Deltares

Enabling Delta Life

Reliability updating for slope stability of dikes

Approach with fragility curves (background report)

Deltares

Reliability updating for slope stability of dikes

Approach with fragility curves (background report)

Dr. ir. Timo Schweckendiek
Dr. ir. Wim Kanning

with contributions from:
Drs. Rob Brinkman
Ir. Wouter-Jan Klerk
Ir. Mark van der Krogt
Ir. Katerina Rippi
Dr. Ana Teixeira

1230090-033

Deltares

Title
Reliability updating for slope stability of dikes

Client	Project	Reference	Pages
Rijkswaterstaat WVL	1230090-033	1230090-033-GEO-0001	70

Classification
none

Keywords
dike, probability of failure, reliability updating, slope stability, past performance

Summary
Slope stability assessments of dikes, just like most geotechnical problems, are typically dominated by the large uncertainties in soil properties, often resulting in rather large estimated probabilities of (slope) failure compared to the actual failure rates observed in the field. Observations of past performance such as survival of significant loading can be incorporated to improve such reliability estimates. The present report describes how survival of observed load conditions can be taken into account quantitatively. In particular, a simplified approach with fragility curves is introduced in order to enable practitioners to adjust slope failure probabilities to field observations.
Besides a scientific description of the theory, the report contains considerations for applying the approach to dike stability problems. Moreover, it describes how the results can be used in the Dutch safety assessment framework for primary flood defenses. In essence, the updated probability of failure obtained from the analysis can be directly compared to the target probability of failure for a specific dike section as defined in the statutory safety assessment of Dutch primary flood defenses (WBI-2017).
Four fictitious benchmark examples illustrate the application of the method to cases with a varying degree of complexity, each highlighting different aspects of the method. All examples confirm that the results obtained with the simplified approach with fragility curves reasonably match the results by conventional reliability methods such as Monte Carlo simulation.
The present work is part of a larger development effort by Rijkswaterstaat to enable practitioners to use reliability updating in Dutch safety assessment practice. Besides this background report, case studies will be elaborated and a manual as well as software will be produced to provide practitioners with all necessary components. The application to the first two real life case studies with a realistic level of detail and complexity for a safety assessment of Dutch dikes is described in a separate report (**?**).

References
see chapter "References"

Version	Date	Author	Initials	Review	Initials	Approval	Initials
01	31 May 2016	T. Schweckendiek		F. Diermanse		M. Sule	
02	26 Aug 2016	T. Schweckendiek		F. Diermanse		M. Sule	
03	22 Nov 2016	T. Schweckendiek		F. Diermanse	FD	M. Sule	

Status
FINAL

Contents

List of Figures

List of Tables

List of symbols

Latin symbols

E_i	discrete scenario
$f_X(x)$	probability density function of X, sometimes abbreviated as $f(x)$
F	failure (event, set)
$F_X(x)$	cumulative distribution function of X, sometimes abbreviated as $F(x)$
$g(\cdot)$	performance function
h	water level (load)
h_c	critical water level (resistance)
$h(\cdot)$	observation function
l_{eq}	equivalent auto-correlation length of a failure mode [m]
L	length of the dike reach [m]
m_d	model uncertainty
n	number of MCS-realizations
p_T	annual target probability of failure (specific loaction and failure mode)
$p_{T,mode}$	annual target probability of failure per reach for a specific failure mode
$p_{T,sys}$	total annual target probability of failure per reach (all failure modes)
$P(\cdot)$	probability operator
$\hat{P}(\cdot)$	probability estimator
R	resistance (random variable)
s	dominant load variable
S	load (random variable)
u	standard normal random variable
w	standard uniform random variable
$\mathbf{X}$	vector of random variables

Greek symbols

α	influence coefficient or importance factor (FORM)
β	reliability index
β_T	target reliability index
ε	evidence or observation (event, set)
$\Phi(\cdot)$	standard normal cumulative distribution function
ω	share of the failure mode of the total acceptable probability of failure
ρ	linear correlation coefficient

Abbreviations

CDF	cumulative distribution function
FC	fragility curves (approximation method)
FORM	First-order reliability method
MCS	Monte Carlo simulation (crude)
PDF	probability density function
RUPP	reliability updating with past performance
SF	stability factor

1 Introduction

1.1 Problem description and context

Slope stability assessments of dikes, just like most geotechnical problems, are typically dominated by the large uncertainties in soil properties. The estimated probabilities of (slope) failure are often rather large compared to the failure rates observed in the field, as experienced in the Dutch VNK2 project (for details refer to Rijkswaterstaat (2014), illustrated in Figure 1.1).

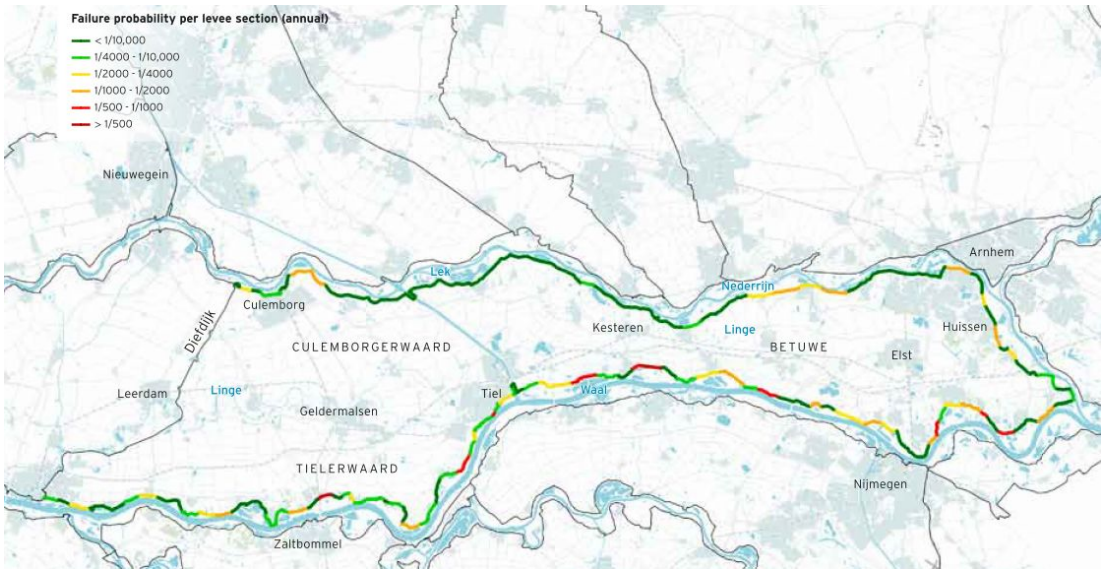


Figure 1.1: *Failure probabilities of the dike system Betuwe/Tieler- en Culemborgerwaarden according to Rijkswaterstaat (2014)*

Reliability analyses as carried out in the VNK2 project rely on physics-based limit state models and probabilistic models of the relevant random variables. The input to the analysis is typically based on site investigation data, laboratory testing and geological insights. Observations of past performance such as survival of significant loading are not incorporated in the assessments, while such information can reduce the uncertainties substantially and lead to more accurate safety assessments. Similar issues have been encountered in risk screenings of the federal levees in the U.S. and dealt with by using so-called likelihood ratios (Margo *et al.*, 2009), yet that approach is not easily incorporated in the Dutch approach with physics-based limit state models.

Rijkswaterstaat is conducting a project to operationalize the concept of Reliability Updating with Past Performance (RUPP; in Dutch often referred to as *bewezen sterkte*) for advanced safety assessments and reinforcement designs of the primary Dutch flood defenses. Reliability updating means to update our estimate of the probability of failure using observations of past performance, here specifically the survival of observed load conditions.

The focus in this first phase of the project is on the failure mode of instability of the inner slope, as many dikes were found not to meet the safety criteria for this failure mode in the statutory safety assessment of the Dutch primary flood defenses (IVW, 2011). The current work builds upon the concepts published in the *Technisch Rapport Actuele Sterkte bij Dijken* (ENW, 2009) and the work by Calle (2005), as well as more recently proposed approaches by Schweckendiek (2014), which have opened up new opportunities.

1.2 Objectives of the long-term development project

The **main objective** of the envisaged development efforts for the long-term project is to enable practitioners to work with use reliability updating in advanced safety assessments and reinforcement designs of the primary Dutch flood defenses. This implies the following sub-objectives:

1 to develop and document a scientifically sound and practicable approach,
2 to confirm and illustrate the practical applicability of the approach on test cases with a level of detail and complexity which is representative for real life conditions.

The long-term development project aims to deliver four main products:

1 **Background report** containing a scientifically sound description of the theory (current report),
2 **Case studies** for testing and illustrating the applicability,
3 **Manual** containing a description of the method and its application for practitioners,
4 **Software** facilitating (a) probabilistic slope stability analysis and (b) use of the RUPP method by practitioners.

These products are envisaged to complement and partially replace earlier guidance on reliability updating with past performance in the so-called TRAS (Dutch: Technisch Rapport Actuele Sterkte; ENW, 2009). The method described in the TRAS has shortcomings as demonstrated in the accompanying test case report (Schweckendiek *et al.*, 2016) and it hardly has been applied in practice (only one known case). Objectives of the current developments are also to overcome the shortcomings of the TRAS method and to provide more explicit guidance to enable and promote use of the approach in practice.

Note that this background report (1) and the accompanying test case report (2) are primarily aimed at an expert reader in order to assess the soundness of the approach and the envisaged application, while the manual (3) will mainly address a broader audience.

1.3 Objectives of this report and approach

The main objective is to describe a method which enables incorporating past performance information in reliability analysis for slope stability of dikes. In the present report we particularly focus on the failure mode 'slope instability', but the method is generic and also applicable to other failure modes. A basic requirement is that the end result is applicable in the Dutch safety assessment framework for flood defenses as described in Schweckendiek *et al.* (2012) and summarized in chapter 2.

1.3.1 Bayesian reliability updating

The basis for the proposed approach is Bayesian posterior analysis, in combination with reliability analysis often called "Bayesian reliability updating". Bayesian reliability updating can be implemented with most conventional reliability analysis methods used in the civil engineering domain such as (Crude) Monte Carlo simulation (MCS), Importance Sampling (IS), the First-order reliability method (FORM) or Numerical Integration (NI) (Straub, 2014). The descriptions in this report (chapter 3) will discuss MCS for illustration purposes, as the implementation of that method is rather straightforward.

The drawback of an implementation with MCS is the large number of required evaluations of the performance function (or limit state). The analysis can be intractable for computationally expensive models, as is the case for slope stability analysis. To this end, we will also de-

scribe an approximation method in chapter 4 using so-called fragility curves, which express the cumulative resistance to the outside water level against the dike.

As some uncertainties cannot be modeled with continuous probability distribution functions, the need for using discrete scenarios arises. Hence, we will also explain the implementation for discrete scenarios in chapter 5.

1.3.2 Application to slope instability

Conceptually, the application of the proposed approach to slope instability of dikes is straightforward. Yet, the underlying probabilistic analyses can be implemented in various ways. Chapter 6 provides an overview of the recommended implementation choices for the Dutch context and addresses modeling issues which require specific attention with past performance-based analyses. A basic choice in this present implementation is to work with 2D limit equilibrium models (LEM) such as Uplift-Van (see Figure 1.2), as these are commonly used for conditions with non-circular slip planes in the Netherlands.

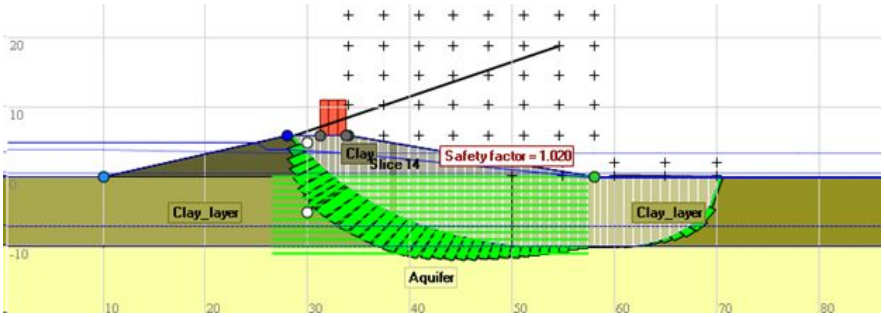


Figure 1.2: *Illustration of a critical slip plane with the Uplift-Van limit equilibrium method*

1.4 Visual outline

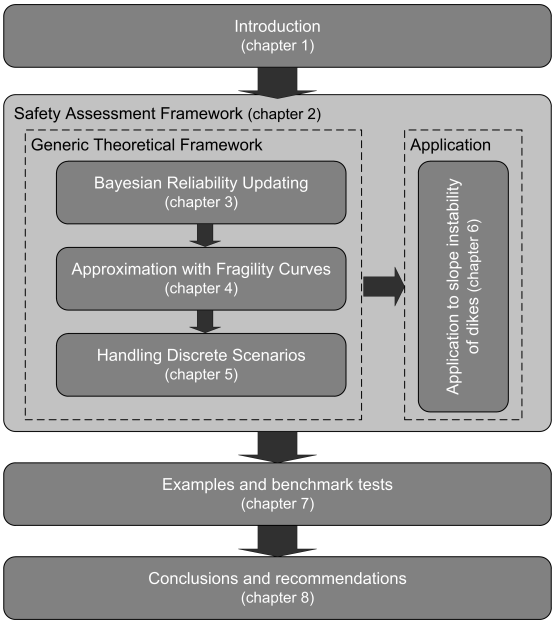


Figure 1.3: *Visual outline of the report*

2 Safety assessment

The presented work aims at application of reliability updating in the safety assessment framework which is envisaged to come into force in the Netherlands in 2017. The essence of the framework is that there is a risk-motivated and legally established acceptable probability of failure for a reach of the system of flood defenses (Deltaprogramma, 2014). Furthermore, Schweckendiek *et al.* (2012) outlines a procedure to determine the target reliability (i.e. acceptable probability of failure) for specific failure modes and dike segments, as described below.

2.1 Legal requirement for a dike reach

The project WV21 (see www.rijksoverheid.nl) investigated updating the safety standards by acceptable risk criteria based on individual risk, group risk and (societal) cost-benefit analysis. The information has served as input for a political decision on new safety standards. The resulting new safety standards for primary flood defenses in the Netherlands are specified in terms of acceptable (annual) probabilities of flooding as illustrated in Figure 2.1.

The project WBI-2017[1] is currently developing safety assessment methods for levees, dunes and hydraulic structures in flood defense systems in the Netherlands with semi-probabilistic as well as fully probabilistic methods and criteria. The proposed approach is appropriate for application to probabilistic assessments as envisaged in the WBI-2017 project for all safety assessments from 2017 onwards.

The basic safety requirement in the Netherlands will be an acceptable annual probability of failure $p_{T,sys}$ or, equivalently, annual target reliability $\beta_{T,sys}$ for a dike segment or flood defense (sub-)system. For practical reasons, these protection standards need to be translated into more specific requirements per levee reach and failure mode.

2.2 Target reliability per failure mode

Practically workable safety requirements are usually expressed per failure mechanism and per element (e.g., homogeneous dike reach) in terms of a specific target reliability (p_T or β_T). To derive such a specific target reliability we need to account for the different failure mechanisms involved as well as for the system reliability aspects such as the length-effect (see e.g. Kanning, 2012). The length-effect arises from the fact that all dike or levee reaches in the protection system contribute to the probability of (system) failure and that the probability of failure increases with the length of an element. Figure 2.2 depicts the conceptual cohesion of the safety framework.

The first step in deriving the specific target reliabilities is assigning target reliability values for each failure mode $\beta_{T,mode}$ for the whole protection system. The requirement in the WBI-2017 approach is that the sum of the target probabilities per failure mode should not exceed the target system probability of failure ($\sum p_{T,mode} < p_{T,sys}$), which is a conservative criterion because the implicit assumption is that the failure modes are mutually exclusive, whereas in practice positive correlation is often present (e.g. through common random variables). The default share of slope instability with respect to the total probability of failure is 4 %, meaning that $p_{T,inst} = 0.04 \cdot p_{T,sys}$.

[1] Information on the WBI-2017 project is provided on http://www.helpdeskwater.nl.

Figure 2.1: *Acceptable annual probabilities of failure for future safety assessments in the Netherlands from 2017 (Deltaprogramma, 2014). The warmer colors represent higher target reliabilities.*

2.3 Length-effect

The second step is to take the so-called length effect (see e.g. Kanning 2012) into account. To this end we use a failure mode-specific equivalent correlation length l_{eq} in deriving the "local" target reliability for slope instability using

$$p_T = \frac{p_{T,inst}}{1 + L/l_{eq}} \tag{2.1}$$

where L is the total length of considered reach (contributing to the probability of failure for instability). The basic theory behind this approach is based on considering the longitudinal variability of the dike as one-dimensional random field and determining the probability of exceeding the limit state as an outcrossing problem. A detailed description of the length-effect and its treatment is beyond the scope of this report; for details reference is made to Vanmarcke (2011), Kanning (2012) and Schweckendiek *et al.* (2012).

The reliability analyses in the proposed approach as discussed in chapter 6 will be based on simple random variables, not random fields. That means that they consider the uncertainty in the random variables as representative for the considered dike segment. Thus, the spatial

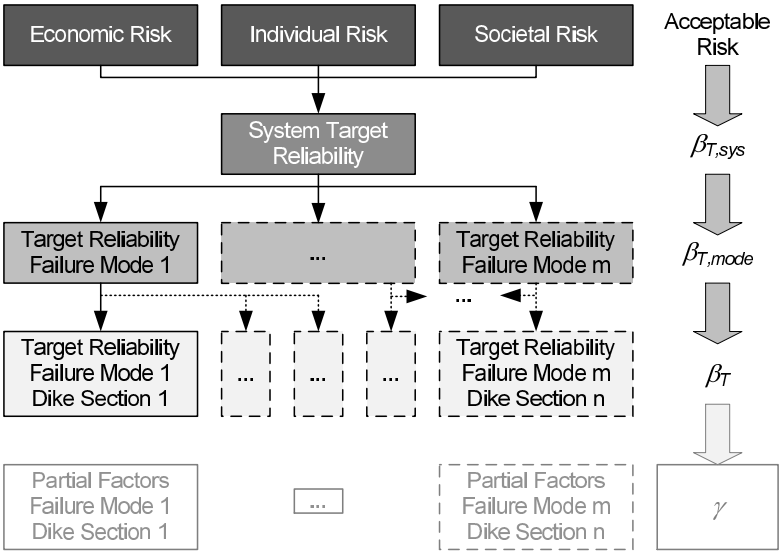


Figure 2.2: *Steps in deriving target reliabilities from acceptable risk criteria adopted from Schweckendiek et al. (2012)*

variability in longitudinal direction is not considered explicitly in the analysis but implicitly. It is noteworthy that this approach effectively means that the analyses works with infinite correlation lengths of the random variables in longitudinal direction and that the effects of spatial variability are considered through the reliability target. Note that the spatial variability in the cross-sectional dimension can be accounted for explicitly, for example by spatial averaging where necessary.

In summary, the target reliability for slope instability of dikes in the proposed approach is determined by:

$$p_T = \frac{\omega \cdot p_{T,sys}}{1 + L/l_{eq}} \tag{2.2}$$

where

ω	share of the failure mode of the total acceptable probability of failure
$p_{T,sys}$	total acceptable annual probability of failure per reach (all failure modes; a reach is typically tens of
L	length of the reach (restricted to the portionof the reach with potential contribution to the probability
l_{eq}	equivalent auto-correlation length [m]

While values for $p_{T,sys}$ will be legally established per reach, appropriate values for the other parameters (L, l_{eq} and ω) will be proposed by the WBI-2017 project or can be substantiated with local data.

The essential implicit assumption in using the target reliability β_T as assessment criterion not only for the prior (i.e. conventional) reliability estimate but also for the posterior (i.e. after reliability updating) is that the length effect does not significantly increase through the posterior analysis relatively speaking. In other words, the equivalent correlation length (l_{eq}) of the considered failure mode does not decrease. Or similarly the ratio of the probability of failure per reach and per segment (i.e. $p_{T,mode}/p_T$) does not increase.

Recent findings by Roscoe *et al.* (2016) support this assumption; in their study the relative length-effect decreases after updating for all contemplated examples. Furthermore, appendix A contains a sensitivity analysis with one-dimensional random fields to examine the change of length effect with reliability updating. The results confirm that mostly the posterior

length-effect is less than or roughly equal to the prior length effect. Only for rather high standard deviations of the resistance we have seen an increase of the length-effect by a factor two (which is not much in terms of probability).

Our recommendation from the information at hand is to stick to the cross-sectional target reliability p_T as formulated above also for updated probabilities of failure. Even though a slight increase of the length effect can occur, this is rather unlikely or rare. We need to bear in mind that the default parameters to account for the length effect in WBI-2017 were chosen conservatively and contain some margin.

3 Reliability updating

This chapter contains the definitions and descriptions of the methods used for reliability analysis and reliability updating in the proposed approach. The finally obtained probabilities of failure or reliability indices can be directly assessed by comparing with the target probabilities of failure p_T and target reliability indices β_T as discussed in chapter 2.

3.1 Reliability analysis (prior analysis)

3.1.1 Failure (undesired event)

Failure refers to an undesired event, not necessarily to the collapse of a structure, which we model by means of a (continuous) performance function $g(\mathbf{X})$ such that the performance function assuming negative values represents the failure domain F:

$$F = \{g(\mathbf{X}) < 0\} \tag{3.1}$$

where $\mathbf{X}$ is the vector of random variables. In the specific case of slope stability, the result of a limit equilibrium analysis is typically a stability factor SF, in which case the performance function can be expressed as $g = SF - 1$ (possibly complemented with a model factor accounting for the uncertainty in the limit equilibrium model), because stability factors are defined such that values below one imply failure (typically based on moment equilibrium).

3.1.2 Probability of failure

Using the definitions of failure and of the performance function, the probability of failure (i.e., unwanted event) is given by:

$$P(F) = P(g(\mathbf{X}) < 0) = \int_{g(\mathbf{X})<0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \tag{3.2}$$

where $f_{\mathbf{X}}(\mathbf{x})$ is the joint probability density function (PDF) of $\mathbf{X}$.

3.1.3 Fragility curves

Fragility curves represent the probability of failure conditioned on a dominant load variable s:

$$P(F|s) = \int_{g(\mathbf{R},s)<0} f_{\mathbf{R}}(\mathbf{r}) d\mathbf{r} \tag{3.3}$$

where $\mathbf{R}$ is the vector of all other random variables except s. That implies that a fragility curve is in fact equivalent to the cumulative distribution function (CDF) of the overall resistance R:

$$P(F|s) = F_R(s) \tag{3.4}$$

where F_R is the CDF[1] of R. Modelling fragility as a CDF is a common assumption in other fields too, such as seismic risk analysis.

[1] Formally speaking, there may be conditions where the fragility curve does not reach one for increasing load levels, in which case the fragility curve is not a proper CDF. In the envisaged area of application, this formal restriction is hardly ever relevant.

3.2 Reliability updating (posterior analysis)

Posterior analysis, also called "Bayesian Updating", is the essential ingredient of reliability updating. The description in this section is restricted to the so-called "direct method" in combination with "inequality information". For a more general treatment refer to, for example, Straub (2014) or Schweckendiek (2014).

3.2.1 Direct reliability updating

Bayes' Rule (Bayes, 1763) forms the basis for updating (failure) probabilities with new evidence:

$$P(F|\varepsilon) = \frac{P(F \cap \varepsilon)}{P(\varepsilon)} = \frac{P(\varepsilon|F)P(F)}{P(\varepsilon)} \tag{3.5}$$

where F is the failure event to be estimated and ε the observed event or evidence.

The *indirect method* entails updating the probability distributions of the basic random variables first and using the updated distributions in a reliability analysis. On the other hand, the *direct method* for reliability updating exploits the definition of the conditional probability of failure, $P(F|\varepsilon) = P(F \cap \varepsilon)/P(\varepsilon)$, by defining a new limit state of the intersection (cut set) of failure and the observation ($F \cap \varepsilon$). While the direct and indirect updating are mathematically equivalent, the direct method is easier to implement, especially with sampling type of reliability methods such as Monte Carlo simulation. In this report all descriptions are restricted to the direct method.

Also using the direct method the updated joint probability distribution of the basic random variables can be inferred (depending on the reliability method used), which is very useful for interpretation of the results. Note that only the complete joint probability distribution can be used for further analysis, as the updating process can change the correlation structure. Even though a-priori the basic random variables are uncorrelated, there may be correlation a-posteriori. The updated marginal distributions can still be useful for illustration purposes.

3.2.2 Inequality information

There are two types of information that are distinguished mainly due to the difference in implementation for reliability updating, equality and inequality information. For the present scope, we only deal with inequality information. When the evidence implies that our observed quantity is greater than or less than some function of the random variables of interest, the evidence ε can be formulated as:

$$\varepsilon \equiv \{h(\mathbf{x}) < 0)\} \tag{3.6}$$

where $h(\cdot)$ is the observation function. Typical examples of inequality information are failure (i.e., exceedance of a limit state), survival or the loads reached in incomplete load tests. Consequently, the posterior probability of failure can be found as follows:

$$P(F|\epsilon) = \frac{P\left(\{g(\mathbf{X}) < 0\} \cap \{h(\mathbf{X}) < 0\}\right)}{P(h(\mathbf{X}) < 0)} \tag{3.7}$$

Note that for multiple observations, the total evidence is the intersection of the individual observations (i.e. their outcome spaces):

$$\varepsilon \equiv \bigcap_k \{h_k(\mathbf{x}) < 0\} \tag{3.8}$$

Effectively, equation 3.8 implies that if we have multiple observations, after updating we only consider the parameter space which is still plausible after accounting for all the individual pieces of evidence.

3.3 Reducibility of uncertainties and auto-correlation in time

For the engineering purposes at hand, we define uncertainties as reducible if it is feasible to acquire, interpret and incorporate information that has a significant impact on the magnitude of uncertainty. Such uncertainty is commonly called epistemic uncertainty (or knowledge uncertainty). Typical examples of epistemic uncertainties related to the contemplated problems are the probability distributions of soil strength properties or uncertainties in stratification, including so-called anomalies or adverse geological details. These properties or features are time-invariant, at least on an engineering time scale.

On the other hand, the uncertainty in an annual maximum river water level at a certain location is practically irreducible. We commonly call this type of uncertainty aleatory uncertainty or intrinsic variability. While it is true that each year we obtain new evidence, because each year a new maximum level is realized, such information usually does not change the probability distribution significantly, of course depending on the amount of data already in the data set (i.e. statistical uncertainty).

In conclusion, reducibility can be considered as a matter of correlation in time. We can only reduce uncertainties (i.e. learn) of random variables which we assume to be time-invariant and, hence, epistemic. In other words, we assume them to be the same at the time of the observation as for the future event to be estimated.

If however, a situation is dominated by aleatory variability (in time), we can hardly "learn" from an observation. The effect of updating will be insignificant, as in this case the observation and the predicted event are statistically independent in time (i.e., zero auto-correlation), practically speaking.

For dike instability most soil properties and geohydrological parameters can be assumed time-invariant with epistemic uncertainty, whereas most (external) loads such as the water level, the phreatic level or traffic loads are typically classified as aleatory. For a detailed list of random variables refer to section 6.4.

Model uncertainty typically has contributions of both reducible and of irreducible nature (Schweckendiek, 2014). Arguably, for physics-based performance functions in which we model most random load conditions explicitly, the model error covers local bias, i.e. systematic over- or under-predictions of the model in terms of performance at the location in question. Hence, the model error can be assumed time-invariant and reducible.

3.4 Implementation with sampling methods

The direct reliability updating method can be used with virtually any standard reliability method (Straub and Papaioannou, 2014). For the sake of illustration, this section describes a straightforward implementation with Crude Monte Carlo simulation.

A pragmatic approach to deal with the issue of auto-correlation in time of the individual random variables, i.e. if they are epistemic and fully redcucible in terms of uncertainty or aleatory and irreducible, is to define two categories of random variables: epistemic and aleatory. In reality, most random variables will represent contributions of both, epistemic and aleatory uncertainty, yet often one of the two is clearly dominant. Furthermore, we use two sets of random variables, $\mathbf{X}^p$ and $\mathbf{X}^f$, where p stands for the (past) observed event and f for the (future) event to be predicted. Both types of random variables (epistemic and aleatory) are included in $\mathbf{X}^p$ and $\mathbf{X}^f$, but are treated differently, as explained below.

The steps below describe a prior and subsequent posterior analysis using these definitions with Crude Monte Carlo Simulation (MCS):

1 **Simulation of the event to be predicted** Generate n realizations of the basic random variables according to their (prior) joint probability distribution. The j-th realization of the i-th random variable is denoted as X_{ij}^f and the j-th realization of the vector of basic random variables is denoted as $\mathbf{X}_j^f$.

2 **Prior probability of failure** The prior probability of failure is the number of realizations in which the performance function assumes a negative value ($\mathbf{1}[\,\cdot\,]$ is the indicator function), divided by n:

$$\hat{P}(F) = \frac{1}{n}\sum_j \mathbf{1}[g(\mathbf{X_j^f} < 0)] \tag{3.9}$$

3 **Simulation of the observed conditions** The realizations of all variables with (fully) reducible uncertainty obtain the same value as the event to be predicted (full auto-correlation in time or time-invariance):

$$X_{ij}^p = X_{ij}^f \tag{3.10}$$

for all i where the uncertainty is assumed reducible. The random variables assumed to be intrinsically variable obtain new independent realizations (no auto-correlation in time) according to their (joint) probability distribution.

4 **Posterior probability of failure** The updating is achieved by conditioning on the observation (in general form $\varepsilon_k = \{h_k(\mathbf{X^p}) < 0\}$,) and evaluating the following term:

$$\hat{P}(F|\varepsilon) = \frac{\sum_j \left(\mathbf{1}[g(\mathbf{X_j^f} < 0)] \cdot \prod_k \mathbf{1}[h_k(\mathbf{X_j^{p,k}}) < 0]\right)}{\sum_j \prod_k \mathbf{1}[h_k(\mathbf{X_j^{p,k}}) < 0]} \tag{3.11}$$

The term $\prod_k \mathbf{1}[h_k(\mathbf{X_j^{p,k}}) < 0]$ implies that an observation can imply several limit states and/or several observations. Note that if observations were made at different points in time, the independent realizations of the random variables $\mathbf{X^{p,k}}$ are required (i.e. the random variables representing aleatory uncertainty).

The implementation with computationally more efficient reliability methods such as Importance Sampling, Directional Sampling or Subset Simulation is rather straightforward and essentially requires solving equation 3.7, in which the numerator represents a combined limit state of a parallel system.

4 Approximation using fragility curves

4.1 Problem description and objective

For application with computationally expensive performance functions, such as stability analyses, the reliability updating approach may not be tractable with Crude Monte Carlo (or other sampling-based techniques) in terms of computation times. For example;

- Suppose one evaluation of the performance function takes *1 second*,
- and the performance function needs to be evaluated *1 million times*,
- the total computation time amounts roughly 278 hours or almost to *12 days*.

This is typically not feasible or acceptable in engineering projects. For high reliability requirements such as for Dutch flood defenses, the required number of computations can even be orders of magnitude higher.

Below we describe an approximation method using fragility curves inspired by the experience with probabilistic stability analyses in the Dutch VNK2-project, which requires significantly less computation time. Furthermore, the proposed approach, in which the fragility curves can be derived with several FORM analyses, can provide very insightful results which allow easier interpretation of the results in terms of sanity checks by practitioners in our experience.

4.2 Fragility curves

Fragility curves are functions describing the conditional probability of failure given a (dominant) load variable (see 3.1.3). For dikes, typically the (water-side) water level h is used as the load of reference:

$$P(F|h) = P\left(g(\mathbf{X}, h) < 0\right) \tag{4.1}$$

in which case $\mathbf{X}$ becomes the vector of all random variables except for h. In other words, for dikes a fragility curve quantifies the probability of failure of the dike, conditional on the occurrence of a given water level, typically but not necessarily assuming a steady state pore pressure response implying a long duration load condition (at least for slope stability analyses).

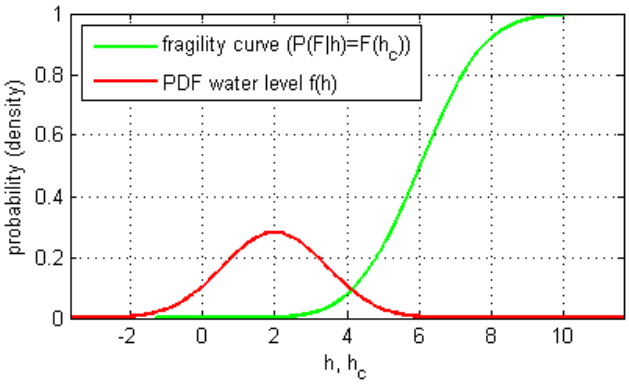


Figure 4.1: *Example fragility curve (see sec for details)*

While the following elaboration focuses on the water level h to be used in fragility curves, it is important to realize that we can use any other load variable instead.

The definition of fragility curves implies that the curve at the same time represents the cumulative distribution function (CDF) F_{h_c} of the critical water level h_c, which is the water level at which the dike fails[1]. This can be illustrated by defining the performance function as $g = h_c - h$, for which case the probability of failure is given by:

$$\begin{aligned} P(F) = P(h_c < h) &= \iint_{h_c<h} f(h_c)f(h)\, dh_c\, dh \\ &= \int F_{h_c}(h)f(h)\, dh \\ &= \int P(F|h)f(h)\, dh \end{aligned} \tag{4.2}$$

The fact that fragility curves represent the probability distribution of the overall resistance (quantified as the 'critical water level') is the key concept used in the approximate approach described in the remainder of this document. Before elaborating how reliability updating works with fragility curves in section 4.4, section 4.3 explains how we can derive fragility curves and how we can sample from them.

4.3 Beta-h curves and critical water level

In reliability analysis for slope stability of dikes it is common practice in the Netherlands to first estimate the probability of failure conditional to several water levels in a relevant range using the First-Order Reliability Method (FORM). The results are represented as beta-h curves as depicted in Figure 4.2, which is just another representation of a fragility curve with the reliability index β on the vertical axis instead of the probability of failure. The "fragility points" are the result of the reliability analyses per water level. The red lines in Figure 4.2 indicate that we assume that the conditional reliability for other water levels than in the fragility points can be reasonably approximated by linear interpolation between the fragility points (in beta-h space). Note that the fragility points can in principle be determined using any other reliability method, not necessarily FORM.

As pointed out in section 4.2, such a beta-h curve represents the CDF of the overall resistance term, in our applications typically the critical water level h_c. Note that h_c is a random variable representing all random variables except h, some of which may be labeled load or load effect variables and not resistance in the conventional sense. For practical implementation, we can define the random variable h_c in terms of the corresponding beta-h curve as follows.

Let the function G be defined as the linear interpolation (and extrapolation) of the conditional reliability index based on the neighboring fragility points (β_1,h_1) and (β_2, h_2):

$$\beta = G(h) = \beta_1 + (\beta_2 - \beta_1)\frac{h - h_1}{h_2 - h_1} \tag{4.3}$$

The critical water level, i.e. water level at which the dike will fail, can now be modeled as a random variable in the following way:

$$h_c = G^{-1}(u) \tag{4.4}$$

[1]Note that fragility curves do not always strictly meet the required properties for being a CDF. For example, there are (rare) cases where increasing the load does not lead to a probability of failure of one (i.e. $\lim_{x\to+\infty} F(h_c) < 1$.). We assume here that these formal issues do not matter for the envisaged practical civil engineering applications.

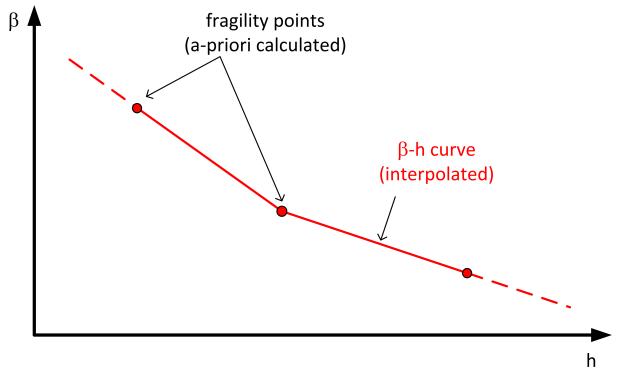


Figure 4.2: *Beta-h curve: The fragility points represent the reliability indices corresponding to the conditional probabilities of failure derived for discrete water levels. The conditional reliability for other water levels is obtained by linear interpolation.*

where u is the realization of a standard normal random variable and G^{-1} is the inverse interpolation of the beta-h curve (i.e. interpolate h_c from given u or β). The definition in standard normal space is particularly useful, as many implementations of reliability analysis work from standard normal space before transforming to real space.

Realizations of h_c can be generated by transforming a standard normal distributed sample using the inverse (interpolated) beta-h curve[2]: $h_{c,i} = G^{-1}(u_i)$ (see Figure 4.3).

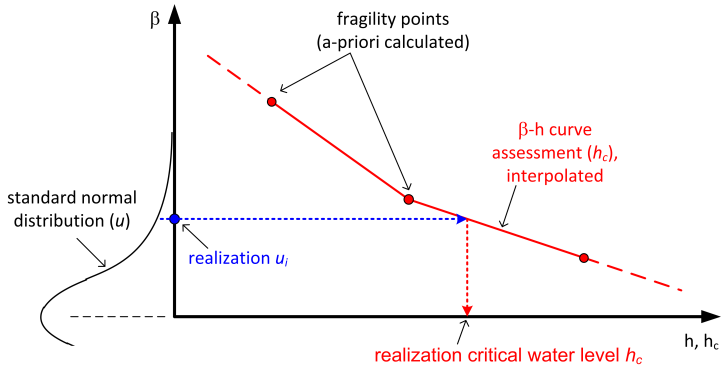


Figure 4.3: *Illustration of sampling realizations of the critical water level directly from (linearly) interpolated beta-h curves*

In the Dutch experience with slope reliability analysis for dikes, linear interpolation in beta space (with sufficient and sensibly located fragility points) is a very reasonable approximation of the exact distribution (see examples in chapter 7), which is the main reason to work with beta-h curves instead of interpolating in probability space.

In summary, beta-h curves can be generated using reliability analyses for discrete water levels, for example with FORM, and allow sampling of the critical water level directly without requiring additional computationally expensive model simulations.

[2]Note that this definition is similar to the relation often exploited for sampling non-uniform random variables: Transforming realizations of a random variable with its CDF leads to a uniformly distributed sample. Inversely, transforming uniformly distributed realizations w_i with an inverse CDF F_X^{-1} leads to a sample with the distribution of X: $x_i = F_X^{-1}(w_i)$.

4.4 Reliability updating with fragility curves

Being able to define the random variable of the critical water level based on beta-h curves as discussed in section 4.3, we can apply the reliability updating approach discussed in section 3.2 directly. To that end we again define the performance function $g = h_c - h$, where h_c is the critical water level and h is the water level, both for the (future) conditions to be assessed, implying that failure is defined as the water level exceeding the critical water level:

$$F = \{g < 0\} = \{h_c < h\} \tag{4.5}$$

Furthermore, we define the observation or evidence (ε) as the critical water level at the observation $h_{c,obs}$ exceeding the water level at the observation h_{obs} (can also be a random variable due to measurement uncertainty etc.):

$$\varepsilon = \{h_{c,obs} > h_{obs}\} \tag{4.6}$$

Note that the conditions at the time of the observation may differ from the assessment conditions, in which case it is necessary to derive a separate beta-h curve for the observation. There are many potential reasons for such differences such as subsidence, degradation or human interventions.

Having defined failure under assessment conditions (F) and the evidence in terms of survival of the observed conditions (ε), the basic formulation of reliability updating with fragility curves directly follows from Equation 3.5:

$$P(F|\varepsilon) = \frac{P(F \cap \varepsilon)}{P(\varepsilon)} = \frac{P(\{h_c < h\} \cap \{h_{c,obs} > h_{obs}\})}{P(\{h_{c,obs} > h_{obs}\})} \tag{4.7}$$

As Straub and Papaioannou (2014) have illustrated, Equation 4.7 can be solved by standard reliability methods. The numerator represents a parallel system reliability problem of two limit states, whereas the denominator is a classical component reliability problem.

Note that reliability updating will only have an effect if the resistance of the dike in the future (h_c) is correlated with the resistance at the time of the observation ($h_{c,obs}$), as will be discussed in the subsequent section 4.5.

4.5 Correlation between assessment and observation

As discussed in section 3.3, we can only reduce the epistemic (knowledge) uncertainty, while aleatory uncertainty will persist. The proposed pragmatic approach is to divide the random variables in two categories in terms of the uncertainty they represent:

1 epistemic, reducible uncertainty (i.e. time invariant)
2 aleatory, irreducible uncertainty (i.e. intrinsically variable)

In reality, most random variables will represent contributions of both epistemic and aleatory uncertainty, yet mostly one of the two is clearly dominant.

We can use the information of auto-correlation in time of the individual basic random variables to estimate the correlation between the dike resistance in the assessment conditions (h_c) and at the time of the observation ($h_{c,obs}$) using the influence coefficients (α) from deriving the fragility curves. According to Vrouwenvelder (2006), the (linear) correlation coefficient ρ between the two resistance terms can be approximated by :

$$\rho \approx \sum_i \alpha_i^p \alpha_i^f \rho_i^{p,f} \tag{4.8}$$

where α_i^p and α_i^f are the FORM influence coefficients (also attainable from other reliability methods) of variable i for the observation (p for past) and for the assessment conditions (f for future), respectively. The correlation coefficient $\rho_i^{p,f}$ describes the correlation of variable i between the observation and the assessment, thus effectively the auto-correlation in time for individual variables. As discussed, we would assume either time-invariance ($\rho_i^{p,f} = 1$) or no correlation at all ($\rho_i^{p,f} = 0$) for each basic random variable. But of course, better estimates can be used if available.

In the approach with fragility curves, the influence coefficients α_i are obtained in the fragility points. The α_i can differ between the fragility points. There are essentially two practical approaches to deal with this issue in estimating the correlation coefficient:

1 by averaging over all fragility points on which the fragility curve is based, or
2 by interpolation of the α_i in the design point of the fragility curve, similar to the interpolation of the conditional reliability index β.

Both options can be used with most standard implementations of reliability methods. Averaging (option 1) works fine for near-linear fragility curves, yet has disadvantages for strongly non-linear of the fragility, in which case typically also the influence coefficients change significantly. An example of such behavior be a sudden drop of the fragility curve a higher water levels due to infiltration of overtopping water and subsequent saturation of the inner slope. When such phenomena play a dominant role, the critical sliding surface change with the water level and so do the influence coefficients. In such conditions interpolation in the design point (option 2) determines ρ in the most relevant area or point of the resistance distribution and provides more accurate results than averaging. Note that for all options it is required to normalize the influence coefficients after averaging or interpolation (i.e. $\alpha_{i,normalized}^2 = \alpha_i^2 / \sum \alpha_i^2$). Interpolating the squared influence coefficients (respecting the sign) removes the need for normalization.

In section 7.4 we illustrate the sensitivity of the posterior reliability to the correlation between assessment and observation using option 2 (interpolation in the design point). Some tests (not reported here) have given us confidence that option 2 is sufficiently accurate in comparison with a full Monte Carlo analysis. Nevertheless, it is recommended to keep on benchmarking the results of approximations with alternative reliability methods which do not require these approximations.

4.6 Implementation with Monte Carlo simulation

In this section we illustrate the implementation of reliability updating with fragility curves as described in section 4.4 using Crude Monte Carlo simulation (MCS) for the special case of full correlation between the resistance in assessment and observation conditions, as well as for partial correlation (i.e. the general case).

Suppose we have derived fragility curves for the assessment conditions as well as for the conditions at the time of the observed survival, allowing us to produce samples of the corresponding critical water levels with realizations $h_{c,i}$ and $h_{c,obs,i}$ respectively. The posterior probability of failure can now be estimated as described in section 3.4:

$$\hat{P}(F|\varepsilon) = \frac{\sum_i \left(\mathbf{1}[h_{c,i} < h_i] \cdot \mathbf{1}[h_{c,obs,i} > h_{obs,i}]\right)}{\sum_i \mathbf{1}[h_{c,obs,i} > h_{obs,i}]} \tag{4.9}$$

In other words, we essentially count the realizations for which both failure and the observation are true and divide by the number of realizations which comply with the observation. Incorpo-

rating the observation in this fashion works as a filter in the Monte Carlo simulation removing implausible realizations (i.e. which do not match the observation).

For the special case that the critical water levels h_c and $h_{c,obs}$ are fully correlated (i.e. $\rho = 1$; see section 4.5), the sampling scheme as described in section 4.3 can be adopted by using the same realization of the standard normal variable u to interpolate both realizations from the fragility curves for h_c and $h_{c,obs}$ (see Figure 4.4).

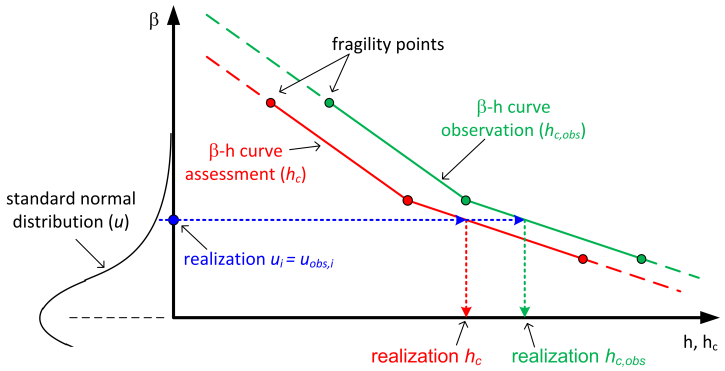


Figure 4.4: *Illustration of sampling realizations of the critical water level directly from (linearly) interpolated beta-h curves for both, the assessment and the observation in case of full auto-correlation in time (special case)*

For the general case, where the correlation between assessment and observation resistance is smaller than 1, we can generate correlated realizations of standard normal variables u and u_{obs} with correlation ρ (e.g. by using the inverse of the bi-variate normal distribution function with correlation ρ) as depicted in Figure 4.5.

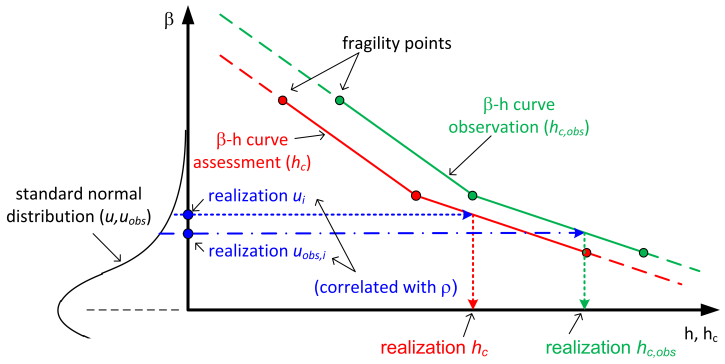


Figure 4.5: *Illustration of sampling realizations of the critical water level from (linearly) interpolated beta-h curves for both, the assessment and the observation in case of partial auto-correlation in time (general case)*

Practical aspects for generating fragility curves for the assessment and observation conditions distinctly for a slope instability problem will be discussed in chapter 6.

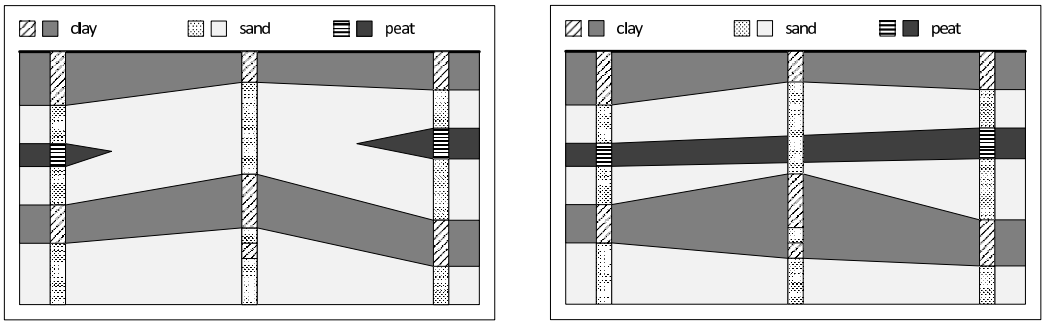
5 Handling discrete scenarios

This chapter contains a description of how to handle discrete scenarios in the reliability updating method described hitherto. Special attention is paid to the auto-correlation in time of scenarios.

5.1 Why discrete scenarios?

The need for defining discrete scenarios typically arises from the inability to capture some uncertainties in the continuous stochastic input variables of the model. The two most relevant types of scenarios in a dike stability context are:

1 **Stratification:** Due to limited site investigation the precise composition of subsoil layers in a given dike section may be uncertain. Such uncertainty can relate to the presence of specific soil strata (see Figure 5.1) or local features such as clay lenses.
2 **Geohydrological response:** In the common tools used for slope stability analysis in the Netherlands, the geohydrological response in the dike cannot entirely be modeled stochastically, practically speaking (Kanning and Van der Krogt, 2016). This is especially true for the phreatic surface.



(a) Interpretation 1 (b) Interpretation 2

Figure 5.1: *Illustration of different stratification scenarios inferred from the same borings, adopted from Schweckendiek (2014).*

When defining discrete scenarios is unavoidable, the reliability (updating) analyses can be carried out conditional on the scenarios (i.e, for each scenario individually).

5.2 Prior analysis with scenarios

As discussed in section 5.1, some ground-related as well as geo-hydrological uncertainties will be modeled as (discrete) subsoil scenarios E_i. The total probability of failure over all (mutually exclusive and collectively exhaustive: $\sum P(E_i) = 1$) scenarios is given by the law of total probability:

$$P(F) = \sum_i P(F|E_i)P(E_i) \tag{5.1}$$

Similarly, the combined fragility curve over all scenarios can be determined by $P(F|h) = \sum_i P(F|h, E_i)P(E_i)$. In contrast to prior analysis, the distinction between time-invariant (epistemic) conditions represented by scenarios and conditions that are intrinsically variable in time (aleatory) will matter in the posterior analysis as discussed in the section below.

5.3 Posterior analysis with scenarios

Section 3.2 described reliability updating for continuous probability distributions, which are the most common representations for uncertainty in physical quantities involved in failure and observation (limit state) functions. This section describes the posterior analysis for situations with discrete scenarios or discrete probability mass functions (PMF), which can be also combined with the fragility curves approach as described in chapter 4.

5.3.1 Probabilities of observation and assessment scenarios

The general idea behind defining scenarios is that we have a set of conditions the dike strength depends on, and that some of these conditions are time-invariant and others are variable in time. For example, the dike strength depends on the stratification of the subsoil under the dike (see Figure 5.1), which is often uncertain. The subsoil composition typically does not change between observation and assessment (i.e. is time-invariant) and, hence the uncertainty is of epistemic nature. On the other hand, we may define discrete scenarios of the phreatic surface's response to external forcings, for example due to practical problems with modeling the response with continuous probability distributions. The phreatic surface response typically depends on more factors than just the dike composition such as rainfall before and during the high water event. Hence, the uncertainty represented is of aleatory nature.

In order to use a similar probabilistic framework for discrete scenarios as for continuous random variables, we will use the following definitions:

- E_i is the event that scenario i is true in the assessment conditions with associated probability $P(E_i)$,
- $E_{obs,j}$ is the event that scenario j is or was true in the observation conditions with associated probability $P(E_{obs,j})$,
- $P(E_i|E_{obs,j})$ is the conditional probability that scenario i is true in the assessment conditions, given scenario j is or was true in the observation conditions.

Note that for the time-invariant or epistemic type of scenarios

- $P(E_i|E_{obs,j}) = 1$ for $i = j$ and
- $P(E_i|E_{obs,j}) = 0$ for $i \neq j$,

as the conditions do not change between assessment and observation. If a scenario was true in the past it will be true in the future, and no other scenario can be true. Likewise, for aleatory type of scenarios holds $P(E_i|E_{obs,j}) = P(E_i)$, as the observation conditions do not give information regarding the assessment conditions.

Essentially all possible combinations of assessment and observations scenarios with associated probabilities need to be considered for the overall updated probability of failure, as elaborated in section 5.3.2 below. A requirement by the total probability theorem is that the whole set needs to be mutually exclusive and exhaustive (i.e. $\sum P(E_i \cap E_{obs,j}) = 1$).

5.3.2 Updating failure probabilities with discrete scenarios

As we have seen in section 3.2, the general formulation for reliability updating with inequality information can be written as $P(F|\varepsilon) = P(F \cap \varepsilon)/P(\varepsilon)$. In order to relate the probabilities defined in section 5.3.1, we define the following short-hand notation:

- $P(F_i) = P(\{g(\mathbf{X}|E_i) < 0\})$ is the probability of failure in the assessment conditions given scenario i is true,
- $P(\varepsilon_j) = P(\{h(\mathbf{X}|E_{obs,j}) < 0\})$ is the probability of the observation being true given scenario j is or was true,
- $P(F_i \cap \varepsilon_j) = P(\{F \cap \varepsilon\}|E_i, E_{obs,j})$ is the probability that both failure in the assessment conditions and the observation are true, given scenario i is true for the assessment and scenario j is true for the observation.

The updated or posterior probability of failure for a given combination of assessment and observation scenario is then given by:

$$P(F_i|\varepsilon_j) = \frac{P(F_i \cap \varepsilon_j)}{P(\varepsilon_j)} = \frac{P(\{F \cap \varepsilon\}|E_i, E_{obs,j})}{P(\varepsilon|E_{obs,j})} = P(F|\varepsilon, E_i, E_{obs,j}) \tag{5.2}$$

This implies that an observed survival during an event with an observed scenario j, gives information about the future failure probability in case scenario i occurs. The posterior probability of failure can then be obtained from the weighted sum of the individual conditional posterior probabilities:

$$P(F|\varepsilon) = \sum_i \sum_j P(F_i|\varepsilon_j) P(E_i \cap E_{obs,j}) \tag{5.3}$$

where the summation goes over all possible combinations of i and j and with

$$P(E_i \cap E_{obs,j}) = P(E_i|E_{obs,j}) P(E_{obs,j}) \tag{5.4}$$

Note that some of these combinations will be irrelevant and do not need to be analyses because their probability can be zero, as explained in section 5.3.1.

5.3.3 Implementation options

The two general options to deal with discrete scenarios computationally are:

1 **Two-stage procedure:** In the two-stage procedure we first evaluate all possible (and relevant) combinations of observation and assessment scenarios individually to obtain the conditional posterior probabilities of failure $P(F_i|\varepsilon_j)$ (Eq. 5.2) before combining the results with the corresponding scenario probabilities to the overall posterior probability of failure $P(F|\varepsilon)$ (Eq. 5.4).

2 **Integrated Monte Carlo simulation:** In an integrated Monte Carlo simulation, first we sample realizations of the discrete scenarios and, subsequently, we sample the realizations of the other (continuous) random variables conditional on the scenarios. This holds for both the assessment and the observation. Appendix B describes an algorithm for a sampling strategy which also takes of the auto-correlation of discrete scenarios in time into account.

Both approaches are equivalent and lead to the same result (within the error margins of the sampling methods). The two-stage approach has the advantage that it immediately provides

information on the individual updated probabilities of failure (and fragility curves) per discrete scenario, which allows a richer interpretation of the results. Hence, throughout the remainder of this report the two-stage implementation is used. The example in appendix section B.2 illustrates the application and equivalence of both procedures.

6 Application to dike instability and survival information

This chapter addresses several important aspects when applying the reliability updating method described hitherto to slope stability of dikes. Also the TRAS (ENW, 2009) contains valuable and relevant guidance, whicn is not all repeated here, but should certainly be incorporated in future guidance on the topic. Special attention is paid to modeling choices for the assessment and the observation conditions as well as to the limitations of the approximation with fragility curves.

Reliability updating with survival information of observed loads can be carried out with a multitude of reliability methods directly using a stability model, or by approximating the overall resistance in terms of the critical water level through fragility curves or, as explained in chapter 4. Though the focus in this chapter is on the approach with fragility curves, as we regard it the most practicable approach for the time being, most considerations equally hold for application of other reliability methods.

6.1 Typically relevant observed loading conditions

The main loading conditions for dikes are (see also Figure 6.1):

1 high (outside) water levels
2 precipitation
3 other external loads (e.g. traffic)

Observation of any significant individual load or load combination can be used for reliability updating. The list above is not exhaustive, any other significant survived condition can be used, as long as the observation can be captured in quantitative terms.

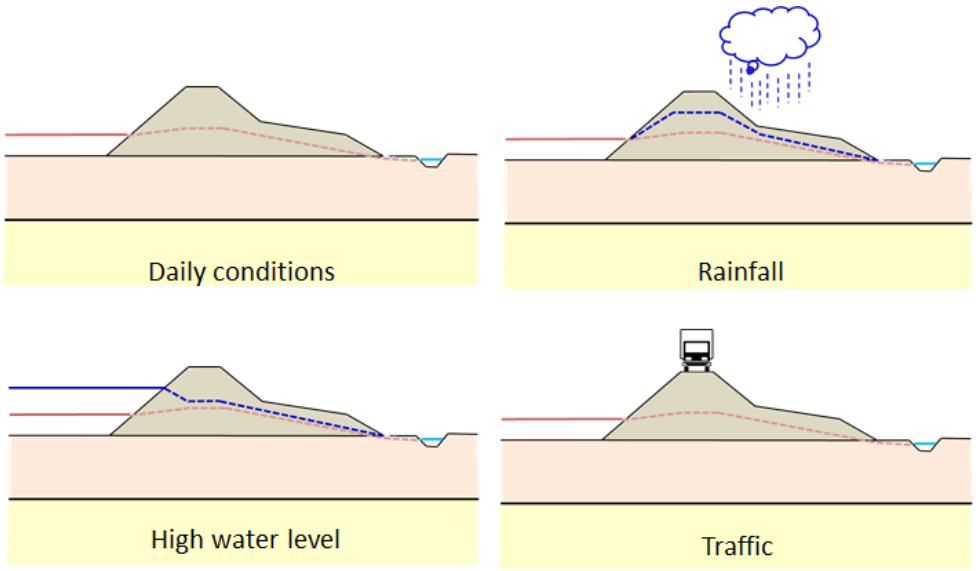


Figure 6.1: *Illustration of the main relevant loads on dikes, observations of which can be used in a reliability updating context*

6.2 How to generate fragility curves

The theory discussed hitherto regarding the approximation with fragility curves requires conditional reliability analyses for given values of the considered load parameter in order to determine the fragility points. The conventional approach in the Netherlands for slope stability is to condition the geo-hydrological response (i.e. pore water pressures) on a given water level. Then, the steps for constructing a fragility curve of beta-h curve are:

1 select a water level
2 condition the geohydrological response on this water level
3 carry out a stability analysis with mean or design values (optional)
4 find the critical sliding plane
5 carry out a FORM analysis for the critical sliding plane
6 repeat steps 1-5 for other water levels

The following remarks should be made regarding the points above:

ad 1) Typically appropriate water levels to choose are between the daily water level and the design water level (and beyond), essentially any water level that can significantly contribute to the probability of failure (combination of conditional failure probability and probability of the load).

ad 2) In most projects hitherto, such as in the VNK2 risk analyses, the geohydrological response was a deterministic cautious estimate. Recently, Kanning and Van der Krogt (2016) described how also parameters of the geohydroglogical response can be modeled as random variables, such as the leakage length and the intrusion length.

ad 3 and 4) In FORM analyses, in principle, we can search for the critical sliding plane in each deterministic slope stability analysis within the FORM iterations. Yet, often the sliding planes do not vary much for a given water level, as the water level and the resulting changes in pore pressures are the main driver for changes in the position of the sliding plane. Therefore, it can be efficient to fix the sliding plane based on a deterministic analysis using either mean or design values. Design values (e.g. characteristic values based on 5%-quantiles of ground properties, sometimes divived by a partial factor) are often the better choice as they are typically closer to the design point values that result from the FORM calculation. Of course, this simplification needs to be treated with care and it is highly recommended to verify if the sliding plane is indeed the critical one for the FORM design point values.

ad 5) The FORM analyses includes all (continuous) stochastic variables except the water level.

ad 6) The choice of additional water levels can be based on the criteria described under ad 1). Also refinement of the grid of fragility points should be sensible for highly non-linear beta-h curves.

It is important to note that, though the water level is typically the dominant load variable, fragility curves can also be generated for other load variables such as the traffic load, if required and sensible.

A major advantage of working with FORM to generate the fragility points is that we can also estimate the correlation between the resistance terms (see Equation 4.8) and assess whether the assumption of full correlation is justified.

For details on dike slope reliability analysis in realistic conditions refer to the accompanying case study report (Schweckendiek *et al.*, 2016).

6.3 Considerations for modeling the observation

Ideally, we model the assessment and observation conditions by expressing our expectations and uncertainties in terms of (continuous and discrete) probability distributions for the relevant random variables. In practical applications however, that is not always feasible, in which case we tend to make conservative assumptions. In this section we will address a number of items in this respect which have to be treated with care in order not to obtain over-optimistic or over-conservative reliability estimates.

6.3.1 Conservative versus optimistic assumptions

Conservative modeling assumptions for the assessment conditions are not conservative for the observed conditions. What do we actually mean by conservative? In a probabilistic context we try to avoid conservatism where we can by modeling uncertainties explicitly. Where that is not possible, conservative assumptions lead to rather high probabilities of failure, meaning that more detailed modeling would lead to a lower estimate for the probability of failure.

For example, since it is hard to model the uncertainty in the response of the phreatic surface to rising water levels, we commonly make a conservative model of the pore water pressure response. For the assessment conditions, modeling the maximum conceivable pore pressure response leads to an upper bound of the estimate of the probability of failure.

Now we have to realize what the implications are for modeling the observation conditions in a reliability updating context. In order not to over-estimate the effect of reliability updating, we need to make the opposite assumptions for loads, load effects and resistances compared to the prior analysis. In terms of the same example with the pore pressure response, assuming rather high pore pressures in the observation implies a rather high observed load effect, which leads to a lower bound of the updated probability of failure. This is not conservative in the given definition. In other words, in order to obtain a cautious estimate of the effect of reliability updating with survival information, we need to make "optimistic" modeling assumptions for the observation as summarized in Table 6.1.

	"Conservative" estimate	
	Assessment	Observation
loads	high	low
load effects	high	low
resistances	low	high

Table 6.1: *Conservative assumptions or estimates for the main categories loads, load effects and resistances, distinguishing between assessment and observation conditions in a reliability updating context.*

Where conservative assumptions are unavoidable, this aspect needs to be carefully examined, per relevant aspect or parameter.

6.3.2 Typically neglected resistance contributions

A subcategory of conservative modeling assumptions discussed in the previous section are contributions to the resistance of the dike which we commonly neglect in assessments and design. The reasons for not taking these features into account in current practice are typically (a) that the effect itself is negligible or (b) that the effect leads to a reduction in failure probability but is difficult to quantify, so neglecting them is still being on the safe side in assessments and designs. The list below contains several examples which may be relevant in reliability assessments of dike slope stability:

- **Unsaturated strength:** It is common practice in assessments and designs to use shear strength based on drained behavior (e.g. Mohr-Coulomb) in the partially saturated or unsaturated zone above the phreatic surface. Yet it is known that the suction in the zone above the phreatic surface due to capillary forces can lead to a higher shear strength (see illustration in Figure 6.2).
- **Three-dimensional failure planes:** The two-dimensional limit equilibrium models commonly used in assessments and designs assume a plane-strain condition, i.e. a sliding surface of infinite length in the longitudinal direction of the dike. In reality, sliding planes in dikes are of finite length, typically 50 to 200 m long. That implies that the resistance of the three-dimensional failure surface will provide more shear resistance than the plane-strain model suggests. For rather homogenenous dikes sections this effect is accounted for with the model uncertainty m_d (the assumed performance function is $g = SF \cdot m_d - 1$, see Duinen (2015) for background information). In case of variable geometries, e.g. in dike sections with buildings or other objects, there may be an effect of the shape of the failure plane beyond what is considered in the model uncertainty.
- **Trees or other objects on the dike:** Objects on the dike or even penetrating it can affect the stability, both positively as well as negatively.

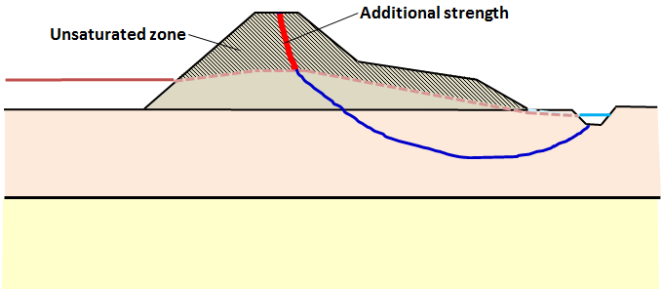


Figure 6.2: *Illustration of unsaturated or partially saturated zone in a dike above the phreatic surface and the part of the sliding plane with potentially "additional strength" compared to conventional safety assessment slope analyses where the effects of unsaturated strength are typically neglected.*

Note that the list above is by no means exhaustive and merely an illustration of the types of resistance contributions that may need to be considered depending on the local conditions.

A pragmatic approach to such resistance contributions is to start with a sensitivity analysis in order to scrutinize whether the effect can be discarded as negligible or needs to be addressed explicitly. Phenomena with significant effect will need to be addressed in some way, even though the effect is favorable for stability, in contrast to assessment and design.

6.4 Epistemic versus aleatory uncertainty

There are various variables that affect the reliability of a dike for slope stability of the inner slope. For reliability updating, it is important to know whether a variable is considered as time-invariant or as a random process representing aleatoric variability. The uncertainty in time-invariant properties is necessarily epistemic (i.e. knowledge uncertainty) and can, hence, be reduced. The list below provides an overview of the most relevant categories of variables, including arguments why these properties typically below to either category - predominantly epistemic or aleatoric uncertainty:

- **Soil properties:** Soil properties describing inherent properties of the material such as volumetric weight, friction angle or cohesion are typically time-invariant (in engineering terms). State variables, on the other hand such as the overconsolidation ratio (OCR) may be subject to changes in time (e.g. due to ageing). The assumption of time-invariance needs to be taken for each individual case considering the contemplated time scale. Note that changes in time do not necessarily imply that a variable enters the aleatory category, as changes in time can also be modeled explicitly as processes with epistemic uncertainty (e.g. degradation).
- **Geo-hydrological response:** The geo-hydrological response of a dike (changes in pore water pressures) to outside water levels can be based on parametric response model using e.g. leakage lengths and intrusion lengths to model the increase in pore water pressure as a function of outside water level (Kanning and Van der Krogt, 2016). Even though the pore water pressure varies in time, the parametric response properties can be assumed as time-invariant. An exception may be the phreatic surface, which does not only respond to the external water level but also to precipitation events, which themselves are of aleatory nature.
- **External loads:** External loads such as traffic loads may be different and uncorrelated between assessment and observation; these are therefore typically assumed as aleatory.
- **Model uncertainty:** Model uncertainty contains epistemic and aleatory components, as discussed in Schweckendiek (2014) in more detail. It is always safe to assume model uncertainty to be aleatory in a reliability updating context, Yet, arguably, in case of slope stability analysis, model uncertainty will be dominated by local bias which is time-invariant and reducible. Hence, if model uncertainty is not dominant, we may choose to assume it to be epistemic; if it is dominant, we should assess thoroughly, if the epistemic assumption does not lead to a severe over-estimation of the effect of reliability updating.
- **Stratification:** The presence of soil strata or geological anomalies are classic examples of epistemic uncertainty. They do not change, appear or disappear on an engineering time scale. We are just uncertain about their presence and sometimes size based on the limited site investigation data at hand.

In general we can state that the larger the share of epistemic uncertainty with respect to the total uncertainty, the greater the potential impact of reliability updating. Naturally the list of categories/variables above is not exhaustive and the choice to model a variable as time-invariant or not can vary from case to case.

For concrete examples of the choices for realistic slope stability problems refer to the accompanying case study report (Schweckendiek *et al.*, 2016).

6.5 Sliding surfaces to be considered

6.5.1 General method

The general reliability updating method based on Bayes' rule as described in chapter 3 virtually does not have limitations in terms of applicability except for practical reasons like modeling and computational efforts. When used with sampling methods like Monte Carlo simulation, we may consider the critical sliding surface (i.e. lowest factor of safety) for each realization to evaluate the (non)exceedance of the stability limit state.

The approximation with fragility curves as explained in chapter 4 is based on simplifications introducing additional requirements in terms of the sliding surfaces considered, which we address in the following sections. If these specific requirements cannot be met, we may have to resort to other reliability methods like Monte Carlo simulation directly with the stability model instead of approximating with fragility curves.

6.5.2 Critical sliding surfaces

Common practice is to first identify the a-priori critical sliding surface(s) and base the reliability (updating) analysis on that very surface (or several ones). It is important to note that the notion of a critical sliding surface is slightly different than in a deterministic setting. In a reliability analysis the probability of slope instability is ideally considered as the probability of all potential sliding planes combined (i.e. $P(slope\ failure) = P(\cup_{i=1}^{n} sliding\ plane\ i)$), taking into account the mutual correlations. Practical limitations like computation times often impede following this rigorous approach.

In this context, the implicit assumption of considering the critical sliding plane, being the one with the highest probability of failure among the potential ones, is that the combined probability of failure will be approximately equal to the one determined for the critical sliding plane. This assumption is often justifiable as typically all other potential sliding planes are highly correlated with the critical one (for perfect correlation: $P(slope\ failure) = \max_{i=1}^{n} P(sliding\ plane\ i)$).

6.5.3 Sliding surface in observation

In the approach with fragility curves, an overall resistance variable is used for both the assessment and the observation conditions. In other words, we aggregate all information of the individual random variables in one resistance variable, typically the critical water level. It is assumed that these resistances represented by the fragility curves are correlated, which implies that they need to be comparable in the sense that we address the same resistance. In practical terms for slope stability, this boils down to the requirement that the failure surface (slip plane) needs to be essentially the same for observation and assessment.

For practical application, that means that we define the critical sliding surface for the assessment conditions and then consider the same sliding surface in the observed, survived conditions. The critical sliding surfaces in the assessment and the observation conditions can be different. The important notion here is that in reliability updating with fragility curves, we contemplate an overall resistance term which is represented by a specific sliding surface or failure mode. And that needs to be the same in assessment and observation.

6.5.4 Several potentially critical sliding planes

Sometimes, a critical sliding surface cannot be identified a-priori. For example, if uplift conditions are critical, the location of the critical sliding plane may differ substantially between low and high water levels (Figure 6.3).

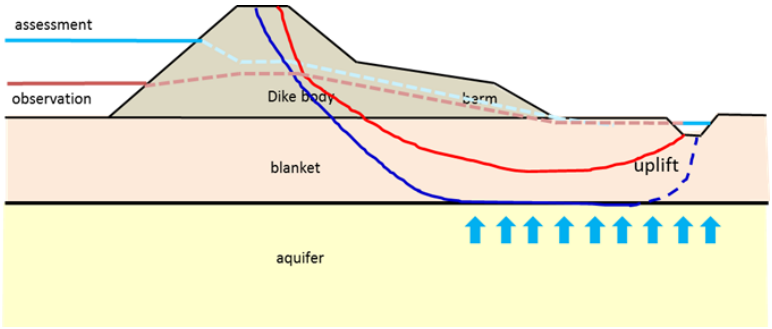


Figure 6.3: *Illustration of how the occurrence of uplift can change the location of the critical sliding plane.*

Also realize that in a reliability updating context, the critical sliding plane a-posteriori can be different from the a-priori one, because several sliding planes my experience more or less impact from incorporating the survival information. In cases where the (prior or posterior) critical sliding plane cannot be determined a-priori with confidence, it is highly recommendable to contemplate all potential critical sliding planes separately in both the prior and the reliability updating analysis.

6.5.5 Changes to the cross cross section

The approximation with fragility curves can in principle also be used in a reinforcement design setting, as long as the failure mechanism in question is not altered by the reinforcement measures in terms of the failure mechanism (e.g. rotation of principal stresses). For example, a change in the outer slope angle or the replacement of the revetment will typically not affect the inner slope stability, in which case the reliability of the inner slope can still be updated with the approximation using fragility curves.

7 Examples and benchmark tests

This chapter contains examples to illustrate the application of the reliability updating method as explained hitherto. The examples also serve as benchmarks for comparing the approximation using fragility curves with exact solutions (e.g. MCS directly with model simulations).

7.1 Example 1: Critical water level versus water level

The first benchmark example concerns a simple linear performance function with two normal distributed random variables. The purpose of this example is a step-wise illustration of the approach for a problem with known exact results. For implementation details and more information refer to appendix C.

7.1.1 Input data and prior reliability

The performance function is

$$g = h_c - h \tag{7.1}$$

with h_c being the critical water level (i.e. the water level at which the dike fails) and h the water level. The probability distributions of both are defined in Table 7.1 and illustrated in Figure 7.1, along with the observed water level which we assume to have survived and the estimated critical water level at the time of the observation.

Symbol	Unit	Description	Distribution	Mean	Standard deviation
h_c	[m]+REF	critical water level	Normal	6	$\sqrt{2}$
h	[m]+REF	water level	Normal	2	$\sqrt{2}$
h_{obs}	[m]+REF	observed water level	Deterministic	5	N/A
$h_{c,obs}$	[m]+REF	critical water level at observation	Normal	7	$\sqrt{2}$

Table 7.1: *Example 1: Probability distributions of variables*

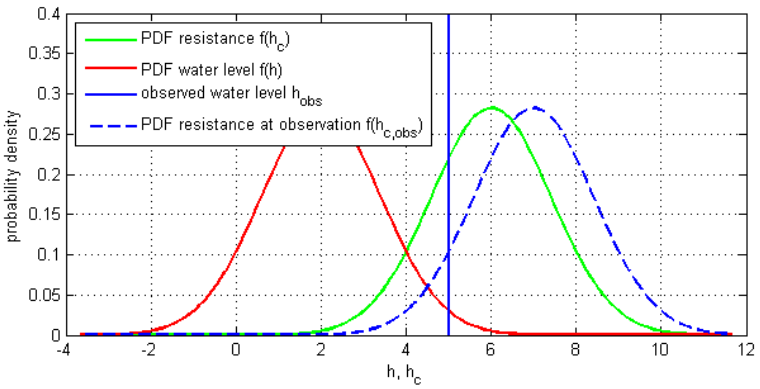


Figure 7.1: *Example 1: Probability distributions of the (critical) water levels for assessment and observation*

Note that the expected critical water level at the time of the observation is estimated to have been higher than for the assessment. Qualitatively speaking, this is a typical situation we

encounter in practical applications due to degradation processes (e.g. settlements) or differences in the presence of traffic loads etc.

For such linear performance functions with normal-distributed random variables, the exact prior reliability index can be obtained analytically by

$$\beta = \frac{E[g]}{\sigma_g} = \frac{\mu_{h_c} - \mu_h}{\sqrt{\sigma_{h_c}^2 + \sigma_h^2}} = \frac{6-2}{\sqrt{2+2}} = 2 \tag{7.2}$$

7.1.2 Prior analysis with fragility curves

Figure 7.2 depicts the fragility curve for the contemplated problem, whereas Figure 7.3 shows the corresponding beta-h curve (linear interpolation between the known fragility points).

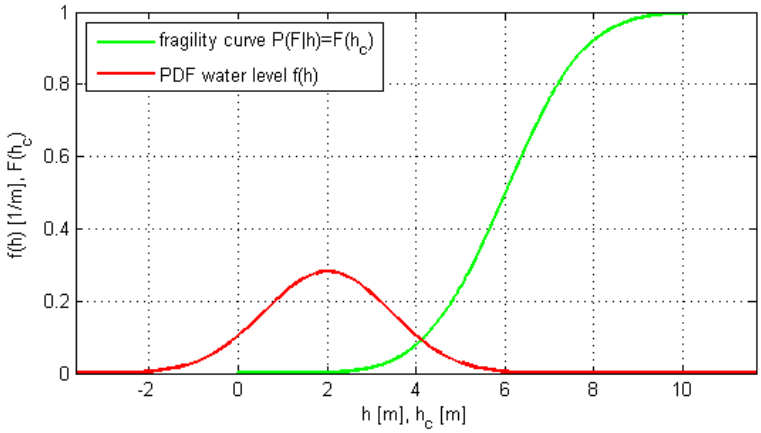


Figure 7.2: *Example 1: Fragility curve*

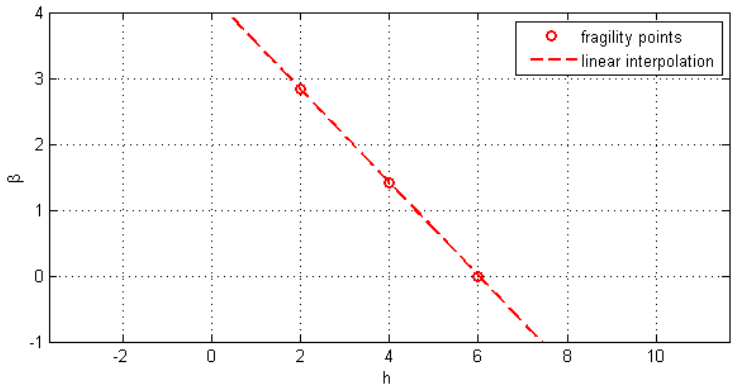


Figure 7.3: *Example 1: Beta-h curve (linear interpolation between pre-determined fragility points)*

Note that the linear interpolation between the fragility points in the beta-h curve (Figure 7.3) does not introduce an approximation error in this specific case, as the beta-h curve for a Normal-distributed resistance variable is actually a straight line.

Monte Carlo simulation based on the beta-h curve as described in section 4.6 confirms the prior reliability index of 2, which shows that the approach produces exact results for linear beta-h curves. For details refer to appendix C.

7.1.3 Posterior analysis exact and with fragility curves

In order to incorporate the information of the observed survived water level of $h_{obs} = 5$ m above reference level, we follow the sampling approach as described in section 4.4 with the beta-h curves as depicted in Figure 7.4. Note that the horizontal shift between the two lines reflects the estimated (mean) difference in resistance between the conditions for the assessment and for the observation.

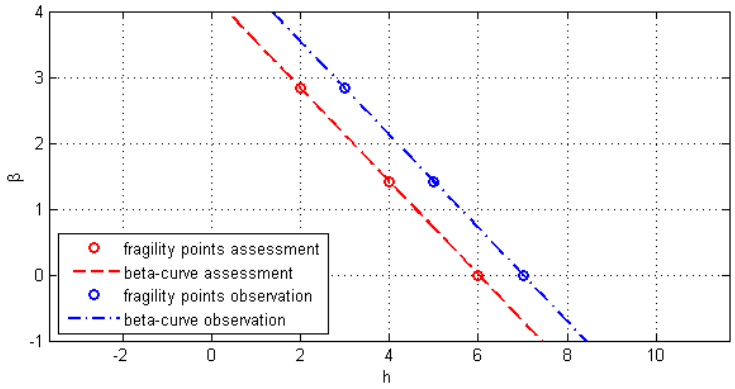


Figure 7.4: *Example 1: Beta-h curves for the assessment and observation conditions (linear interpolation between pre-determined fragility points)*

	Reliability index	Probability of failure
Prior	2.00	$2.3 \cdot 10^{-2}$
Posterior	2.35	$9.4 \cdot 10^{-3}$

Table 7.2: *Example 1: Prior and posterior reliability estimates*

Table 7.2 shows that the updated probability of failure is roughly a factor 2.5 lower than the prior value. Figure 7.5 shows that the posterior resistance distributions are truncated at a value somewhat lower than the observed load. The reason is that the assessment conditions exhibit a somewhat lower resistance compared to the observed conditions. In this specific case with two exactly parallel beta-h curves, the shift in truncation value is precisely the horizontal distance between the assessment and the observation curve, which in turn is the difference in mean critical water level (resistance).

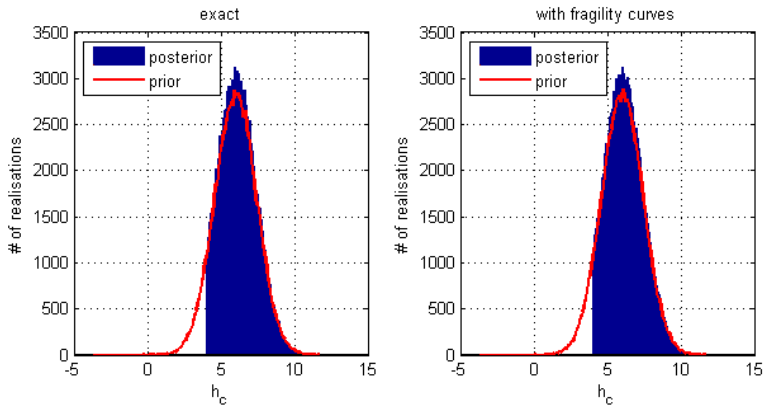


Figure 7.5: *Example 1: Histograms of the prior and posterior realizations of the critical water level comparing direct MCS and sampling from fragility curves*

7.1.4 When to expect more or less effect based on the fragility curves?

This section provides considerations on when to expect more or less effect by reliability updating based on characteristics of the fragility curves for assessment and observation. For the sake of illustration we assume the fragility curves parallel to each other and separated by Δ (in horizontal direction).

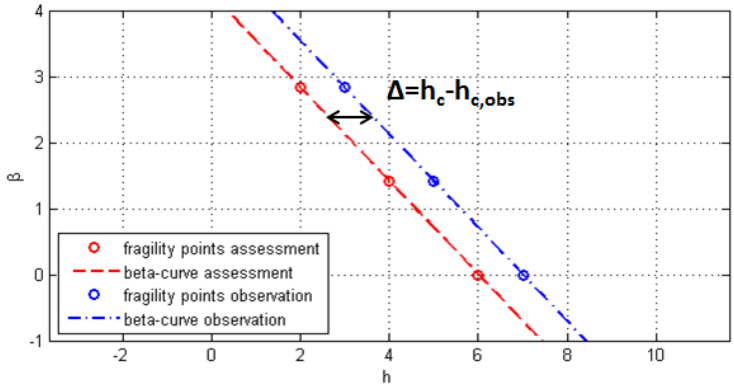


Figure 7.6: *Example 1: Difference between h_c and $h_{c,obs}$ in fragility curve*

Note that this implies that the critical water level in the observation is estimated to be exactly Δ higher than in the assessment conditions (i.e. $\Delta = h_c - h_{c,obs}$) also illustrated in the PDF of h_c of Example 1 in Figure 7.7. Figure 7.7 also illustrates that the probability distribution of h_c is effectively truncated at $h_{obs} - \Delta$, removing area A from the PDF. Unsurprisingly, for $\Delta = 0$ the probability of the critical water level below the observed water level becomes zero.

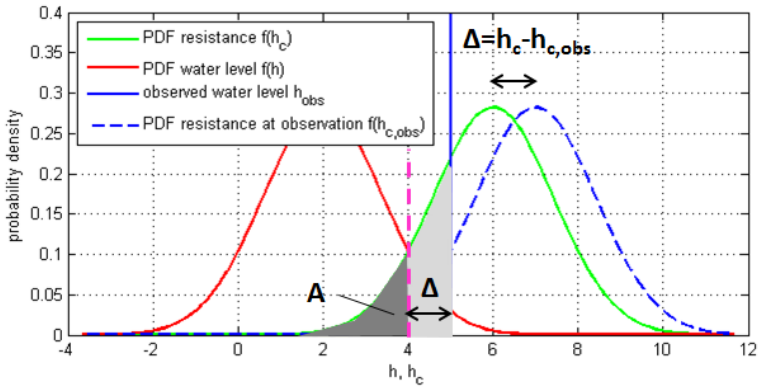


Figure 7.7: *Example 1:Truncation of the PDF of the critical water level ($f(h_c)$ at the value $h_{obs} - \Delta$ for two linear and parallel fragility curves).*

Consequently, we expect a relatively large effect of reliability updating for rather flat fragility curves (i.e. large uncertainty in the resistance) with a relatively high observed load and little distance between the assessment and observation curve. Especially for rather flat curves, relatively small differences between assessment and observation can lead to a high Δ (horizontal distance) and, hence, a rather low effect on the probability of failure.

7.2 Example 2: Internal erosion with Bligh's rule

While the previous example has illustrated the effect of probability updating based on survival information for a single resistance variable (h_c), this second example shows that the same principles are valid with a multidimensional problem (i.e. various resistance variables) and a non-normal distributed overall resistance term. The theory is applied to Bligh's rule (i.e. internal erosion) as it is described in Schweckendiek (2014). For details refer to appendix C.

7.2.1 Input data and prior reliability

The performance function based on Bligh's rule reads:

$$g = m_B \frac{L}{c} - h \tag{7.3}$$

where m_B is the model uncertainty factor as defined in Kanning (2012), based on a Bayesian Analysis of dike failures and survivals, L [m] is the length of the seepage length, c is the percolation factor which depends on the grain size of the erodible material in the aquifer and in this case h [m] is the head difference over the structure (see Figure 7.8).

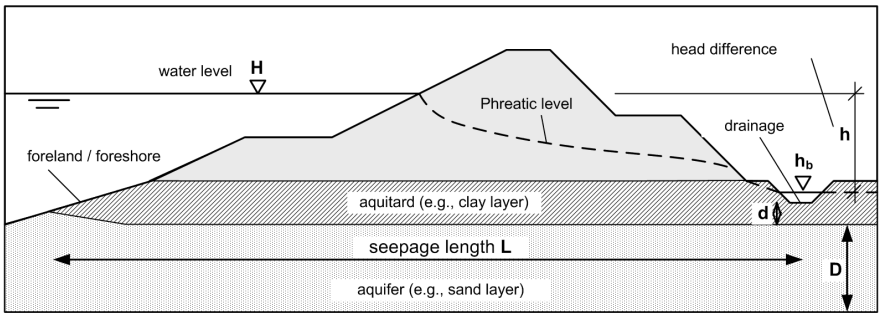


Figure 7.8: *Definitions for Bligh's rule model (Schweckendiek, 2014)*

The example parameters are defined in Table 7.3, both for the assessment as well as for the observation (survival) conditions. The parameter Δ_L reflects the estimated mean difference between the length of the seepage path, implying that we assume the seepage length (mean value) to have been larger at the time of the observation than for the assessment (see also Figure 7.9), for example because the structure was amended due to a change in usage.

Symbol	Unit	Description	Distribution	Parameters
m_B	[-]	model uncertainty factor	Lognormal	$\mu = 1.76, \sigma = 1.69$
L	[m]	seepage length	Normal	$\mu = 50, \sigma = 2.5$
c	[-]	percolation factor	Deterministic	$c = 18$
h	[m]	head difference	Gumbel	$\alpha = 0.53, \beta = 0.406$
h_{obs}	[m]	observed head difference	Deterministic	2.4
Δ_L	[m]	difference between observation and assessment	Deterministic	10.0

Table 7.3: *Example 2: Probability distributions and parameters*

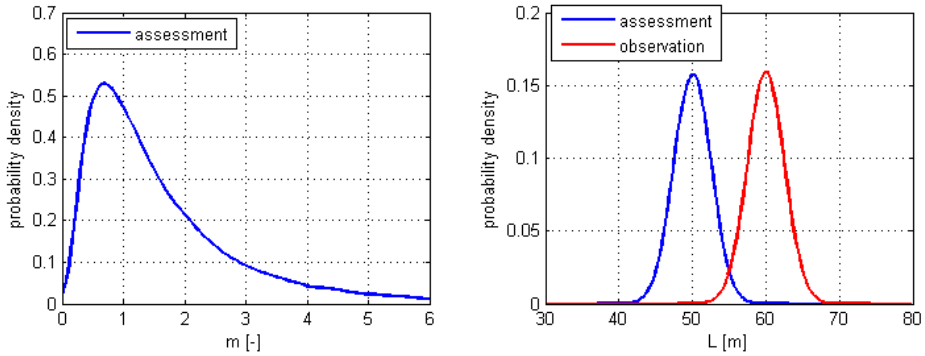


Figure 7.9: *Example 2: Probability distributions of m_B and L parameters for assessment and observation*

Note that the critical head difference in this case is given by $h_c = m_B \frac{L}{c}$. The resulting distribution of the overall resistance term is depicted in Figure 7.10 (black line) along with the probability distribution of the head difference (green line) and the observed water level (blue dashed line).

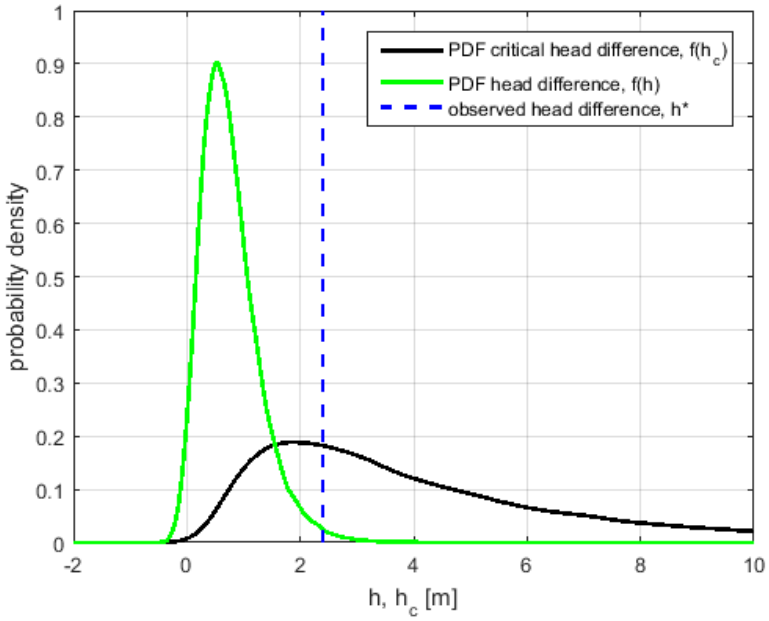


Figure 7.10: *Example 2: Probability distributions of the (critical) head difference for assessment and observation*

The prior reliability estimated using Monte Carlo simulation (MCS) is 1.65, with a corresponding probability of failure of $4.8 \cdot 10^{-2}$.

7.2.2 Prior analysis with fragility curves

Knowing the exact result for the prior reliability we now investigate the performance of the approximation with fragility curves on the prior results. As depicted in Figure 7.11 the beta-h curve is clearly non-linear, implying a non-normal distributed overall resistance term (i.e. h_c). Hence, the discretization of the curve matters for the interpolation between the fragility points.

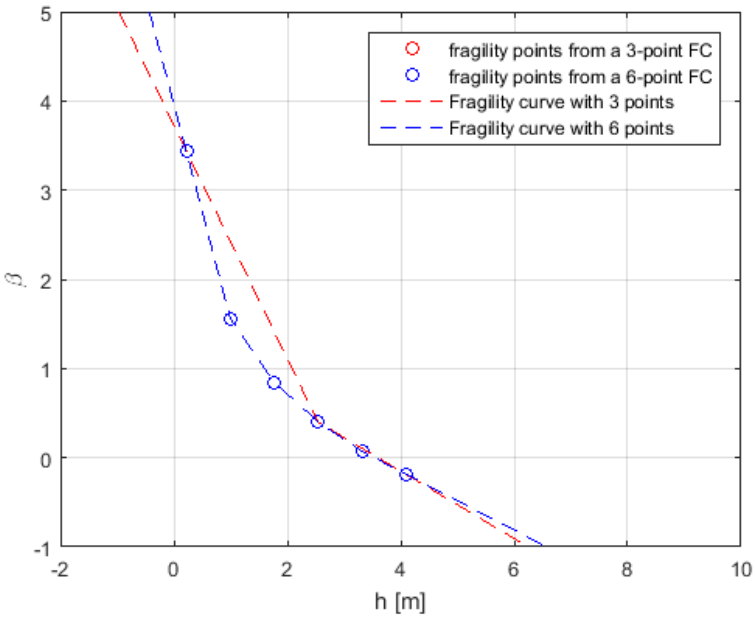


Figure 7.11: *Example 2: Beta-h curves for the assessment conditions with 3 (red line) and 6 (blue line) fragility points respectively*

Table 7.4 shows the reliability estimates for both the exact method and the approximation method with fragility curves for difference discretizations (i.e. number of fragility points). The approximation method produces results close to the exact method for an appropriate level of discretization, in this case 6 fragility points. The approximation with only 3 fragility points leads to an overestimation of the reliability index due to the high non-linearity of the beta-h curves in the region with significant probability density of the head difference (load).

		MCS	FC
3 points	$PriorP(F)$	$4.8 \cdot 10^{-2}$	$1.3 \cdot 10^{-2}$
	β	1.65	2.23
6 points	$PriorP(F)$	$4.8 \cdot 10^{-2}$	$4.2 \cdot 10^{-2}$
	β	1.65	1.73

Table 7.4: *Example 2: Prior probability of failure and reliability index comparing the exact results obtained with Monte Carlo simulation (MCS) with the approximation using fragility curves (FC) with 3 and 6 fragility points for the construction of the beta-h curve respectively (see Figure 7.11)*

7.2.3 Posterior analysis exact and with fragility curves

Table 7.5 presents the results of the posterior analysis for both the exact MCS and the approximation with beta-h curves with 3 and 6 fragility points respectively. Note that the beta-h curve for the observation in Figure 7.12 lies slightly above the assessment curve, as we assume the seepage length at the time of the observation to be larger ($\Delta_L = 10$ m).

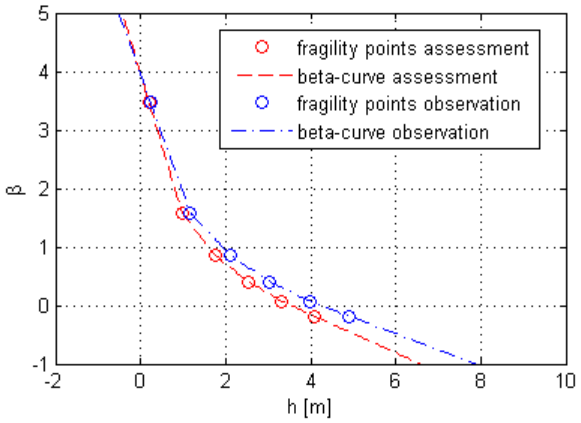


Figure 7.12: *Example 2: Beta-h curves for assessment and observation ($\Delta_L = 10$ m)*

Similar to the prior analysis, also the posterior reliability indices computed with the approximation are slightly overestimated with respect to the exact MCS results for 3 fragility points. The refined grid of 6 fragility points in turn provides virtually exact results. Furthermore, it is noteworthy that the effect of reliability updating is less pronounced, if we assume the structure was stronger at the time of the observation than it is now ($\Delta_L = 10$ m), which is plausible.

		$\Delta_L = 10$		$\Delta_L = 0$	
		MCS	FC	MCS	FC
3 points	$PosteriorP(F\|\varepsilon)$	$2.7 \cdot 10^{-3}$	$3.5 \cdot 10^{-3}$	$8.0 \cdot 10^{-4}$	$1.3 \cdot 10^{-3}$
	β	2.78	2.69	3.14	3.01
6 points	$PosteriorP(F\|\varepsilon)$	$2.7 \cdot 10^{-3}$	$2.5 \cdot 10^{-3}$	$8.0 \cdot 10^{-4}$	$1.0 \cdot 10^{-3}$
	β	2.78	2.80	3.14	3.10

Table 7.5: *Example 2: Posterior reliability estimates with and without difference between assessment and observation ($\Delta_L = 10$ m) comparing Monte Carlo simulation (MCS) and the approximation with fragility curves (FC) for 3 and 6 fragility points as shown in Figure 7.12*

Figure 7.13 depicts the prior and posterior joint probability density functions (JPDF) of the seepage length and the Bligh model factor. Comparing the prior and the posterior illustrates that the posterior JPDF is practically truncated and the probability density in the implausible parameter space given the observation becomes zero, while the probability mass is redistributed over the feasible region.

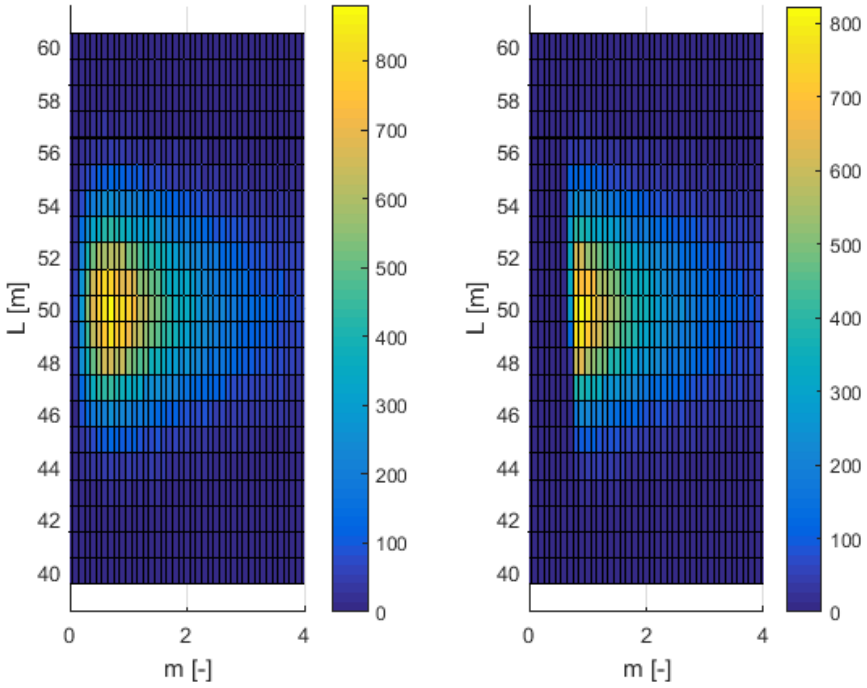


Figure 7.13: *Example 2:Prior (left figure) and posterior (posterior figure) JPDF plot of a two-dimensional histogram for m_B and L parameters*

7.3 Example 3: Simple slope stability problem

As opposed to examples 1 and 2 with simple closed-form performance functions, this example treats a more complex implicit performance function for slope stability with multiple load and resistance variables.

7.3.1 Input data and prior reliability

Figure 7.14 illustrates the geometry and main characteristics of the case at hand. The contemplated slope stability problem consists of a clay dike (*Clay*) on a clay blanket (*Claylayer*) on top of a sand aquifer (*Aquifer*). For the sake of simplificity, all stability analyses are done with a fixed slip circle (with the lowest stability factor for mean values of all stochastic variables).

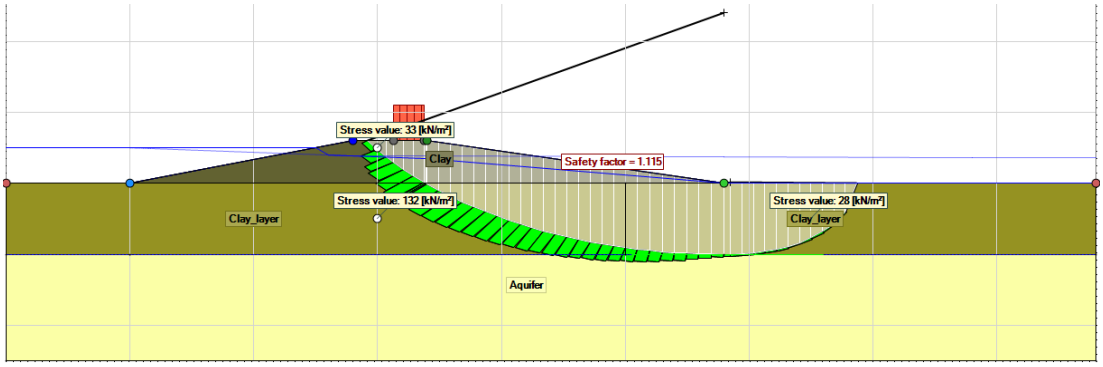


Figure 7.14: *Example 3: Simple slope stability problem with clay dike on clay blanket on top of a sand aquifer (the white circles demonstrate the yield stress points as defined in D-Geostability).*

Name	Unit	Description	Distribution	Parameters
$Clay\_Su$	[-]	undrained shear strength ratio	Lognormal	$\mu = 0.35, \sigma = 0.10$
$Claylayer\_Su$	[-]	undrained shear strength ratio	Lognormal	$\mu = 0.30, \sigma = 0.03$
$Clay\_m$	[-]	strength increase exponent	Lognormal	$\mu = 0.90, \sigma = 0.02$
$Claylayer\_m$	[-]	strength increase exponent	Lognormal	$\mu = 0.90, \sigma = 0.02$
$Aquifer\_c$	$[kN/m^2]$	cohesion	Deterministic	0.0
$Aquifer\_phi$	[°]	friction angle	Deterministic	$35°$
$Clay\_yield$	$[kN/m^2]$	yield stress	Lognormal	$\mu = 38, \sigma = 6$
$Claylayer\_under\_yield$	$[kN/m^2]$	yield stress	Lognormal	$\mu = 137, \sigma = 6$
$Claylayer\_next\_yield$	$[kN/m^2]$	yield stress	Lognormal	$\mu = 28, \sigma = 6$
$Water\ level$	$[m]$	outside water level	Gumbel	$shift = 1.5, scale = 0.4$
m_d	[-]	model uncertainty	Lognormal	$\mu = 0.995, \sigma = 0.033$

Table 7.6: *Example 3: Probability distributions of variables.*

The performance function we evaluate in the reliability analysis is:

$$g = SF \cdot m_d - 1 \tag{7.4}$$

where SF is the stability factor that is retrieved from the D-GeoStability[1] and m_d is the model uncertainty. The dike material and the clay blanket are modeled using undrained shear

[1]The exact computation procedures and definition of the stability factor can be found in the D-Geostability manual, which can be downloaded from https://download.deltares.nl/en/download/geotechnical-software/.

strength, whereas the shear strength of the aquifer is modeled using Mohr-Coulomb. The input parameters are given in Table 7.6.

As a benchmark, the probability of failure was estimated using Monte Carlo simulation (MCS) with $20,000$ samples. More precisely, Monte Carlo sampling of all stochastic variables, including the water level, results in a prior reliability index of 2.15, which corresponds to a probability of failure of $1.6 \cdot 10^{-2}$.

7.3.2 Prior analysis with fragility curves

Figure 7.15 depicts the fragility curve for the assessment conditions. One of the main advantages of using fragility curves in a probabilistic slope stability analysis is that the fragility points are calculated for fixed water levels and, hence for a fixed (steady state) pore pressure field. For a fixed pore pressure field, experience shows that the critical sliding plane (i.e. with lowest SF) is rather stable even for changing soil parameters, which often allows us to work with a fixed slip plane in all FORM iterations.

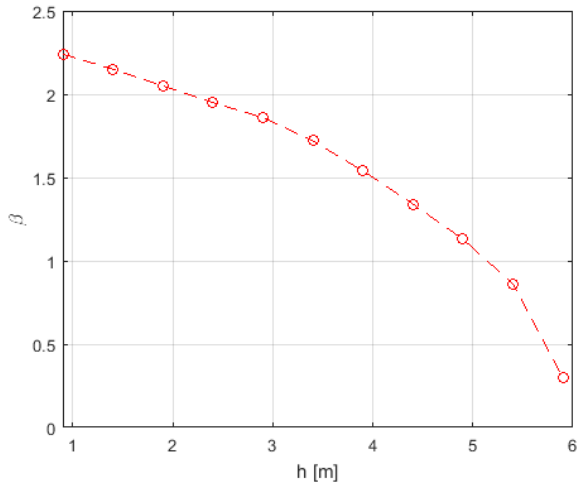


Figure 7.15: *Example 3: The fragility curve (beta-h curve) for the assessment conditions.*

Table 7.7 shows that the prior reliability index as estimated using MCS and approximation using fragility curves (FC) are very close. The small approximation error can be due to (a) the approximation by the fragility curve and/or (b) the MCS sample size (which can be tested by repeating the experiment several times).

	β	D-Geo Stability calculations
MCS (exact)	2.15	20,000
FC	2.07	250

Table 7.7: *Example 3: Prior reliability indices from MCS and the approximation with fragility curves (FC), including the number of D-GeoStability analyses performed per method.*

7.3.3 Posterior analysis exact and with fragility curves

The observed survived water level is considered deterministic: $h_{obs} = 2$ m. For both reliability updating approaches (MCS and FC), in general we would need to configure two D-Geostability models, one for the assessment conditions and one for the observed conditions. In the current example, the only difference is the water level, so we only need to estimate one fragility curve. In general, many conditions can differ between the two models to accomodate known changes between the observation and the assessment (e.g. the traffic load).

Table 7.8 shows the posterior results of both methods, MCS and FC, which compare very well, even though the approximation with fragility curves requires substantially less stability analyses. Note that the approach actually involves two approximations, the linearization of the beta-h curves and the estimation of the fragility points by FORM.

	β	D-Geo Stability calculations
MCS (exact)	3.05	100,000
FC	3.05	500

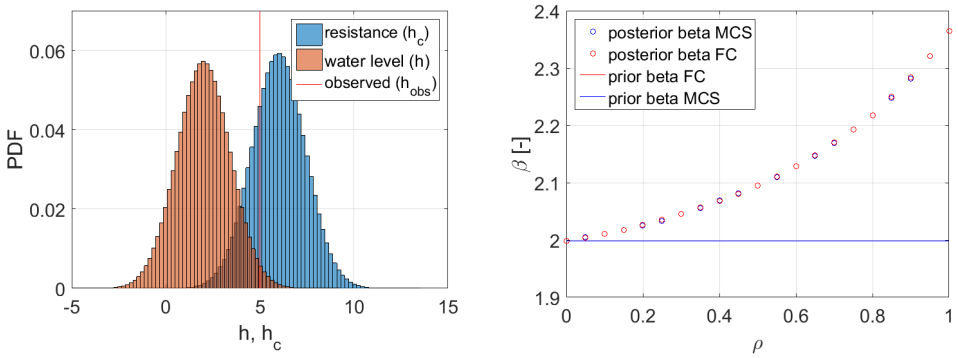
Table 7.8: *Example 3: Posterior reliability indices from Monte Carlo simulation (MCS) and the approximation with fragility curves (FC), including the number of D-GeoStability calculations.*

7.4 Example 4: Correlation between assessment and observation

This section investigates the effect of correlation between the resistance in assessment and observation as explained in section 4.5 considering the simple linear performance function from Example 1 ($g = h_c - h$). In first instance, we will used the exact same parameters as in section 7.1 and contemplate a range of correlation coefficients ρ (between h_c and $h_{c,obs}$). Subsequently, we will change the example parameters in a sensitivity study.

7.4.1 Base case

The prior reliability index of the base case of 2.0 is the same as for Example 1 (section 7.1). Accordingly, the posterior reliability index assuming perfect auto-correlation in time of the critical water levels was assumed (i.e. $\rho = 1$) amounts to 2.35. Figure 7.16 depicts how the posterior reliability index decreases with decreasing correlation coefficient ρ, approaching the prior value (no updating effect) as the correlation approaches zero.



***(a)** Histograms of the prior realizations (MCS)* ***(b)** Posterior reliability index*

***Figure 7.16:** Example 4: Influence of the correlation coefficient ρ on the posterior reliability $for h_{obs} = 5$ (with $h_c \sim N(6, \sqrt{2})$ and $h \sim N(2, \sqrt{2})$.*

Figure 7.16 also provides a comparison of the approach with fragility curves with full Monte Carlo simulation of the underlying basic random variables. The results are virtually the same. Figure 7.17 illustrates that there is a gradual increase in the posterior reliability index with increasing correlation coefficient also for lower and higher observed loads.

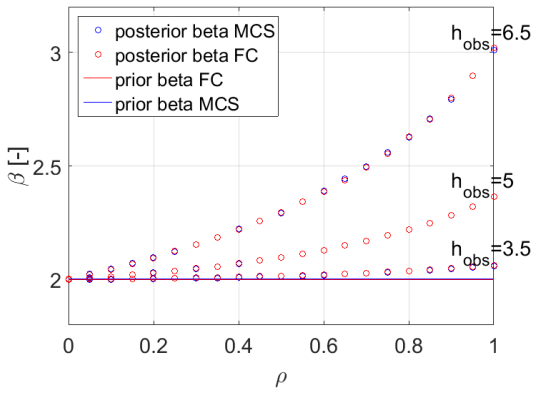


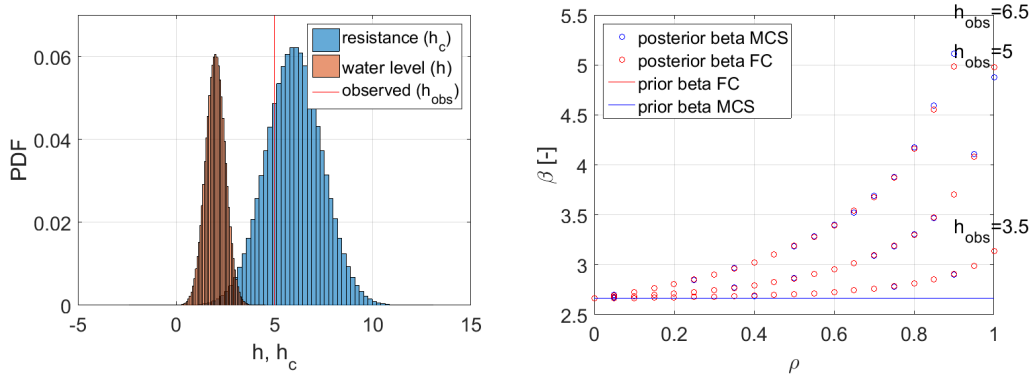
***Figure 7.17:** Example 4: Influence of the correlation coefficient ρ on the posterior reliability for different observed loads (with $h_c \sim N(6, \sqrt{2})$ and $h \sim N(2, \sqrt{2})$.*

7.4.2 Variations

In this section we vary the mean and standard deviation of the critical water level (h_c) and the observed water levels in order to examine the influence of the correlation between assessment and observation for different reliability levels and different relative contributions of load and resistance to the total uncertainty.

Lower standard deviation of the load

The standard deviation of the load (h) is reduced compare to the base case. The results are shown in Figure 7.18. The increase in posterior reliability index is less gradual with increasing correlation coefficient compared to Figure 7.17. It seems that with a relatively large contribution of the resistance, the so-called "hockey stick curve" is resembled here.

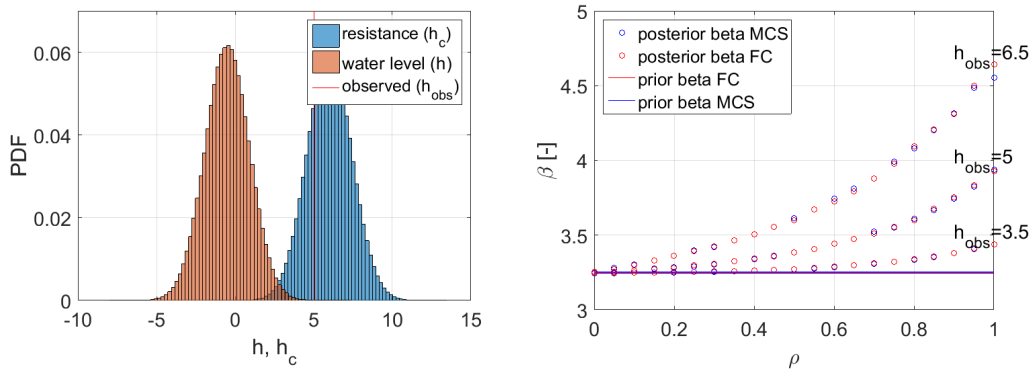


***(a)** Histograms of the prior realizations (MCS)* ***(b)** Posterior reliability index for different h_{obs}*

***Figure 7.18:** Example 4: Influence of the correlation coefficient ρ on the posterior reliability for different observed loads ($h_c \sim N(6, \sqrt{2})$ and $h \sim N(2, 0.5)$).*

Higher mean resistance

A higher mean value of the resistance (keeping the standard deviation as in the base case) leads to a higher prior reliability index. The posterior results are shown in **??**. The behavior is similar to the previous variations.



***(a)** Histograms of the prior realizations (MCS)* ***(b)** Posterior reliability index for different h_{obs}*

***Figure 7.19:** Example 4: correlation coefficient ρ analysis for the performance function $g = h_c - h$, where $h_c \sim N(6, \sqrt{2})$ and $h \sim N(2, sqrt2)$.*

Summary of the relative effects

The influence of the correlation coefficient on the posterior reliability is summarized in Figure 7.20 for various variations of the parameters. Depending on the relative contributions of load and resistance to the total uncertainty there is a more or less pronounced or gradual effect of the correlation.

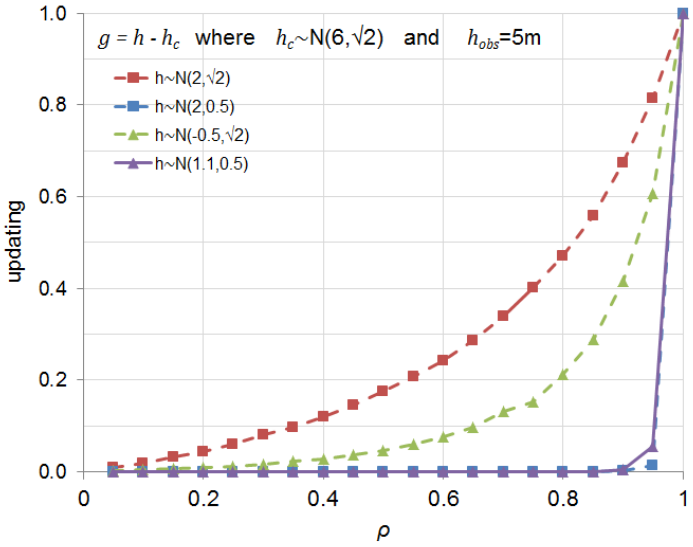


Figure 7.20: *Example 5: Reliability updating effect as a function of the correlation coefficient ρ. The updating effect on the vertical axis is defined as 1 standing for the effect (difference between prior and posterior reliability index) achieved with correlation 1, and zero with zero correlation.*

7.5 Concluding remarks for all examples

The examples in this chapter have illustrated the application of the theory described in this report to problems of varying complexity and highlighting different aspects each one. We point out the following observations:

- The approximation with fragility or beta-h curves produces an exact result for a linear performance function with a normal-distributed resistance term (see example 1).
- The method can also be applied to multi-variate, non-linear problems with non-normal distributions. The error in the approximation depends on the discretization of the beta-h curves (number and location of fragility points, see example 2).
- Building beta-h curves for slope stability problems with a limited number of fragility points using FORM requires substantially less computations than sampling based reliability methods like Monte Carlo simulation. Note that FORM often exhibits convergence problems when applied to conditions with varying water levels (and pore pressure fields), which is avoided in analyses conditioned on fixed water levels (see example 3). Using FORM does introduce an error, though, which could be avoided or reduced by using other reliability methods.
- The correlation between the resistance in assessment conditions and the resistance in observation conditions clearly impacts the posterior reliability. Decreasing correlation results in a lower effect by reliability updating. It is therefore recommended to always estimate and take into account the correlation explicitly.

For more realistic conditions in terms of the level of complexity and detail of slope reliability analysis of dikes refer to the accompanying case study report (Schweckendiek *et al.*, 2016).

8 Conclusion

This report has described how survival of observed load conditions can be taken into account quantitatively in general and in particular for slope stability of dikes. In particular, a simplified approach with fragility curves has been introduced, building upon the work by Calle (2005) and the approach outlined in ENW (2009). Working with fragility curves is common practice for slope reliability analyses in the Netherlands (e.g. VNK2-project). In this sense, the approach with fragility curves is a straightforward extension of the prior reliability analysis.

Furthermore, we have shown in chapter 2 how the results can be used in the Dutch safety assessment framework for primary flood defenses. In essence, the updated probability of failure obtained from the analysis can be directly compared to the acceptable probability of failure for a specific dike section as defined in WBI-2017.

The presented method can handle the distinction between epistemic (reducible) and aleatory (irreducible) uncertainties and it can accommodate the (known) differences between the conditions at the time of the observation and the conditions at the time of the assessment explicitly. The only clear limitation of the proposed approximation with fragility curves in the context of slope stability analyses is that the failure mechanism, in this case the critical sliding plane, needs to be essentially the same in the observations and assessment conditions.

The four fictitious benchmark examples in chapter 7 have illustrated the application of the method to cases with a varying degree of complexity, each highlighting different aspects of the method, including an example of a simple dike slope stability problem. All examples confirm that the results obtained with the simplified approach with fragility curves reasonably match the results by conventional reliability methods such as Monte Carlo simulation.

The present work is part of a larger development effort by Rijkswaterstaat to enable practitioners to use reliability updating in Dutch safety assessment practice. Besides this background report, case studies will be elaborated and a manual as well as software will be produced to provide practitioners with all necessary components. The application to the first two real life case studies with a realistic level of detail and complexity for a safety assessment of Dutch dikes is described in a separate report (Schweckendiek *et al.*, 2016).

We have the following main recommendations for further method development and operationalisation:

1 For further testing of the applicability and robustness of the approach it is necessary to investigate more case studies with varying characteristics which are representative for the majority of the dikes in the primary Dutch flood protection system. Additional cases will also help to obtain better insight into how much impact can be expected in which conditions. Obvious candidates are dikes in tidal estuaries, canal dikes and river dikes in the upstream area (in Dutch: bovenrivierengebied).
2 Reliability updating can also have a significant impact for other failure mechanisms dominated by epistemic (knowledge) uncertainties. This is typically the case for geotechnical failure modes such as internal erosion (piping).
3 The approach with fragility curves is an approximation which brings about limitations. Other methods should be investigated in order to use reliability updating to its full potential, including updating of the (joint) probability distribution of the basic random variables. While some more advanced reliability approaches may not be suitable for practical application, they can provide valuable insights and benchmark results for more practicable approximate methods. A particularly promising method is Subset Simulation, which may

be able to combine the advantages of using sampling methods directly with the limit state models and practical applicability.

4 The current description of the approach has shortcomings for situations with multiple external loads (e.g. water level and traffic load) with different scales of fluctuation in time. To address such conditions, a suitable probabilistic load model including temporal correlation needs to be implemented.

References

Bayes, T., 1763. "An Essay towards solving a Problem in the Doctrine of Chances." *Biometrika* 45: 293–315.

Calle, E. O. F., 2005. "Observed strength of dikes (in Dutch: Bewezen sterkte bij dijken)." *Geotechniek* 2005 (1).

Deltaprogramma, 2014. "Synthesedocument achtergrondrapport B1 bij Deltaprogramma 2015 - Synthesedocument deelprogramma Veiligheid."

Duinen, A., 2015. *Modelonzekerheidsfactoren Spencer-Van der Meij model ongedraineerder schuifsterkte*. Tech. rep.

ENW, 2009. *Technnisch Rapport Actuele Sterkte bij Dijken*. Tech. rep., Expertise Netwerk Waterveiligheid.

IVW, 2011. *Derde toets primaire waterkeringen*. Tech. rep., Inspectie van Verkeer en Waterstaat.

Kanning, W., 2012. "The weakest link." (July).

Kanning, W. and M. van der Krogt, 2016. *Pore water pressures for slope stability*. Tech. rep., Deltares memo number 1230090-034-GEO-0008.

Margo, D., A. Harkness and J. Needham, 2009. *Managing Our Water Retention Systems - Levee Screening Tool*. Tech. rep., USSD.

Rijkswaterstaat, 2014. *The National Flood Risk Analysis for the Netherlands*. Tech. rep., Rijkswaterstaat VNK2 Project Office.

Roscoe, K., A. Hanea and A. Vrouwenvelder, 2016. "Levee system reliability modeling: The length effect and Bayesian updating." *Structural Safety* (submitted.

Schweckendiek, T., 2014. *On Reducing Piping Uncertainties - A Bayesian Decision Approach*. Phd thesis, Delft University of Technology, Delft.

Schweckendiek, T., A. Teixeira, M. van der Krogt and W. Kanning, 2016. *Reliability updating for slope stability of dikes - Test cases report*. Tech. rep., Deltares report number 1230090-037-GEO-0001.

Schweckendiek, T., A. Vrouwenvelder and E. Calle, 2014. "Updating piping reliability with field performance observations." *Structural Safety* 47: 13–23. DOI 10.1016/j.strusafe.2013.10.002.

Schweckendiek, T., A. C. W. M. Vrouwenvelder, E. O. F. Calle, W. Kanning and R. B. Jongejan, 2012. "Target Reliabilities and Partial Factors for Flood Defenses in the Netherlands." In P. Arnold, G. A. Fenton, M. A. Hicks and T. Schweckendiek, eds., *Modern Geotechnical Codes of Practice - Code Development and Calibration*, pages 311–328. Taylor and Francis.

Straub, D., 2014. "Value of information analysis with structural reliability methods." *Structural Safety* 49 (July 2014): 75–85. DOI 10.1016/j.strusafe.2013.08.006, ISSN 01674730.

Straub, D. and I. Papaioannou, 2014. "Bayesian Updating with Structural Reliability Methods." *Journal of Engineering Mechanics* 141 (3): 014134–1–13.

Vanmarcke, E., 2011. “Risk of Limit-Equilibrium Failure of Long Earth Slopes: How it Depends on Length.” In *Proceedings of Geo-Risk 2011, Atlanta, USA*, 609. ASCE, Atlanta.

Vrouwenvelder, A. C. W. M., 2006. “Spatial effects in reliability analysis of flood protection systems.” In *International Forum on Engineering Decision Making*. Lake Louise, Canada.

Zhang, J., L. M. Zhang and W. H. Tang, 2011. “Slope Reliability Analysis Considering Site-Specific Performance Information.” *Journal of Geotechnical and Geoenvironmental Engineering* 137 (3): 227–238. ISBN 1090-0241/2011/3-227.238.

Zhang, L. M., 2004. “Reliability Verification Using Proof Pile Load Tests.” *Journal of Geotechnical Engineering and Geoenvironmental Engineering* 130 (11): 1203–1213.

APPENDIX

A Length effect prior and posterior

A.1 Problem statement

This example examines the differences between prior and posterior length effect (for the principles of the length effect refer to Kanning (2012) or Schweckendiek *et al.* (2012)). The proposed approach treats the length-effect in an implicit fashion through the reliability target assigned to a cross section (see section 2.3). The implicit assumption made is that the posterior length effect is not greater than the prior length effect, at least not greater than the length-effect accounted for in the derivation of the target reliability.

Similar to the explanation in section 2.3, we can account for the length effect in deriving the target probability of failure for a cross section by

$$P_{cs} < P_{cs,req} = \frac{P_{segment}}{1 + \frac{L}{l_{eq}}} = \frac{P_{segment}}{LE} \tag{A.1}$$

where we define the so-called length-effect factor LE as the ratio of the failure probabilities of a segment (of finite length) and of a cross section. In the remainder of this appendix we will examine how LE changes with reliability updating.

A.2 Approach

A.2.1 Limit state

To support the argument that the length effect is not (much) greater a-posteriori than a-priori we use the simplified limit state used throughout this report also for fragility curves:

$$g = h_c - h \tag{A.2}$$

where h_c is the critical water level (i.e. the water level at failure) and h the water level.

A.2.2 Spatial variability

In terms of spatial variability we suppose that the water level has a very long auto-correlation distance (practically infinite), allowing us to model h as a simple random variable. On the other hand, we model h_c as a one-dimensional (Gaussian) random field with scale of fluctuation θ_{h_c}.

A.2.3 Operational limit state definitions

For the analysis we simulate $i = 0...n$ discrete cross sections of the random field of the critical water level, resulting in the values $h_{c,i}$. The cross section analysis is done in the first cross section of the random field, i.e. with $h_{c,0}$.

The **failure** definitions for the cross section and for the random field are defined as follows:

- Cross section: $h_{c,0} < h$.
- Random field: $\cup_i h_{c,i} < h$ (i.e. failure of any cross section in the segment implies failure of the segment).

For the posterior analysis, we also need to define the **evidence** for both cross section and random field:

- Cross section: $h_{c,obs,0} > h_{obs}$
- Random field: $h_{c,obs,i} > h_{obs}$ (i.e. all cross sections in the realization of the random field need to comply with the observed water level for the evidence to be true).

In other words, we discard all random field realizations for which at least one point does not meet the (cross sectional) criterion.

These definitions allow us to compute the prior and posterior probabilities of failure for both the cross section and the random field.

A.2.4 Prior and posterior length-effect

The length-effect is defined here as the ratio of the probability of failure in a one-dimensional random field (i.e. segment) divided by the probability of failure of the cross section. Consequently, we can define the prior and posterior length effect, LE_{prior} and $LE_{posterior}$, as follows:

$$\begin{aligned} LE_{prior} &= P_{f,rf,prior}/P_{f,cs,prior} \\ LE_{posterior} &= P_{f,rf,posterior}/P_{f,cs,posterior} \end{aligned} \tag{A.3}$$

The prior probabilities of failure for a random field and $P_{f,rf,prior}$ and for a cross section $P_{f,cs,prior}$ can also be calculated or approximated with the following analytical expressions for the current example (assuming Gaussian random variables):

$$P_{f,cs,prior} = \Phi\left(-\beta_{prior}\right) \quad with \quad \beta_{cs,prior} = \frac{\mu_{h_c} - \mu_h}{\sqrt{\sigma_{h_c}^2 + \sigma_h^2}} \tag{A.4}$$

$$P_{f,rf,prior} = P_{f,cs,prior} \cdot \left(1 + \frac{L}{l_{eq}}\right) \quad with \quad l_{eq} = \frac{\theta_{h_c}\sqrt{\pi}}{\beta_{prior}} \tag{A.5}$$

where L is the segment length and l_{eq} is the equivalent auto-correlation length.

A.3 Example (base case)

First, the prior and posterior reliability index and length effect are calculated for a base case, from which variations of several parameters will be made as sensitivity studies in the sequel. The input parameters for the base case are presented in Tables B.1 and A.2.

Symbol	Unit	Description	Distribution	Mean	Standard deviation
h_c	[m]+REF	critical water level	Normal	9	$\sqrt{2}$
h	[m]+REF	water level	Normal	2	$\sqrt{2}$
h_{obs}	[m]+REF	observed water level	Deterministic	7	N/A
$h_{c,obs}$	[m]+REF	critical water level at observation	Normal	10	$\sqrt{2}$

Table A.1: *Input parameters for the base case of the study into the prior and posterior length-effect*

θ_{h_c}, correlation length (in x-direction)	100 m
L (dike length)	1000 m
cellsize	20 m

Table A.2: *Correlation length, dike length and cellsize for random field generation in the base case*

Table A.3 contains the results of the prior and posterior reliability index and the length-effect. The following observations can be made:

- The reliability index for the random field is lower than for the cross section (which is the whole point of the length-effect);
- The analytical expressions give good approximations of the prior reliability indices;
- The posterior length effect is roughly the same as the prior length effect in this particular case.

	β	$\beta_{analytical}$	β_{rf}	$\beta_{rf,analytical}$	LE
prior	3.49	3.50	2.64	2.59	17.5
posterior	3.87	-	3.13	-	15.8

Table A.3: *Prior and posterior reliability indices and length-effect factors for the base case*

Figure A.1 illustrates four realizations of the random fields for this example. Note that the realizations of the assessment conditions and the observation conditions are fully correlated, resulting in virtually parallel fields.

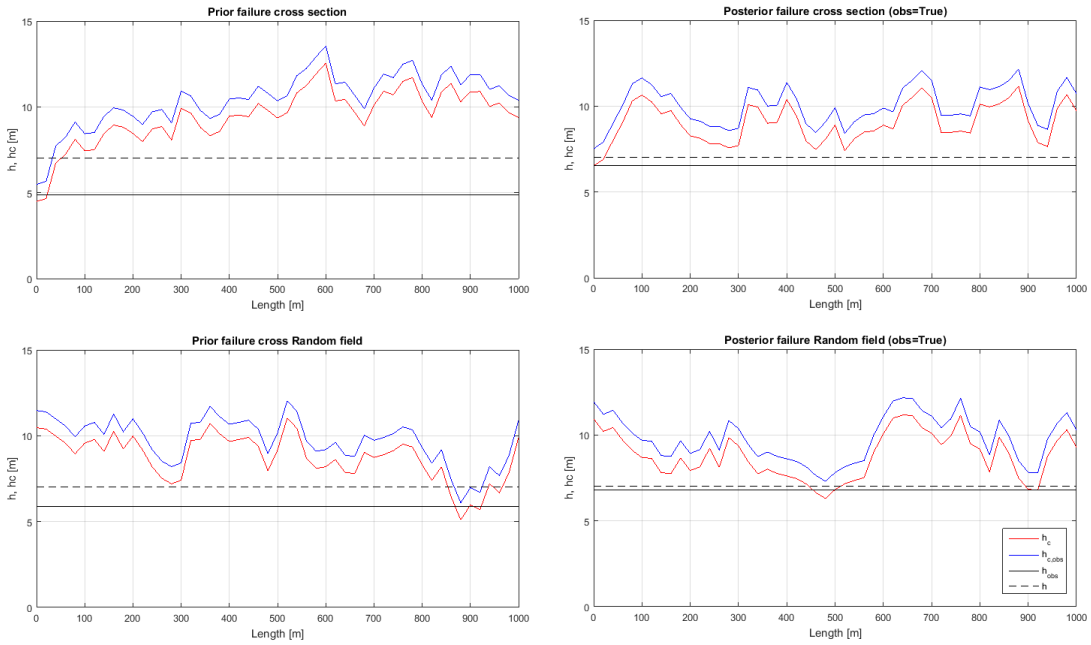


Figure A.1: *Prior and posterior length effect factor and reliability indices*

A.3.1 Variation 1: Correlation length

Table A.4 contains the results for variations of the correlation length (θ_{h_c}). The length effect LE is plotted in Figure A.2. We observe the following:

- The length effect decreases with increasing correlation length (well-known effect).
- The posterior length effect is less than or roughly equal to the prior, more specifically:
  - for low correlation lengths the posterior length effect is lower than the prior;
  - for high correlation lengths, the posterior length effect is roughly equal to the prior.

Notice that for increasing correlation length the length-effect virtually vanishes and so do the differences between prior and posterior. The interesting range for stability analyses will be between roughly 50 m and 200 m, though.

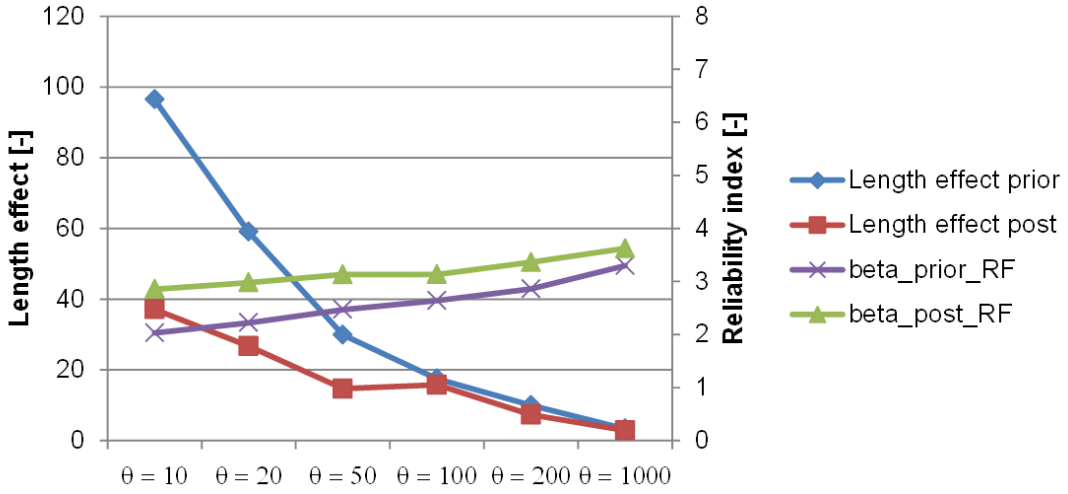


Figure A.2: *Prior and posterior length effect factors and reliability indices for variations of the correlation length ($\mu_{h_c} = 9$ and $\sigma_{h_c} = \sqrt{2}$)*

		β	β_{rf}	LE
10	prior	3.51*	2.03	96.5
(cellsize=2)	posterior	3.85*	2.85	37.1
20	prior	3.51	2.22	59.0
(cellsize=4)	posterior	3.87	2.98	26.7
50	prior	3.51	2.47	29.9
(cellsize=10)	posterior	3.85	3.13	14.7
100	prior	3.49	2.64	17.5
(cellsize=20)	posterior	3.87	3.13	15.8
200	prior	3.52	2.86	10.0
(cellsize=40)	posterior	3.88	3.36	7.3
1000	prior	3.62	3.30	3.3
(cellsize=200)	posterior	3.89	3.62	2.9

*) Variation calculated with a maximum $COV(P_f)$ of 0.1, else 0.05.

Table A.4: *Prior and posterior length effect factors and reliability indices for variations of the correlation length ($\mu_{h_c} = 9$, $\sigma_{h_c} = \sqrt{2}$ and $h_{obs} = 7$)*

A.3.2 Variation 2 - Higher standard deviations

Variation 2 involves variations with higher standard deviations of h_c. In order to have a high reliability level, we need rather higher mean values as well. Notice that relatively flat beta-h curves as sometimes found for dike stability often represent a relatively large uncertainty in the resistance term (i.e. the critical water level), which would be reflected by large standard deviations. The results of the analysis are shown in Figure A.3 which depicts the length effect and reliability for several variations of mean and standard deviation. We observe in most cases the posterior length effect is comparable or slightly less than the prior, except for the combination with a very high standard deviation, where the posterior length-effect factor about twice as large as the prior length-effect factor.

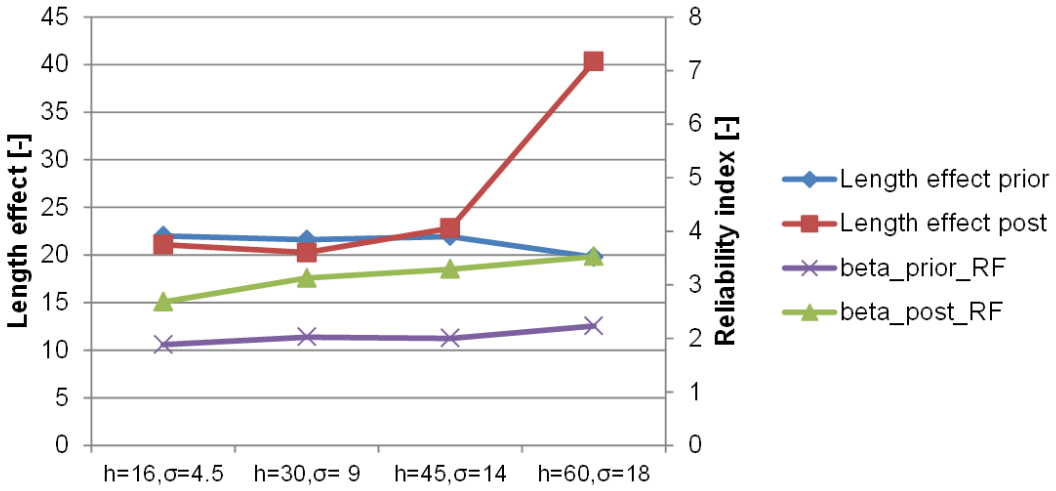


Figure A.3: *Prior and posterior length effect factors and reliability indices for variations of the mean and standard deviation of the resistance (correlation length 100 m)*

A.4 Conclusion

Overall we see that the posterior length-effect is smaller or comparable to the prior, except for very high standard deviations of the resistance (critical water level). The implications for the proposed approach are discussed in section 2.3.

B Reliability updating with discrete scenarios

This appendix describes algorithms and benchmarks for working with discrete scenarios.

B.1 Algorithms Integrated Monte Carlo approach (IMC)

This section describes algorithms to implement the reliability updating approach with fragility curves as explained in the main report (section 5.3.3) with discrete scenarios using a fully integrated Crude MCS approach to evaluate

$$\hat{P}(F|\varepsilon) = \frac{\sum_i F_i \cdot \varepsilon_i}{\sum_i \varepsilon_i} = \frac{\sum_i \mathbf{1}[h_{c,i} < h_i] \cdot \mathbf{1}[h_{c,obs,i} > h_{obs}]}{\sum_i \mathbf{1}[h_{c,obs,i} > h_{obs}]} \tag{B.1}$$

N assessment scenarios and M observation scenarios (uncorrelated)

For each MCS realization i ($i = 1...n$):

1 sample water level h_i
2 sample assessment scenario number j ($j = 1...N$)
3 sample observation scenario number k ($k = 1...M$)
4 sample standard normally distributed variable u and with it
  4.1 determine $h_{c,i}$ from the fragility curve of assessment scenario E_j
  4.2 determine $h_{c,obs,i}$ from the fragility curve of assessment scenario $E_{obs,k}$
5 evaluate $F_i = \mathbf{1}[h_{c,i} < h_i]$ and $\varepsilon_i = \mathbf{1}[h_{c,obs,i} > h_{obs}]$

N assessment scenarios and M observation scenarios (all fully correlated)

For each MCS realization i ($i = 1...n$):

1 sample water level h_i
2 sample assessment scenario j ($j = 1...N$)
3 observation scenario is equal to assessment scenario $k = j$
4 sample standard normally distributed variable u and with it
  4.1 determine $h_{c,i}$ from the fragility curve of assessment scenario E_j
  4.2 determine $h_{c,obs,i}$ from the fragility curve of assessment scenario $E_{obs,k}$
5 evaluate $F_i = \mathbf{1}[h_{c,i} < h_i]$ and $\varepsilon_i = \mathbf{1}[h_{c,obs,i} > h_{obs}]$

N assessment scenarios and M observation scenarios (mixed fully correlated and uncorrelated)

For each MCS realization i ($i = 1...n$):

1 sample water level h_i
2 sample assessment scenario j ($j = 1...N$)
3 sample observation scenario number k ($k = 1...M$)
4 IF observation scenario is fully correlated with assessment scenario, THEN $k = j$
5 sample standard normally distributed variable u and with it
  5.1 determine $h_{c,i}$ from the fragility curve of assessment scenario E_j
  5.2 determine $h_{c,obs,i}$ from the fragility curve of assessment scenario $E_{obs,k}$
6 evaluate $F_i = \mathbf{1}[h_{c,i} < h_i]$ and $\varepsilon_i = \mathbf{1}[h_{c,obs,i} > h_{obs}]$

B.2 Benchmark examples

The aim of the present examples is to illustrate the theory of handling discrete scenarios as described in chapter 5 for a simplified example. Therefore, this benchmark example considers a base case similar to the one in Example 1, but now with 3 discrete scenarios for the probability distribution of the resistance, both for assessment and observation. The analysis follows the the two-stage procedure which is compared to the results of an integrated Monte Carlo analysis (IMC) as presented in section 5.3.3 in order to demonstrate their equivalence. Different assumptions for the correlation between scenarios are considered in order to illustrate the effects and differences.

B.2.1 Input data and example setup

The performance function is:

$$g = h_{c,i} - h \tag{B.2}$$

with $h_{c,i}$ being the critical water level (i.e. the water level at which the dike fails) conditional on scenario i and h the water level. The water level follows a standard normal distribution ($h \sim N(0,1)$). The distributions of the critical water levels for the different scenarios are defined in Table B.1. Note that the resistance for the observation conditions is estimated to be slightly higher than for the assessment.

	Assessment			**Observation**		
Scenario	μ	σ	Scenario probability	μ	σ	Scenario probability
1	2.8	1.0	0.2	3.0	1.0	0.2
2	3.2	0.5	0.5	3.5	0.5	0.5
3	3.6	0.5	0.3	4.0	0.5	0.3

Table B.1: *Parameters for the critical water level of three resistance scenarios for assessment and observation*

As discussed in section 3.3, time-invariant scenarios and scenarios representing aleatory uncertainty require different treatment. We will consider three distinct cases:

- **Case A:** all scenarios are time-invariant (e.g. a case where all scenarios represent stratification scenarios),
- **Case B:** all scenarios represent aleatory randomness (e.g. a case where all scenarios represent different geohydraulic responses),
- **Case C:** a mixture of the above (see Table B.2).

			Observation scenario		
Assessment scenario	Stratification	Geohydraulic case	1	2	3
1	A	1	1	1	0
2	A	2	1	1	0
3	B	2	0	0	1

Table B.2: *Dependence of scenarios in assessment and observation. 1 implies that the combination of the occurrence of a scenario combination is possible; 0 implies impossibility of the combination.*

In order to include the sensitivity to the observed load, we consider two different observed water levels: a water level with an annual exceedence probability of 1/100 (2.3 m+NAP) and another with annual exceedence probability 1/1,000 (3.9 m+NAP). For the sake of computational efficiency all calculations were carried out using Monte Carlo with Importance Sampling with a sampling distribution for the water level of $N(1,2)$ and a sample size of 10^6.

B.2.2 Posterior analysis with two-stage procedure and integrated Monte Carlo

Case A: Time-invariant scenarios

For case A, the scenarios are time-invariant, meaning that scenario 1 in the assessment corresponds to scenario 1 in the observation and so on. Table B.3 presents the prior and posterior reliability estimates for the two different water level observations. The results of both methods (two-stage and IMC) are in good agreement.

	Prior		**Posterior**			
			Observation = 2.3 m		*Observation = 3.9 m*	
	Two-stage	IMC	Two-stage	IMC	Two-stage	IMC
Scenario 1	1.980	1.979	2.743	2.744	4.022	4.018
Scenario 2	2.862	2.862	2.906	2.905	3.815	3.816
Scenario 3	3.220	3.220	3.226	3.225	3.807	3.805
All scenarios	**2.394**	**2.393**	**2.899**	**2.899**	**3.841**	**3.840**

Table B.3: *Prior and posterior reliability indices for case A with time-invariant scenarios*

Case B: Aleatory scenarios

Case B uses the same parameters as case A, except we assume the scenarios to represent aleatory uncertainty, implying that realizations in the assessment are independent of realizations in the assessment. Table B.4 shows the results in terms of the prior and posterior reliability indices. Compared to case A case we observe the overall effect of reliability updating to be less, which is plausible. However, this is not the case for all scenarios individually; for scenarios with relatively high reliability the effect of updating is larger than in case A. The reason is that now there is a non-zero probability that the observation originated from a scenario with a relatively high prior reliability index, which implies a significant load effect, resulting in a relatively strong updating effect.

	Prior		**Posterior**			
			Observation = 2.3 m		*Observation = 3.9 m*	
	Two-stage	IMC	Two-stage	IMC	Two-stage	IMC
Scenario 1	1.980	1.979	2.129	2.129	3.343	3.346
Scenario 2	2.862	2.862	2.953	2.954	3.579	3.580
Scenario 3	3.220	3.220	3.317	3.315	3.962	3.960
All scenarios	**2.394**	**2.393**	**2.520**	**2.520**	**3.517**	**3.517**

Table B.4: *Prior and posterior reliability indices for case B without correlation between assessment and observation*

Case C: Mixed time-invariant and aleatory scenarios

In this case we contemplate a mixture of time-invariant and aleatory scenarios as defined in Table B.2. When elaborating on the scenario probabilities, the scenarios become in fact combinations of a scenario for the assessment and the observation, meaning that there are five possible scenario combinations in this case. The probability of such a combination is determined by $P(E_i \cap E_{obs,j}) = P(E_i|E_{obs,j})P(E_{obs,j})$ resulting in the values presented in Table B.5.

Scenario combination $i-j$	$P(E_i)$	$P(E_{obs,j})$	$P(E_{obs,j}\|E_i)$	$P(E_i \cap E_{obs,j})$
$1-1$	0.2	0.2	0.29	0.06
$1-2$	0.2	0.5	0.71	0.14
$2-1$	0.5	0.2	0.29	0.15
$2-2$	0.5	0.5	0.71	0.36
$3-3$	0.3	0.3	1	0.3
				$\Sigma = \mathbf{1}$

Table B.5: *Scenario combinations and corresponding probabilities for case C*

The resulting posterior reliability indices are shown in Table B.6. The overall updating effect is, as expected, in between cases A and B.

	Prior		**Posterior**			
			Observation = 2.3 m		*Observation = 3.9 m*	
	Two-stage	IMC	Two-stage	IMC	Two-stage	IMC
$\beta_{S1,1}$	1.979	1.979	2.743	2.746	4.024	4.026
$\beta_{S1,2}$	1.979	1.979	2.052	2.053	3.945	3.937
$\beta_{S2,1}$	2.862	2.862	3.241	3.244	3.855	3.856
$\beta_{S2,2}$	2.862	2.862	2.907	2.906	3.815	3.816
$\beta_{S3,3}$	3.219	3.219	3.226	3.226	3.806	3.806
β_{total}	**2.393**	**2.394**	**2.564**	**2.565**	**3.843**	**3.844**

Table B.6: *Prior and posterior reliability indices for a case C (mixed correlation).*

B.2.3 Discussion

The results show in all cases that the two-stage procedure and the integrated Monte Carlo approach produce the same results. The advantage of the two-stage approach is that all individual combinations of assessment and observation scenarios can be contemplated individually and the number of the Monte Carlo samples per combination can be adjusted per combination.

The differences between the cases highlight the importance of correlation assumptions for discrete scenarios. The overall reliability updating effect is larger for correlated scenarios than for non-correlated ones. However, per scenario the effects can be different as pointed out in the discussion of case B.

C Examples and Benchmarks

Example 1: $g = h - h_c$ (Matlab code incl. plots)

```
%% Benchmark example reliability updating (g = hc - h)
%  by Timo Schweckendiek
%  version: 29 March 2016
clear; close all; clc;

%% Input Parameters

% Prior distribution of the critical water level (resistance, Normal distributed)
mu_hc = 6;           % mean value
sigma_hc = sqrt(2);  % standard deviation

% Prior distribution of the water level (load, Normal distributed)
mu_h = 2;            % mean value
sigma_h = sqrt(2);   % standard deviation

% Observed water level (load)
h_obs = 5;           % observed value
% P_exc = 1 - normcdf(h_obs, mu_h, sigma_h); % exceedance probability

% Observed resistance parameters
delta_hc = 1.0;      % difference between observation and assessment
rho = 1;             % correlation resistance between observation and future

% Number of realisations in Monte Carlo analyses
n = 1e+5;

% water levels for beta-h curve (discretization)
hgrid = [mu_h (mu_hc+mu_h)/2 mu_hc];

% Dummy variable (for plots etc.)
theta = mu_h-4*sigma_h : 0.1 : mu_hc+4*sigma_hc;

%% Prior distributions and observation
figure('name', 'Priors with observed water level');
plot(theta, normpdf(theta, mu_hc, sigma_hc), 'g-'); grid on; box on; hold on;
plot(theta, normpdf(theta, mu_h, sigma_h), 'r-');
plot([h_obs h_obs], [0 0.4], 'b-');
plot(theta, normpdf(theta, mu_hc + delta_hc, sigma_hc), 'b--');
xlabel('h, h_c'); ylabel('probability density');
legend('PDF resistance f(h_c)','PDF water level f(h)',...
    'observed water level h_{obs}','PDF resistance at observation f(h_{c,obs})',...
    'Location', 'NorthWest');


%% Prior probability of failure with Monte Carlo simulation (MCS)

% Random realisations of resistance and load
var_hc = sigma_hc .* randn(n, 1);
hc = mu_hc + var_hc;
h = mu_h + sigma_h .* randn(n, 1);

% Performance (limit state) function and probability of failure
g = hc - h; F = (g<0);
Pf = sum(F)/n; beta = norminv(1-Pf)

%% Scatter plot prior MCS
figure('name', 'Prior MCS');
```

```
plot([-2 12], [-2 12], 'k-'); hold on; grid on;
plot(hc(g<0), h(g<0), 'r.');
plot(hc(g>=0), h(g>=0), 'b.');
xlabel('h_c'); ylabel('h');
legend('limit state','failure', 'non-failure');

%% Prior fragility curve
figure('name', 'Prior Fragility');
[hc_cdf,x] = ecdf(hc);
plot(x, hc_cdf, 'g-'); hold on; grid on; box on;
plot(theta, normpdf(theta,mu_h,sigma_h), 'r-');
xlabel('h, h_c'); ylabel('probability (density)'); % P(F|h), F(h_c), f(h)
axis([min(theta) max(theta) 0 1]); hold off;
legend('fragility curve P(F|h)=F(h_c)','PDF water level f(h)',...
    'Location', 'NorthWest');


%% Prior analysis with approximated fragility curves (MCS)

% Determine beta per discrete water level
betagrid = norminv(1-normcdf(hgrid,mu_hc,sigma_hc));

% Plot beta-h curve
figure('name', 'Beta-h');
plot(hgrid, betagrid, 'ro'); hold on; grid on; box on;
plot(theta, interp1(hgrid,betagrid,theta,'linear','extrap'),'r--');
xlabel('h'); ylabel('\beta');
axis([min(theta) max(theta) -1 4]);
legend('fragility points','linear interpolation',...
    'Location', 'NorthEast'); hold off;

% Simulate hc with fragility curve method
u = normrnd(0,1,n,1); % standard normal
hc_fc = interp1(betagrid,hgrid,u,'linear','extrap');

% Performance (limit state) function and probability of failure
F_fc = (hc_fc < h);
Pf_fc = sum(F_fc)/n; beta_fc = norminv(1-Pf_fc)


%% Histogram of realizations from fragility curve
figure('name', 'Hist');
[nhc,xout] = hist(hc_fc,50); %Compute histogram
bar(xout,nhc); hold on; grid on; box on;
plot(theta, normpdf(theta,mu_hc,sigma_hc)*trapz(xout,nhc),'r-');
xlabel('h_c'); ylabel('bin count');
legend('histogram','original PDF (scaled)',...
    'Location', 'NorthWest'); hold off;


%% Posterior probability of failure with MCS (direct method) - exact

% Generate fully correlated resistance at observations
hc_obs = mu_hc + delta_hc + var_hc;

% Reliability updating
Eps = (hc_obs > h_obs);
Pf_post = sum(F & Eps) / sum(Eps); beta_post = norminv(1-Pf_post)


%% Posterior probability of failure with MCS (direct method) - exact

% Fragility curve at observation (difference in hc)
hgrid_obs = hgrid + delta_hc;

% Simulate hc_obs with fragility curve method (use same u as for assessment!)
```

```
hc_obs_fc = interp1(betagrid,hgrid_obs,u,'linear','extrap');

% Posterior probability of failure
Eps_fc = (hc_obs_fc > h_obs); % % observation / evidence (binary vector)
Pf_post_fc = sum(F_fc & Eps_fc) / sum(Eps_fc);
beta_post_fc = norminv(1-Pf_post_fc)


%% Beta-h curves for observation and assessment
figure('name', 'Beta-h obs');
plot(hgrid, betagrid, 'ro'); hold on; grid on; box on;
plot(theta, interp1(hgrid,betagrid,theta,'linear','extrap'),'r--');
plot(hgrid_obs, betagrid, 'bo');
plot(theta, interp1(hgrid_obs,betagrid,theta,'linear','extrap'),'b-.');
xlabel('h'); ylabel('\beta');
axis([min(theta) max(theta) -1 4]);
legend('fragility points assessment','beta-curve assessment',...
    'fragility points observation','beta-curve observation',...
    'Location', 'SouthWest'); hold off;


%% Comparison posterior and prior resistance

figure('name', 'Post hc');

subplot(1,2,1)
bar(theta, hist(hc(Eps), theta)./(sum(Eps)/n)); grid on; box on; hold on;
plot(theta, hist(hc, theta), 'r-'); hold off;
xlabel('h_c'); ylabel('# of realisations');
legend('posterior','prior', 'Location', 'NorthWest');
title('exact');

subplot(1,2,2)
bar(theta, hist(hc_fc(Eps_fc),theta)./(sum(Eps_fc)/n)); grid on; box on; hold on;
plot(theta, hist(hc_fc, theta), 'r-'); hold off;
xlabel('h_c'); ylabel('# of realisations');
legend('posterior','prior', 'Location', 'NorthWest');
title('with fragility curves');
```

Example 2: $g = m_B \frac{L}{c} - h$ (Matlab code incl. plots)

```
%% Benchmark Bligh Rule
%  by Timo Schweckendiek and Katerina Rippi
%  version: 28 April 2016
clear; close all; clc;

%% Input Parameters

% Prior moments of the model factor and seepage length
mu_m = 1.76;         % mean
sigma_m = 1.69;      % standard deviation
mu_L = 50;         % mean
sigma_L = 2.5;     % standard deviation
c = 18; %[-] percolation coefficient (deterministic)

% Distribution parameters of the load (water level)
alpha_h = 0.53; %location parameter in gumbel (mu)
beta_h = 0.406;  %scale paramater in gumbel (sigma)

%  Observed water level (load) and resistance parameters (from new estimations)
% m_obs = 1.2;    % [-] (model factor)
% L_obs = 50;   % [m] (seepage length)
h_obs = 2.4;       % observed value
```

```
delta_m = 0.0;  % difference between observation and assessment
delta_L = 10;    % difference between observation and assessment

% Number of realisations in Monte Carlo analyses
n = 1e+5;


%% Prior probability of failure with Monte Carlo simulation (MCS)

% Random realisations of the model factors
[lambda_m, zeta_m] = lognpar(mu_m, sigma_m);
m = lognrnd(lambda_m, zeta_m, n, 1); % mu lognormally distributed
L = normrnd(mu_L, sigma_L, n, 1); % L normally distributed
m_obs = lognrnd(lambda_m+delta_m, zeta_m, n, 1); % mu lognormally distributed
L_obs = normrnd(mu_L+delta_L, sigma_L, n, 1); % L normally distributed

% Resistance hc
hc = m.*L./c;
mu_hc = mean(hc);
sigma_hc = std(hc);

% Random realisations of the load, h
h = gevrnd(0,beta_h, alpha_h, n, 1);

% Illustration of observed value with prior distributions
figure('name', 'Prior PDFs');
m_grid = 0:0.05:6;
L_grid = 30:0.3:80;
subplot(1,2,1)
plot(m_grid, ksdensity(m, m_grid), 'b-', 'LineWidth', 2);
hold on; grid on; box on;
legend('PDF model uncertainty, f(m)');
xlabel('m [-]'); ylabel('probability density');
subplot(1,2,2)
plot(L_grid, ksdensity(L, L_grid), 'b-', 'LineWidth', 2);
hold on; grid on; box on;
plot(L_grid, ksdensity(L_obs, L_grid), 'r-', 'LineWidth', 2);
legend('PDF seepage length, f(L)','PDF seepage length at observation, f(L_{obs})');
xlabel('L [m]'); ylabel('probability density');

figure('name', 'Priors with observed head differences')
hc_grid = -2.0:0.05:10.0;
plot(hc_grid, ksdensity(hc, hc_grid), 'k-', 'LineWidth', 2);
hold on; grid on; box on;
plot(hc_grid, ksdensity(h, hc_grid), 'g-', 'LineWidth', 2);
hold on; grid on; box on;
plot([h_obs h_obs], [0 1], 'b--', 'LineWidth', 2);
legend('PDF critical head difference, f(h_c)', 'PDF head difference, f(h)', 'observed head differe
xlabel('h, h_c [m]'); ylabel('probability density');
hold off;

% Performance (limit state) function and probability of failure
g = hc - h;
Pf = sum(g<0) / n
beta = norminv(1-Pf)

% Scatter plot
figure('name', 'Prior MCS');
plot([-2 8], [-2 8], 'k-');
hold on; grid on;
plot(hc(g<0), h(g<0), 'r.');
plot(hc(g>=0), h(g>=0), 'b.');
xlabel('h_c [m]'); ylabel('h [m]');
legend('limit state','failure', 'non-failure');
axis([-1 8 -1 8]); hold off;
```

```
%% Posterior probability of failure with MCS (direct method)

%Generate fully correlated resistance at observations
mc_obs = m + delta_m;
Lc_obs = L + delta_L;
hc_obs = mc_obs.*Lc_obs./c;

% Observation function
Eps = h_obs-hc_obs;

% Posterior probability of failure
Pf_post = sum(g<0 & Eps<0) / sum(Eps<0)
beta_post = norminv(1-Pf_post)

% Scatter plot
figure('name', 'Posterior MCS');
plot(hc_obs(g>=0), h(g>=0), 'b.'); hold on; grid on;
plot(hc_obs(Eps>=0), h(Eps>=0), 'g.');
plot(hc_obs(g<0 & Eps<0), h(g<0 & Eps<0), 'r.');
%Plot limit state:this line has a 45 degrees inclination only if the
%difference between observation and assessment (deltas) is zero
plot([0.85 10000], [0 10000], 'k-'); %
plot([h_obs h_obs],[-1 10000],'k--');
xlabel('h_{c,obs} [m]'); ylabel('h [m]');
legend('non-failure, posterior', 'no match with observation',...
    'failure, posterior',...
    'limit state', 'observation');
axis([-1 8 -1 6]); hold off;

%Fragilty curves
figure('name', 'Prior and Posterior fragility curves')
plot(hc_grid, ksdensity(hc_obs(Eps<0), hc_grid, 'support','positive','function', 'cdf'), '
hold on; grid on; box on;
plot(hc_grid, ksdensity(hc, hc_grid, 'support','positive','function','cdf'), 'r-', 'LineWi
legend('F_{h_c|\epsilon} (h)', 'F_{h_c} (h)');
xlabel('h [m]'); ylabel('F(h)');

% Prior vs. posterior m and L parameters
figure('name', 'Prior vs. Posterior m and L parameters');
subplot(1,2,1)
hist3([m L], {0:.1:6; 40:1:60});
xlim([0 4])
set(gcf,'renderer','opengl');
set(get(gca,'child'),'FaceColor','interp','CDataMode','auto');
view(2); xlabel('m [-]'); ylabel('L [m]');
colorbar;
subplot(1,2,2)
hist3([m(Eps<0) L(Eps<0)], {0:.1:6; 40:1:60});
xlim([0 4])
set(gcf,'renderer','opengl');
set(get(gca,'child'),'FaceColor','interp','CDataMode','auto');
view(2); xlabel('m [-]'); ylabel('L [m]');
colorbar;

%% Prior probability of failure with fragility curves (FC)

%===========================
% hgrid_fc with more points
%===========================
mult = 1;
steps = 6; % number of points
for i = 1:steps
    hgrid_fc(i) = mu_hc-sigma_hc*mult;
    mult = mult - 1/steps;
end
```

```
%hgrid_fc = [mu_hc-sigma_hc mu_hc-sigma_hc/2 mu_hc];
for i = 1:length(hgrid_fc)
    Pex_grid(i) = sum(hc < hgrid_fc(i))/n;
end
betagrid = norminv(1-Pex_grid);

figure('name', 'Beta-h');
plot(hgrid_fc, betagrid, 'ro'); hold on; grid on; box on;
plot(hc_grid, interp1(hgrid_fc,betagrid,hc_grid,'linear','extrap'),'r--');
xlabel('h [m]'); ylabel('\beta');
axis([min(hc_grid) max(hc_grid) -1 5]);
legend('fragility points', 'Fragility curve',...
    'Location', 'NorthEast'); hold off;

% Simulate hc with fragility curve method
u = normrnd(0,1,n,1); % standard normal
hc_fc = interp1(betagrid,hgrid_fc,u,'linear','extrap');

% Performance (limit state) function and probability of failure
g_fc = (hc_fc < h);
Pf_fc = sum(g_fc)/n; beta_fc = norminv(1-Pf_fc)

%% Posterior probability of failure with fragility curves (FC)

%Resistance at observation (difference in m and L)
hc_delta = (m + delta_m).*(L + delta_L)./c;
mu_hc_delta = mean(hc_delta);
sigma_hc_delta = std(hc_delta);
%hgrid_obs = [mu_hc_delta-sigma_hc_delta mu_hc_delta-sigma_hc_delta/2 mu_hc_delta];

%============================
% hgrid_obs with more points
%============================
mult = 1;
steps = 6; % number of points
for i = 1:steps
    hgrid_obs(i) = mu_hc_delta-sigma_hc_delta*mult;
    mult = mult - 1/steps;
end

%Simulate hc_obs with fragility curve method
hc_obs_fc = interp1(betagrid, hgrid_obs, u, 'linear', 'extrap');

Eps_fc = (hc_obs_fc>h_obs);
Pf_post_fc = sum(g_fc & Eps_fc) / sum(Eps_fc);
beta_post_fc = norminv(1-Pf_post_fc)

% Beta-h curves for observation and assessment
figure('name', 'Beta-h obs');
plot(hgrid_fc, betagrid, 'ro'); hold on; grid on; box on;
plot(hc_grid, interp1(hgrid_fc,betagrid,hc_grid,'linear','extrap'),'r--');
plot(hgrid_obs, betagrid, 'bo');
plot(hc_grid, interp1(hgrid_obs,betagrid,hc_grid,'linear','extrap'),'b-.');
xlabel('h [m]'); ylabel('\beta');
axis([min(hc_grid) max(hc_grid) -1 5]);
legend('fragility points assessment','beta-curve assessment',...
    'fragility points observation','beta-curve observation',...
    'Location', 'SouthWest'); hold off;
```