 3 and 8 in Figure 6.7) make the biggest contribution. In fact, product moment correlation (ρ), correlation ration (CR), or rank correlation (ρ_r) certify this conclusion. Yet, in the next step, sorting the parameters according to their real influence

on the probability of failure depends on the criteria which are selected. According to the previous discussion in Section 6.5.5 the product moment is appropriate when the relation between the basic variables, X_i, and the predicted variable, Z, is linear. The correlation ratio needs interpolation of variables to be calculated, as is shown in Equation 6.5. Here, the third order polynomial interpolation is assumed according to the visualization of data and regression coefficients. However, the value of CR is sensitive to the interpolation function. For instance, the $E(Z_8|X_3)$[2] and $E(Z_8|X_5)$, where the Z_8 is the vector of safety factors of the flood wall for MSL +8 ft (2.4 m), X_3 is the soil number 8, and X_5 is the soil number 5, are interpolated by the Equation 6.8 and Equation 6.9, respectively.

$$E(Z_8|X_3) = -0.0182 + 0.0055X_3 - 0.0000{X_3}^2 + 0.0000{X_3}^3. \tag{6.8}$$

$$E(Z_8|X_5) = 0.2110 + 0.0110X_5 - 0.0001{X_5}^2 + 0.0000{X_5}^3. \tag{6.9}$$

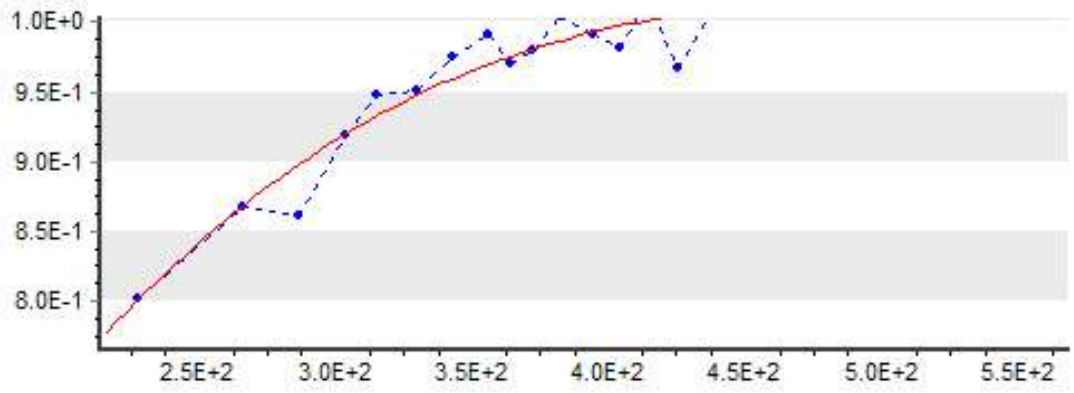
To make the interpolation more clear, Figures 6.13(a) and (b) present the graph of Equations 6.8 and 6.9, and the related graph of expectation of Z_8 given X_3 and X_5. In these Figures it is clear that the stability of the flood wall is sensitive to layer number 3 and almost indifferent to layer number 5.

The rank correlation has two advantages and apparently provides a good criteria for ranking of variables. First, it figures out nonlinear correlations; second, it is a suitable choice when there is monotonicity, which is also an essential assumption in DB technique. Therefore, the Spearman ratio or rank correlation will be a good option for ranking the variables in geotechnical flood defence problems. This conclusion requires further investigation.

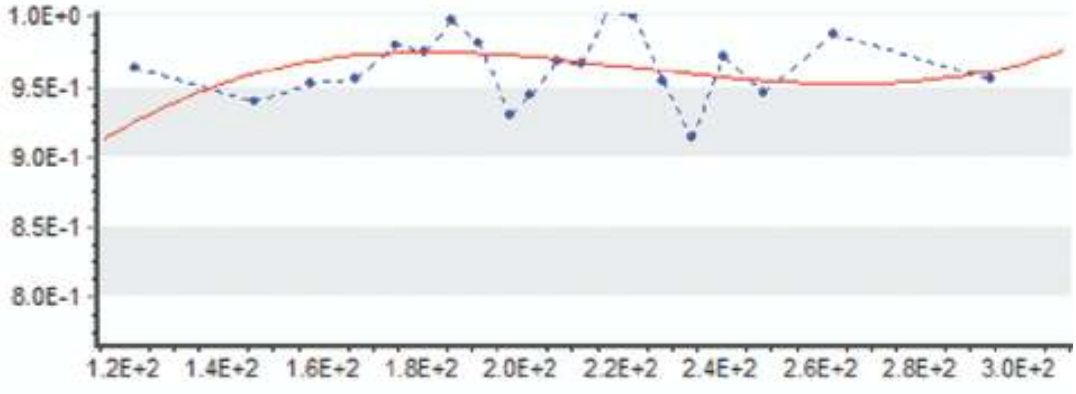
6.6 Dynamic bounds (DB) applied to the flood wall

Dynamic bounds (DB) provide two main advantages when they are coupled with Monte Carlo simulations. First, it can speed up the simulation when there are a limited number of variables. Second, it makes it possible to store and use the generated bounds for the next Monte Carlo simulations. These two advantages may bring the probabilistic finite elements from research into the practical field of risk assessment for flood defences. Besides, considering the fact that the efficiency of DB increases for a higher number of simulations, this method is suitable when

[2]The expectation of Z_8 given X_3; Z_8 is the vector of safety factors of the flood wall for MSL +8 ft (2.4 m), X_3 is the soil number 3



(a) $E(Z_8|X_3)$ versus variation of soil number 3.



(b) $E(Z_8|X_5)$ versus variation of soil number 8.

Figure 6.13: The expectation of failure of the 17th Street Flood Wall for MSL +8 ft (2.4 m) versus variation of Marsh under levee (soil number 3) in figure (a) and Sensitive layer(soil number 5) in figure (b). The third order polynomial interpolation curve is depicted together with the pedestrian graph of expectation presented by the dotted line.

Material	Soil Nu.	Water Level +4 (ft)			Water Level +8 (ft)			Water Level +10 (ft)		
		ρ(%)	CR(%)	ρ_r(%)	ρ(%)	CR(%)	ρ_r(%)	ρ(%)	CR(%)	ρ_r(%)
Marsh Under Levee	3	39.1	13.5	68.9	35.9	16.8	49.3	54.6	31.2	67.7
Gray Clay Horizontal	8	27.8	10.3	43.8	24.2	14.9	54.7	28.4	14.8	36.5
Gray Clay	2	22.2	10.2	16.5	19.4	10.3	23	24.4	9.9	22.6
Intermix Zone	7	11.7	3	6.3	15.54	4.2	12.3	17.1	4.7	16.8
Gray Clay Vertical	9	11.9	0.8	14.9	9.3	0.69	18.3	12.2	1.6	16.4
Marsh Free Field	4	20.8	3	35.3	10.4	0.7	28.3	11.6	1.2	11
Sensitive Layer-Free Field	6	3.1	0.2	6.3	2.1	0.05	4.2	4.7	0.0	4.6
Sensitive Layer-Under Levee	5	1.3	0.001	6.1	0.7	0.01	2.5	3.5	0.04	3.5
Brown Clay	1	3.4	0.02	6.6	2.3	0.05	2.1	4.6	0.04	3.4
Sand	10	0.3	0.1	0.3	2.8	0.04	4.5	3.5	0.03	0.04

Table 6.4: The variation of soil parameters considered in the probabilistic finite element analysis. [1] CV is the coefficient of Variation. ρ is the product moment correlation (see Equation 6.2). CR is the correlation ratio according to Equation 6.4. ρ_r is the rank correlation (see Equation 6.7). In order of converting units, 1 ft=0.3 meter.

Material	Soil Nu.	WL=+4 ft		WL=+8 ft		WL=+10 ft	
		ρ	ρ_r	ρ	ρ_r	ρ	ρ_r
Marsh Under Levee	3	.391	.689	.359	.493	.546	.677
Gray Clay	8	.278	.438	.242	.547	.284	.365

Table 6.5: The first couple of influential variables in the stability of the flood wall; using product moment correlation, rank correlation, or correlation ratio conduct to these variables.

a higher number of Monte Carlo is required. Therefore, the probabilistic finite element with its great advantages can be used to calculate the probability of failure of many structures. To present this approach, the coupled system of dynamic bounds with Monte Carlo is presented in this section. The accuracy of product moment correlation and rank correlation in ranking of influential variables is also investigated.

6.6.1 DB considering two variables

The first two variables of Table 6.4 are selected according to their correlation and presented in Table 6.5. As a matter of fact, accepting each of the methods introduced in Section 6.5.5 leads to this selection. These variables are the main influential soil layers and they play the major role in failure of the flood wall. Previous analyses of the 17th street canal, moreover, shows that the Marsh layer under the embankment(layer number 3) with the Gray layer(layer number 8) are the main layers, see Rajabalinejad et al. (2007a).

Applying the DB technique to these variables conclude Table 6.6. The second column of this table shows the calculated number of DB in which a low variance is obtained[3]. The calculated probability of failure for different water levels in this table are close to the values presented in Table 6.3; it also verifies that these two variables control the stability condition of the flood wall.

6.6.2 DB considering three variables

In this section three are implemented in DB to simulate a more accurate behavior of the 17th Street Flood Wall, and estimate its probability of failure. In this case, the first three influential variables are considered. Yet, different ranking criteria give different outputs. In fact, a good and efficient criterion is essential for ranking variables in DB. The results of this section can provide us two conclusions; first, having a comparison between the results of three dimensional DB and Classical Monte Carlo clarifies that if the obtained accuracy is enough, or using a higher

[3]It is assumed that $V(\hat{p}_f) < 0.05$ is acceptable.

W.L. (feet)	DB	Failures	Stables	Equival. MC	$\hat{p}_f$	$V(\hat{p}_f)$
+4 (1.2 m)	145	2594	7395	10000	25.9	0.017
+6 (1.8 m)	97	601	1294	1895	31.7	0.038
+8 (2.4 m)	119	603	897	1500	40.2	0.031
+10 (3.0 m)	72	469	131	600	77.8	0.022
+12 (3.6 m)	38	431	24	455	93.5	0.012

Table 6.6: The probability of failure calculated by the 2D-DB of variables of Table 6.5.

Material	Soil Nu.	WL=+4 ft		WL=+8 ft		WL=+10 ft	
		ρ	ρ_r	ρ	ρ_r	ρ	ρ_r
Marsh Under Levee	3	.391	.689	.359	.493	.546	.677
Gray Clay	8	.278	.438	.242	.547	.284	.365
Gray Clay	2	.222	.165	.194	.23	.244	.226

Table 6.7: The first three influential variables in stability of the 17th Street Flood Wall; variables are selected according to the product moment correlation, ρ (see Equation 6.2).

dimension of DB is necessary. Second, the influence of variables according to product moment correlation and rank correlation can be investigated to clarify which criteria are more efficient for ranking.

Therefore, the first three influential variables are selected according to the product moment correlation (ρ) and rank correlation (ρ_r) in Table 6.7 and Table 6.6.2, respectively. The correlation ratio (CR) is not considered because of the fact that its results are dependent on the interpolation function. Meanwhile, the higher degree of interpolation function does not necessarily yield better results.

Table 6.7 shows that soil number 3, 8, and 2 are the most important parameters in the failure of the flood wall for different water levels. Table 6.6.2, however, presents that soil number 3, 8, and 4 are the main influential variables and their sequence is changed when the structure's behavior is nonlinear in W.L.=+8 ft (2.4 m).

In the next step, three dimensional DB are applied to the variables presented in Tables 6.7 and 6.6.2. The results are presented in Tables 6.8 and 6.6.2. Now, a comparison between results of these tables with the results of Classical Monte Carlo (CMC) shows that the results in Table 6.6.2 are closer to the CMC results. This is an important point for using DB technique.

As a result, it can be concluded that both the product moment correlation (ρ) and the rank correlation (ρ_r) provide similar results. Therefore, it is advised to do more research in order to determine which of these methods is better.

W.L. (feet)	DB	Failures	Stables	Equival. MC	p_f	$V(\hat{p}_f)$
+4 (1.2 m)	190	424	1076	1500	28.2	0.041
+6 (1.8 m)	231	670	830	1500	44	0.029
+8 (2.4 m)	221	781	719	1500	52	0.024
+10 (3.0 m)	52	162	32	194	82.2	0.035
+12 (3.6 m)	23	102	6	108	95.1	0.022

Table 6.8: The probability of failure calculated by DB coupled with MC for the first three variables according to product moment correlation (Table 6.7).

Material	Soil Nu.	WL=+4 ft		WL=+8 ft		WL=+10 ft	
		ρ	ρ_r	ρ	ρ_r	ρ	ρ_r
Gray Clay	8	.278	.438	.242	.547	.284	.365
Marsh Under Levee	3	.391	.689	.359	.493	.546	.677
Marsh Free Field	4	.208	.353	.104	.283	.116	.11

Table 6.9: The first three influential variables in stability of the 17th Street Flood Wall; variables are selected according to rank correlation, ρ_r (see Equation 6.7).

W.L. (feet)	DB	Failures	Stables	Equival. MC	$\hat{p}_f$	$V(\hat{p}_f)$
+4 (1.2 m)	171	377	1123	1500	25.1	< 0.05
+6 (1.8 m)	207	526	939	1465	35.6	< 0.05
+8 (2.4 m)	202	322	438	1500	42.9	< 0.05
+10 (3.0 m)	66	157	37	194	80.9	< 0.05
+12 (3.6 m)	29	102	8	110	92.7	< 0.05

Table 6.10: The probability of failure calculated by DB coupled with MC for the first three variables according rank correlation (Table 6.6.2).

6.7 Summary of results

According to the monotonic property, presented in many geotechnical flood defence systems, dynamic bounds are applied to this research; they not only maintain the accuracy of a Monte Carlo method, but also speed up the simulation process and store the boundaries.

The performance of rank correlation (ρ) is investigated and compared to product moment correlation (ρ_r). The results show that rank correlation is a good tool for ranking the basic variables. Figure 6.14 shows contribution of soil parameters to the failure of the 17th Street Flood Wall according to rank correlation.

The dynamic bounds (DB) method is coupled with Monte Carlo to simulate the behavior of the 17th Street Flood Wall. Figure 6.15 presents the calculated probability of failures with different methods and different water levels. The presented methods are classical Monte Carlo, two dimensional DB according to Table 6.5 (shown in the figure as DB2D), three dimensional DB ranked by product moment correlation (see Table 6.7), and three dimensional DB ranked by rank correlation (see Table 6.6.2) . As a result, three dimensional DB can be used by product moment correlation or rank correlation to estimate the reliability of the flood wall.

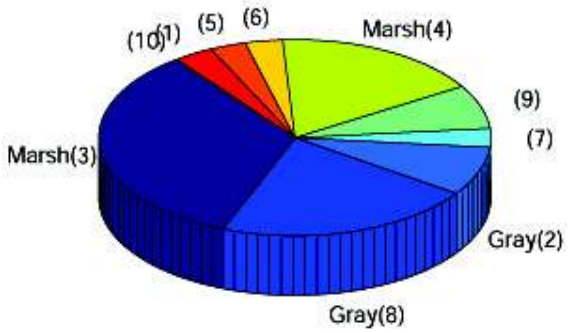
Figure 6.15 also presents behavior of the flood wall under a water level beyond the design criteria which can be considered as resiliency. In fact, the colored part of this figure presents the probability of failure of the 17th Street Flood Wall in overflowing or overtopping under hydrostatic pressure. In general, having this plot for a flood defence system, we would be able to assess its behavior, risk, and reliability.

6.8 Conclusion

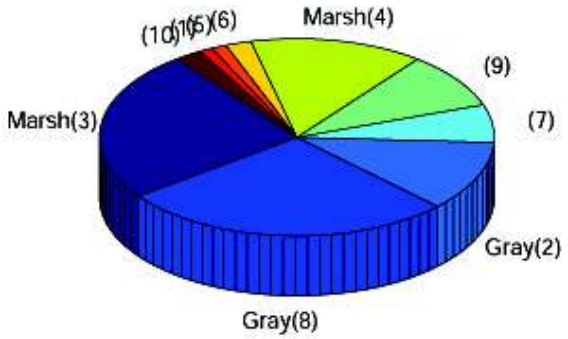
A probabilistic approach provides a better understanding of the failure mechanisms, occurrence probabilities, as well as consequences of failure. However, to achieve these advantages, a well defined model of the structure together with a robust reliability method are needed. In flood defence systems, the finite element method is a good tool for the accurate modeling as well as integrating different failure mechanisms. Also risk and reliability assessment of flood defences is very important for the improvement of the condition and prevent from catastrophes as a result of a weak flood defence.

In the present study, an attempt is made to introduce a probabilistic method integrated with finite element analysis to estimate[4] the probability of failure of flood defences. The behavior of the 17th Street Flood Wall as a case study is investigated by this method. It is shown that DB speeds up the Monte Carlo

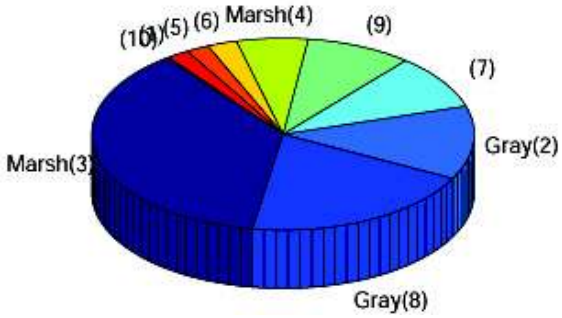
[4]As it is indicated in the introduction, the variation of soil parameters are assumed to estimate the reliability of this structure; this assumption will infect the calculated probability of failure.



(a) W.L.=+4 ft (1.2 m)



(b) W.L.=+8 ft (2.4 m)



(c) W.L.=+10 ft (3.0 m)

Figure 6.14: The contribution of different variable parameters (in Table 6.1) on the probability of failure of the 17th street flood wall, New Orleans; they are ranked according to the rank correlation (ρ) and the soil numbers are according to the Table 6.1. Every figure is related to a specified water level.

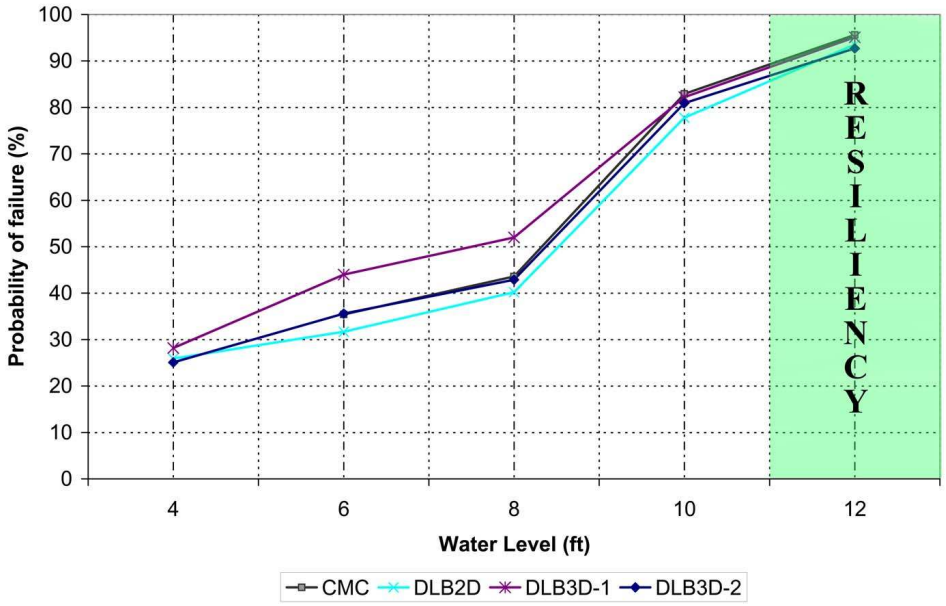


Figure 6.15: The estimated probability of failure of the 17th Street Flood Wall for different water levels. Different reliability techniques are coupled with finite elements analysis: CMC presents the Classical Monte Carlo, DB2D presents two dimensional dynamic bounds (DB), DB3D-1 presents three dimensional DB considering the rank correlation, and DB3D-2 presents three dimensional DB considering product moment correlation.

simulation technique and stores the boundary values. The stored boundary values, then, can be used to make the next simulations very cheap and available.

Resilient flood defences should be seriously taken into account in design processes. If as is shown for the 17th Street Flood Wall, overtopping or overflowing occurs how reliable this structure is? The results also certify the conclusion of the research team that "Resilience was not an element in the New Orleans Hurricane Protection System design" concluded by USACE-e (2006).

It was theoretically concluded and also observed in the 17th Street Flood Wall that both of the product moment correlation and rank correlation provide a good tool to rank the influence of variables for the problems having monotonicity. A good ranking criterion is important when dynamic bounds are used.

7

Conclusion and further research

7.1 Conclusion

In this dissertation, some drawbacks of the applied reliability methods were discussed, and an attempt is made to introduce new methods for evaluation of an implicit or explicit limit state equations (LSE) assuming that the LSE is a time-consuming process which makes it impractical to use normal reliability methods mainly because of the prohibitive computing effort.

Monte Carlo (MC) is accepted as a basic method which presents high accuracy given enough simulations. In this method, all possible combination of the variables are considered given enough number of realizations. The method of dynamic bounds (DB), based on the Monte Carlo, is introduced to take the advantage of monotonicity and a limited number of variables, which is usually the case in engineering problems. This method, as a result, is fast and robust and can be integrated with complicated limit state equations like finite elements. Its main advantage over direct Monte Carlo simulation is that only a fraction of the limit state function evaluations (finite element analyses) is needed, without loss of accuracy.

By breaking up the simulation in two or more stages, initial estimates of the computing effort to attain a required level of accuracy can be updated at intermediate stages, resulting in good predictions of computation costs. Besides, the bounds can be stored for the next series of simulations, which presents another important advantage of this method. Moreover, the method can be coupled with the importance sampling technique, further reducing the required calculations, speeding up the whole procedure.

DB divides the range of a LSE into three parts of stable, unstable, and unqualified. Attempts can be made to shrink the unqualified part as it is addressed by improved dynamic bounds (IDB), given the order of the LSE with respect to

its variables. As a result, the DB is improved by defining a lower bound of the response surface in the unqualified region. This is done on the base of a minimization process. Improved dynamic bounds (IDB) presents more efficiency and accuracy without implementing more uncertainty into the model. It still can be coupled with importance sampling technique, and more importantly the bounds can be stored for the next series of simulations.

The method of Bayesina Monte Carlo (BMC) is introduced where the Bayesian technique is implemented in Monte Carlo simulations. In this study, an attempt is made to get the prior information of the model incorporated to the current level of the analysis. This is a step forward in Monte Carlo simulations. In fact, a link was presented among the information of each pixel and its neighbors. In other words, information of each point passes through this link and effect the others. Besides, this approach provides a nice tool to get the other priors incorporated to the Monte Carlo simulations. For instance, the dynamic bounds method, presented in Chapter 3, assumes monotonicity which can be implemented as an informative prior.

Give more information, do less computation is the common base of DB, IDB, and BMC. The Monte Carlo method is reluctant to further information over the limit state equation (LSE). The method of DB takes the monotonic behavior of a model into account which is mainly presented in engineering problems. The method of IDB requests some information over the form of a LSE to shrink the size of unqualified space based upon the certain bounds. BMC is a more general and promising method, which takes into account the prior information over the LSE or from the previous simulations. In conclusion, All of the suggested procedures can speed up the Monte Carlo simulations integrated with finite elements or the other highly complex and time consuming processes. The proposed methods also provide a tool for implementing informative priors regarding the considered model.

In addition to the theoretical background, the theories discussed above are in correspondence with real problems. For instance, the method of dynamic bounds is implemented to a complicated flood defence structure in the 17th Street Flood Wall to assess its probability of failure with less effort than the MC. DB technique is implemented in this research which speeds up the Monte Carlo and stores the stable and unstable bounds. The stored bounds, then, can be used to make the next simulations very cheap and accessible.

It was theoretically concluded and also observed (in the 17th Street Flood Wall) that the rank correlation provides a good tool to rank the influence of variables for the problems having monotonicity. A good ranking criterion is important when dynamic bounds are going to be applied.

7.2 Further research

Dynamic bounds (DB) is strongly recommended to be applied to the important structures like flood defences in the Netherlands.

Improved dynamic bounds (IDB) suggests extending the dynamic bounds of a limit state equation (or a model) by applying some extra information. However, the Monte Carlo process and development of the stable, unstable, and unqualified spaces are stochastic processes. Therefore, the unqualified part of the range of a LSE can be considered as a part of a stochastic process. Therefore it is possible to extended dynamic bounds without any certain assumption over the order of the polynomial; however, one expects a controlled amount of uncertainty contributing to the result.

Bayesian Monte Carlo is a promising method which can be implemented to many problems without any severe limitation.

Toward applications. The materials of this dissertation are suggested for solving practical problems, and it is recommended to utilize these theories and bring the reliability analysis more into the everyday engineering work.

References

Allsop, W. (2007). Failure mechanisms for flood defence assets. Technical Report T04-06-01, HR Wallingford Ltd.

Baker, R. (2007). Comments on 'on two definitions of the factor of safety commonly used in the finite element slope stability analysis' by hong zheng, l.g. tham and defu liu [computers and geotechnics 33 (2006) 188-195]. *Computers and Geotechnics*, 34(2):124–125.

BAKER, R. and GARBER, M. (1978). THEORETICAL-ANALYSIS OF STABILITY OF SLOPES. *GEOTECHNIQUE*, 28(4):395–411.

Bjerager, P. (1990). Methods for structural reliability computations. *Course on: General Principles and Applications in Mechanics of Solids and Structures, International Center for Mechanical Sciences, Udine, Italy*.

Box, G. and Wilson, K. (1954). The exploration and exploitation of response surfaces: some general considerations and examples. *Biometrics*, 10(1):16–60.

Bretthorst, G. L. (1992). Bayesian interpolation and deconvolution. Cr-rd-as-92-4, The advanced sensor diretorate research, Alabama 35898-5000.

de Normalisation, C. E. (1994). *Eurocode 1: Basis of Design and Actions on Structures: Part 1 Basis of Design*. European Committee for Standardization.

Deák, I. (1980). Three digit accurate multiple normal probabilities. *Numerische Mathematik*, 35(4):369–380.

Fiessler, B., Rackwitz, R., and Neumann, H. (1979). Quadratic Limit States in Structural Reliability. *Journal of the Engineering Mechanics Division*, 105(4):661–676.

Griffiths, D. V. and Lane, P. A. (1999). Slope stability analysis by finite elements. *Geotechnique*, 49(3):387–403. Times Cited: 46.

Haldar, A. and Mahadevan, S. (2000). *Reliability Assessment Using Stochastic Finite Element Analysis*. Wiley.

Hammersley, J. and Handscomb, D. (1964). *Monte Carlo methods*. Methuen.

Hasofer, A., Lind, N., Christian, N., and of Waterloo. Solid Mechanics Division, U. (1973). *An Exact and Invariant First-order Reliability Format*. Solid Mechanics Division, University of Waterloo.

Jaynes, E. T. (2003). *Probability theory, the logic of science*. Cambridge University Press.

Jonkman, S. N., Stive, M. J. F., and Vrijling, J. K. (2005). New orleans is a lesson to the dutch. *Journal of Coastal Research*, 21(6):XI–XII. Times Cited: 2.

Kanning, W., van Baars, S., van Gelder, P., and Vrijling, J. K. (2007). Lessons from new orleans for the design and maintenance of flood defence systems. In *European Safety and Reliability Conference (ESREL 2007)*, pages 937–942.

Kurowicka, D. and Cooke, R. (2006). *Uncertainty Analysis with High Dimensional Dependence Modeling*. Delft University of Technology.

Link, L. E., Jaeger, J. J., Stevenson, J., Stroupe, W., Mosher, R. L., Martin, D., Garster, J. K., Zilkoski, D. B., Ebersole, B. A., and Westerwink, J. J. (2006). Performance evaluation of the new orleans and southeast louisiana hurricane protection system: Draft final report of the interagency performance evaluation task force, vol. 1. *United States Army Corps of Engineers (USACE), vols. 1e9*.

Mai, C. V., van Gelder, P., and Vrijling, J. K. (2006). Safety of coastal defences and flood risk analysis. *Safety and Reliability for Managing Risk, Vols 1-3*, pages 1355–1366. Times Cited: 0 Soares, CG European Safety and Reliability Conference (ESREL 2006) SEP 18-22, 2006 Estoril, PORTUGAL.

Melchers, R. E. (1999). *Structural reliability analysis and prediction*. John Wiley, Chichester ; New York, 2nd edition. 98053106 GB99-20653 Robert E. Melchers. ill. ; 26 cm. Includes bibliographical references (p. [409]-430) and index.

Nie, J. and Ellingwood, B. (2000). Directional methods for structural reliability analysis. *Structural Safety*, 22(3):233–249.

Ouypornprasert, W. (1988). Adaptive Numerical Integration for Reliability Analysis. *Report No. 12*, 87.

Rajabalinejad, M. (2004a). Collapse of the divergent tunnel of Silveh dam and its remediation plans. *Road and Tunnel*, I(III):11–21.

Rajabalinejad, M. (2004b). Editorial, the necessity of risk management after Bam earthquake. *Road and Tunnel*, I(II):1–2.

Rajabalinejad, M. (2004c). Tuiserkan tunnel; its collapse and the best reconstruction methods. *Road and Tunnel*, I(I):3–11.

Rajabalinejad, M. (2008). A systematic approach to risk mitigation. In E. Beauchamp-Akatova, J. v. d. T., editor, *The research Agenda of Risk and Design*, volume I of *Anno*, pages 76–86, Delft, the Netherlands. Faculty of Technology, Policy and Management, Delft University of Technology.

Rajabalinejad, M., Kanning, W., van Gelder, P., Vrijling, J. K., and van Baars, S. (2007a). Probabilistic assessment of the flood wall at 17th street canal, new orleans. *Risk, Reliability and Societal Safety, Vols 1-3*, pages 2227–2234. Aven, T European Safety and Reliability Conference (ESREL 2007) JUN 25-27, 2007 Univ Stavanger, Stavanger, NORWAY.

Rajabalinejad, M., Meester, L., van Gelder, P., and Vrijling, J. K. (2007b). Dynamic bounds in monte carlo simulations. *Submitted for publication*, Reference number:STRUCS-D-07-00050R1.

Rajabalinejad, M., Noorzad, A., and Vrijling, J. K. (2006a). Seepage in the weathered foundation of Abshineh dam (IRAN), Part I: Erosion in the foundation of Abshineh dam. In *Proceedings of Sea to Sky Geotechnique*, volume 2.

Rajabalinejad, M., Noorzad, A., and Vrijling, J. K. (2006b). Seepage in the weathered foundation of Abshineh dam (IRAN), Part II: Evaluation of remedial methods. In *Proceedings of Sea to Sky Geotechnique*, volume 2.

Rajabalinejad, M., van Gelder, P., and van Erp, N. (2008a). Application of bayesian interpolation in monte carlo simulation. In S. Martorell, C. Guedes Soares, J. B., editor, *Safety, Reliability and Risk Analysis (ESREL)*, volume I, pages 705–713, Valencia, Spain. Taylor and Francis Group, London, UK.

Rajabalinejad, M., van Gelder, P., and Vrijling, J. (2008b). Probabilistic finite elements with dynamic limit bounds: A case study: 17th street flood wall, new orleans. In Prakash, S., editor, *6th International Conference on Case Histories in Geotechnical Engineering*, volume I, Rolla, Missouri USA. Missouri University of Science and Technology.

Rajabalinejad, M., van Gelder, P., and Vrijling, J. K. (2008c). Improved dynamic limit bounds in monte carlo simulations. In *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, volume I of *American Institute of Aeronautics and Astronautics*, Schaumburg, IL, USA. American Institute of Aeronautics and Astronautics.

Rajabalinejad, M., van Gelder, P. H. A. J. M., Vrijling, J. K., van Baars, S., and Kanning, W. (2007c). The Probabilistic Finite Elements. In *Proceeding of European Symposium on Flood Risk Management Research (From extreme events to citizens involvement)*.

Sivia, D. S. (1996). *Data Analysis: A Bayesian Tutorial*. Clarendon Press.

Smith, I. and Griffiths, D. (2004). *Programming the Finite Element Method*. Wiley.

Team, I. L. I. (2006). Investigation of the Performance of the New Orleans Flood Protection Systems in Hurricane Katrina . Technical Report Chapet 8, Berkely University.

Triebel, H. (1978). *Interpolation theory, function spaces, differential operators*. North-Holland.

Turner, M. J., Martin, H. C., and Leible, R. C. (1964). Further development and applications of stiffness methods. In *Matrix Methods of Structural Analysis*, volume 1, pages 203–266. Macmillian, New York, 1st edition.

USACE-b (2006). Orleans and Southeast Louisiana Hurricane Protection System, Volume II Geodetic Vertical and Water Level . Technical Report 167, U.S. Army Corps of Engineers, Report of the Interagency Performance Evaluation Task Force.

USACE-c (2006). Orleans and Southeast Louisiana Hurricane Protection System, Volume III The Hurricane Protection System . Technical Report 378, U.S. Army Corps of Engineers, Report of the Interagency Performance Evaluation Task Force.

USACE-d (2006). Orleans and Southeast Louisiana Hurricane Protection System, Volume IV The Storm . Technical Report 264, U.S. Army Corps of Engineers, Report of the Interagency Performance Evaluation Task Force.

USACE-e (2006). Orleans and Southeast Louisiana Hurricane Protection System, Volume V The Performance Levees and Floodwalls . Technical Report 86, U.S. Army Corps of Engineers, Report of the Interagency Performance Evaluation Task Force.

USACE-f (2006). Orleans and Southeast Louisiana Hurricane Protection System, Volume VI The Performance Interior Drainage and Pumping . Technical Report 53, U.S. Army Corps of Engineers, Report of the Interagency Performance Evaluation Task Force.

USACE-g (2006). Orleans and Southeast Louisiana Hurricane Protection System, Volume VII The Consequences . Technical Report 210, U.S. Army Corps of Engineers, Report of the Interagency Performance Evaluation Task Force.

USACE-h (2006). Orleans and Southeast Louisiana Hurricane Protection System, Volume VIII Engineering and Operational Risk and Reliability Analysis . Technical Report 59, U.S. Army Corps of Engineers, Report of the Interagency Performance Evaluation Task Force.

Van Gelder, P., Mai, C. V., Wang, W., Shams, G., Rajabalinejad, M., and Burgmeijer, M. (2008a). Data management of extreme marine and coastal hydrometeorological events. *Journal of Hydraulic Research*, 46(2):191–210. Times Cited: 2.

van Gelder, P. and Vrijling, J. K. (1998). Risk-averse reliability-based optimization of sea defenses. *Risk-Based Decision Making in Water Resources Viii*, pages 61–76. Times Cited: 0 Haimes, YY 8th Engineering Foundation Conference on Risk-Based Decision Making in Water Resources OCT 12-17, 1997 SANTA BARBARA, CALIFORNIA.

Van Gelder, P., Wang, W., and Vrijling, J. K. (2007). Statistical estimation methods for extreme hydrological events. *Extreme Hydrological Events: New Concepts for Security*, 78:199–252. Times Cited: 0 Vasiliev, OF NATO Advanced Research Workshop on Extreme Hydrological Events JUL 11-15, 2005 Novosibirsk, RUSSIA.

Van Gelder, P. H., Rajabalinejad, M., and al., e. (2008b). Reliability analysis of flood defence structures and systems in europe. In *Proceedings of the European Conference on Flood Risk Management, Research into Practice*, volume I, pages 715–723, Oxford, the UK. HR Walingford.

Vrijling, J. K., van Hengel, W., and Houben, R. J. (1998). Acceptable risk as a basis for design. *Reliability Engineering and System Safety*, 59(1):141–150. Times Cited: 12.

Vrijling, J. K., Vanhengel, W., and Houben, R. J. (1995). A framework for risk-evaluation. *Journal of Hazardous Materials*, 43(3):245–261. Times Cited: 16.

Vrijling, J. K., VanHengel, W., and Houben, R. J. (1996). Acceptable risk: A normative evaluation. *Stochastic Hydraulics '96*, pages 87–94. Times Cited: 0 Tickle, KS 7th IAHR International Symposium on Stochastic Hydraulics JUL 29-31, 1996 MACKAY, AUSTRALIA.

Waarts, P. (2000). *Structural reliability using finite element analysis*. Ph. D. thesis, Delft University of Technology, Delft, The Netherlands, 189 pp.

Wikipedia (2008). Flood control in the netherlands.

William, H., Brian, P. F., Saul, A. T., and William, T. V. (October 30, 1992). *Numerical Recipes in C : The Art of Scientific Computing*. CambridgeUniversity Press; 2 edition.

www.floodsite.net (2004-2008). Integrated flood risk analysis and management methodologies.

www.kang.nl (2007). Royal dutch geographical society.

A

Prior

A.1 Derivation of the prior

Equation 5.15 presented the joint probability density function of the pixels, given the prior information as

$$\begin{aligned} P(U|) &\equiv P(u_0, \cdots, u_{v+1}|\phi, I) \\ &\propto \int \tfrac{1}{\phi^{v+2}} \exp\left\{-\tfrac{\Delta^2}{2\phi^2} F(u)\right\} \times \underbrace{\cdots du_i \cdots}_{i\neq j}, \end{aligned} \tag{A.1}$$

where,

$$F(u) = [(\frac{u_1 - u_0}{\delta_0})^2 + \sum_{i=1}^{v} (\frac{u_i - u_{i-1}\delta_{r,i} - u_{i+1}\delta_{l,i}}{\delta_i})^2 + (\frac{u_{v+1} - u_v}{\delta_{v+1}})^2]$$

and

$$\Delta = \delta_0 + \cdots + \delta_{v+1}.$$

$(F(u)$ which is a part of the prior may be rewritten to the Matrix form.

$$F(u) = [(\frac{u_1 - u_0}{\delta_0})^2 + \sum_{i=1}^{v} (\frac{u_i - u_{i-1}\delta_{r,i} - u_{i+1}\delta_{l,i}}{\delta_i})^2 + (\frac{u_{v+1} - u_v}{\delta_{v+1}})^2].$$

To write $F(u)$ in the matrix form, it needs to be rewritten as

$$\begin{aligned} F(u) = &\tfrac{u_1^2 - u_0^2 - 2u_1u_0}{\delta_0^2} + \tfrac{u_{v+1}^2 - u_v^2 - 2u_{v+1}u_v}{\delta_{v+1}^2} \\ &+ \sum_{i=1}^{v} \tfrac{u_i^2 - 2\,u_iu_{i-1}\delta_{r,i} - 2\,u_iu_{i+1}\delta_{l,i} + u_{i-1}^2\delta_{r,i}^2 + 2\,u_{i-1}\delta_{r,i}u_{i+1}\delta_{l,i} + u_{i+1}^2\delta_{l,i}^2}{\delta_i^2}. \end{aligned}$$

It still may be extended as

$$F(u) = \frac{u_1^2}{\delta_0^2} - 2\,\frac{u_1 u_0}{\delta_0^2} + \frac{u_0^2}{\delta_0^2} + \frac{u_1^2}{\delta_1^2} - 2\,\frac{u_1 u_0 \delta_{r,1}}{\delta_1^2} - 2\,\frac{u_1 u_2 \delta_{l,1}}{\delta_1^2} + \frac{u_0^2 \delta_{r,1}^2}{\delta_1^2} + 2\,\frac{u_0 \delta_{r,1} u_2 \delta_{l,1}}{\delta_1^2} + \frac{u_2^2 \delta_{l,1}^2}{\delta_1^2} + \frac{u_2^2}{\delta_2^2} - 2\,\frac{u_1 u_2 \delta_{r,2}}{\delta_2^2} - 2\,\frac{u_2 u_3 \delta_{l,2}}{\delta_2^2} + \frac{u_1^2 \delta_{r,2}^2}{\delta_2^2} + 2\,\frac{u_1 \delta_{2,2} u_3 \delta_{l,2}}{\delta_2^2} + \frac{u_3^2 \delta_{l,2}^2}{\delta_2^2} + \dots$$

Then, the rearrangement of the prior leads toward the matrix multiplication as presented in Equation A.2.

$$F(u) = \mathbf{u}^T \mathbf{R}\, \mathbf{u}, \tag{A.2}$$

where

$$\mathbf{u} = (u_0, u_1, \dots, u_{v+1}),$$

and

$$\mathbf{R} \equiv \Delta^2 \times \begin{pmatrix} \frac{1}{\delta_0^2} + \frac{\delta_{r,1}^2}{\delta_1^2} & \frac{-1}{\delta_0^2} - \frac{\delta_{r,1}}{\delta_1^2} & \cdots & \cdots & 0 \\ \frac{-1}{\delta_0^2} - \frac{\delta_{r,1}}{\delta_1^2} & \ddots & \ddots & \vdots & \\ \frac{-\delta_{l,2}}{\delta_2^2} - \frac{\delta_{r,3}}{\delta_3^2} & & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \frac{\delta_{l,v}.\delta_{r,v}}{\delta_v^2} \\ \vdots & \ddots & \ddots & \ddots & \frac{-1}{\delta_{v+1}^2} - \frac{\delta_{l,v}}{\delta_v^2} \\ 0 & \cdots & \cdots & \frac{-1}{\delta_{v+1}^2} - \frac{\delta_{l,v}}{\delta_v^2} & \frac{1}{\delta_{v+1}^2} + \frac{\delta_{l,v}^2}{\delta_v^2} \end{pmatrix}.$$

$\mathbf{R}$ is an $(v+1) \times (v+1)$ matrix, and its more completed representation is given in Equation 5.16. Now using Equation A.2 the prior Equation A.1 can be rewritten as

$$P(U|\phi, I) \propto \frac{1}{\phi^{v+2}} \exp\left\{-\frac{1}{2\phi^2}\mathbf{u}^T \mathbf{R}\mathbf{u}\right\}. \tag{A.3}$$

B

Likelihood

B.1 Derivation of the Likelihood

Given the prior and the data point, Equation 5.21 presented the likelihood. The exponent of Equation 5.21 can be written in the matrix form.

$$\Delta^2 \times \sum_{i=0}^{v+1} \frac{(d_i - u_i z)^2}{\delta_i^2}. \tag{B.1}$$

where

$$z \equiv \begin{cases} 1, & \text{if data is measured in pixel } i, \\ 0, & \text{elsewhere.} \end{cases}$$

and

$$d_i \equiv \begin{cases} d_i, & \text{if data is measured in pixel } i, \\ 0, & \text{elsewhere.} \end{cases}$$

The matrix representation of Equation B.1 can be written as

$$\begin{aligned} \Delta^2 \sum_{i=0}^{v+1} \frac{(d_i - u_i z)^2}{\delta_i^2} &= \Delta^2 \sum_{i=0}^{v+1} \left(\frac{d_i^2 - 2 d_i u_i z + u_i^2 z}{\delta_i^2} \right) \\ &= \Delta^2 \sum_{i=0}^{v+1} \frac{d_i^2}{\delta_i^2} - 2 \sum_{i=0}^{v+1} \frac{d_i u_i z}{\delta_i^2} + \sum_{i=0}^{v+1} \frac{u_i^2 z^2}{\delta_i^2} \\ &= \Delta^2 [\mathbf{d}\delta]^T [\mathbf{d}\delta] - 2\Delta^2 [\mathbf{d}\delta]^T [\mathbf{u}\delta] + \mathbf{u}^{\mathrm{T}} \mathbf{S} \mathbf{u}. \end{aligned} \tag{B.2}$$

Where the vector of data $\mathbf{d}$ has zero value wherever there is no data value assigned to a given pixel, and it has the data value wherever there is a data assigned to a pixel. Therefore, $\mathbf{S}$ is an $(v+1) \times (v+1)$ diagonal matrix with $\frac{\Delta^2}{\delta_i^2}$ on every

diagonal element, where the data has been measured and zero everywhere else. For example, if $\mathbf{d} = [d_1, 0, 0, d_2, 0, d_3, 0]^T$ then the $\mathbf{S}$ is

$$\mathbf{S} = \begin{pmatrix} \frac{\Delta^2}{\delta_1^2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\Delta^2}{\delta_4^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\Delta^2}{\delta_6^2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

And δ is a diagonal matrix of the size $(v+1) \times (v+1)$ where $\frac{1}{\delta_i}$ is located on each correlated cell. It is, therefore, in the form of

$$\delta = \begin{pmatrix} \frac{1}{\delta_0} & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \frac{1}{\delta_1} & 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \frac{1}{\delta_2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & \frac{1}{\delta_v} & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & \frac{1}{\delta_{v+1}} \end{pmatrix}.$$

Using equation Equation B.2, the likelihood function, Equation 5.18, can be rewritten as

$$P(D|\sigma, U, I) \propto \frac{1}{\sigma^N} \exp\left\{-\frac{1}{2\sigma^2}\left(\Delta^2[\mathbf{d}\delta]^T[\mathbf{d}\delta] - 2\Delta^2[\mathbf{d}\delta]^T[\mathbf{u}\delta] + \mathbf{u}^{\mathrm{T}}\mathbf{S}\mathbf{u}\right)\right\}, \quad \text{(B.3)}$$

which is equivalent to Equation 5.21.

B

List of Symbols

This section lists the definition of used symbols in alphabetical order. They are additionally explained in the text when they first appear. Units and abbreviations are not included in this list. They will be defined in the text when first used.

Abbreviations

BMC	Bayesian Monte Carlo
CDF	Cumulative Distribution Function
DB	Dynamic Bounds
DBIS	Dynamic Bounds integrated with Importance Sampling
FORM	First Order Reliability Method
JPDF	Joint Probability Density Function
IDB	Improved Dynamic Bounds
IDBIS	Improved Dynamic Bounds integrated with Importance Sampling
IS	Importance Sampling
LSE	Limit State Equation
LSF	Limit State Function
MC	Monte Carlo
NI	Numerical Integration
PDF	Probability Density Function
Rel.s.e	Relative Standard Error
SORM	Second Order Reliability Method
TE	Top Event in a fault tree

Small letters

$\mathbf{c}$	A vector of the positions of data points: $\mathbf{c} = (c_1, \ldots, c_n)$
$\mathbf{d}$	A vector of the observed data points: $\mathbf{d} = [d_0, \ldots, d_{v+1}]^T$
e_i	Error related to the i*th* pixel in the prior of BMC
e_c	Error related to the i*th* pixel in the likelihood of BMC
p_{s}	Probability of a random vector lies above the upper bound
p_{s}^e	Probability of a random vector lies above the extended upper bound
$\hat{p}_{\mathrm{s}}$	An estimate for p_{s}
$\hat{p}_{\mathrm{s}}^e$	An estimate for p_{s}^e
p_{u}	Probability of a random vector lies below the lower bound
p_{u}^e	Probability of a random vector lies below the extended lower bound
$\hat{p}_{\mathrm{u}}$	An estimate for p_{u}
$\hat{p}_{\mathrm{u}}^e$	An estimate for p_{u}^e
p_{f}	Probability of failure
$\hat{p}_{\mathrm{f}}$	An estimate for p_{f}
n_{f}	Number of failure in Monte Carlo simulation
p_Δ	Probability of a random vector hits the unqualified space
u_j	Target pixel in BMC
$\mathbf{u}$	A vector of pixel values: $\mathbf{u} = [u_0, \ldots, u_{v+1}]^T$
v	Number of internal pixels
x_i	A scaler value

Capital letters

I	Prior information in BMC
D	The JPDF of data points in BMC
$\mathbf{G}$	Combined matrix of the prior and likelihood
N	(mainly) Number of simulation in Monte Carlo
$\mathbf{Q}$	Presents the relation of the pixel values according to the prior
$\mathbf{R}$	The covariance matrix of the pixel values which comes out of the prior
$\mathbf{S}$	A diagonal matrix which relates to the position of data vlues
S	The set of stable bound in DB
S_b	The set of points on the stable bound in DB
S_b^e	The set of points on the extended stable bound in IDB
U	The set of lower bound in DB
U	The JPDF of pixel values in BMC
U_b	The set of points on the unstable bound in DB
U_b^e	The set of points on the extended unstable bound in IDB
Z	A random variable which represents $G(\vec{x})$

Small Greek

α	Influence factor of variables in FORM
β	Reliability index in the FORM
δ_i	Distance of left and right neighboring pixels
$\delta_{l,i}$	Implemented weight to the right neighbor pixel
$\delta_{r,i}$	Implemented weight to the left neighbor pixel
ϵ	The regularizer in BMC, $\epsilon = \frac{\sigma}{\phi}$
σ	Standard deviation of the likelihood
ξ	The horizontal local axis in IDB
η	The vertical local axis in IDB

Capital Greek

Δ	Unqualified space in DB and IDB
Δ	Two times of the distance between u_{v+1} and u_0
Λ	Eigenvalues of A matrix
$\mathbf{\Lambda}$	Eigenvalues of A matrix in the singular decomposition format
$\mathbf{\Lambda}_R$	Eigenvalues of the matrix $\mathbf{R}$
$\mathbf{\Lambda}_{\epsilon R}$	Eigenvalues of the matrix $\epsilon^2\mathbf{R}$
ϕ	Standard deviation of the prior
ϕ_i	Standard deviation of the prior proportional to the δ_i

Mathematics

$Cov(X_i, X_j)$	Covariance of random variables X_i and X_j
$E[X]$	Expected value of the random variable, X
$\sigma[X]$	Standard deviation of a random variable X
f_{X_i}	The pdf of a random variable, X_i
$f(u)$	Prior model in BMC
$G(\vec{x})$	Implicit or explicit limit state equation
$E[X_i\|X_j]$	Expected value of a variable X_i given X_j
$P(X_i > x_i)$	Probability of getting a random value greater than x_i from f_{X_i}
$CR(X_i, X_j)$	Correlation Ratio of of random variables X_i and X_j
$\rho(X_i, X_j)$	Product moment correlation of random variables X_i and X_j
$\rho_r(X_i, X_j)$	Rank Correlation of of random variables X_i and X_j
$\vec{X}$	A vector of random variables: $(X_1, X_2, \cdots, X_n)$
$\vec{x}$	A vector of scalers $(x_1, x_2, \cdots, x_n)$
$V(\hat{p}_\mathrm{f})$	The coefficient of variation of $\hat{p}_\mathrm{f}$
$\propto$	The proportional sign

Acknowledgements

I would like to acknowledge people who helped me to get this research done. However, I can name a few of them here. First, I acknowledge the promotor and co-promotor of this research; many thanks to Prof. *Vrijling* and Dr. *van Gelder*. More importantly, I would like to acknowledge the sincere help and support of my parents, *Ali* and *Ghadamkheir*, then my family. I certainly appreciate the *angelic patience* of my son, *Hossein*, who was growing up apart from his father. I also appreciate Mr. *Duivendijk* who suggested my promoter in TUDelft. This opportunity may have not achieved without the help and encouragement of my brother, *Abdollah*. I also appreciate all the people who directly and indirectly were involved in this research and also my colleagues: *Noel van Erp* for introducing me to Bayesian statistics, *Mark Voorendt*, *Chantal van Woggelum*, *Judith Schooneveld* and *Inge van Roij*. I should also appreciate the people who financially supported this research in the frame work of the *DelftCluster* project and *FloodSite* project.

M. Rajabalinejad
Delft, 1 April 2009

Index