

# Hydra Ring Scientific Documentation

Ferdinand Diermanse  
Kathryn Roscoe  
Juliana Lopez de la Cruz  
Henri Steenbergen  
Ton Vrouwenvelder

1206006-004



**Title**  
Hydra Ring Scientific Documentation

<b>Client</b>	<b>Project</b>	<b>Reference</b>	<b>Pages</b>
Rijkswaterstaat	1206006-004	1206006-004-ZWS-0001	249

**Keywords**  
System reliability, failure probability, dike ring, probabilistic techniques, failure mechanisms, load models, flood risk

**Summary**

This document describes the scientific background of Hydra-Ring, a probabilistic model used for the computation of failure probabilities for interconnected systems of flood defences. It includes background on the probabilistic techniques, the way in which the statistics and correlations of the hydraulic load and strength variables are handled, and the built-in failure mechanism models. The current version of this document is a preliminary work-in-progress document that is largely intended for developers and programmers in the development phases of Hydra-Ring.

The most relevant topics in the scientific documentation are the probabilistic techniques, including techniques for combining failure probabilities, the correlation models which are used to describe the interdependence of load variables, the statistical distributions which are used to describe the load and strength variables, and the description of hydraulic load models. Examples of load models in the Netherlands are presented in appendices.

Version	Date	Author	Initials	Review	Initials	Approval	Initials
	Jan 2013	F. Diermanse					

**State**  
draft

This is a draft report, intended for discussion purposes only. No part of this report may be relied upon by either principals or third parties.



## Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Background	5
1.2	Purpose and scope of this document	5
1.3	Global setup of Hydra-Ring	5
1.4	The role of Hydra-Ring in the safety assessment procedure of flood defences in the Netherlands	7
1.5	Predecessors of Hydra-Ring	8
1.6	Outline	10
<b>2</b>	<b>Probabilistic computation techniques for system reliability</b>	<b>11</b>
2.1	Introduction	11
2.2	Basic techniques and terminology	11
2.2.1	Systems and components	11
2.2.2	Formulation of the probability of failure of a single component	12
2.2.3	Conversion between standard normal variables (U-domain) and real variables (X-domain)	14
2.2.4	Reliability index $\beta$	16
2.2.5	Linearisation of Z-functions	19
2.3	Failure probability for a single component	22
2.3.1	Introduction	22
2.3.2	Numerical Integration	23
2.3.3	Crude Monte Carlo	25
2.3.4	Monte Carlo Importance Sampling	27
2.3.5	Monte Carlo directional sampling	30
2.3.6	First-order reliability method (FORM)	32
2.3.7	Computing $\alpha$ values for other methods than FORM	41
2.3.8	Rationale	43
2.4	Combining failure probabilities for components - generic methods	45
2.4.1	Introduction	45
2.4.2	Combining n components: the Hohenbichler method	45
2.4.3	Upscaling for systems with identical components: numerical integration with constant correlation	64
2.4.4	System analysis with Monte Carlo methods or numerical integration	72
2.4.5	Techniques for time and space dependent processes	73
2.5	Combining failure probabilities of components – choices and implementations in Hydra-Ring	77
2.5.1	General overview	77
2.5.2	Temporal upscaling	77
2.5.3	Combining over wind directions	87
2.5.4	Closure scenarios for flood barriers	94
2.5.5	Spatial upscaling - from cross section to flood defence segment	101
2.5.6	Combination order over mechanisms and flood defense segments	106
2.5.7	Order of combining: wind directions	108
2.6	Procedure for computing the failure probability of a flood defence system	109
2.6.1	Introduction	109
2.6.2	Procedure for FBC	109
2.6.3	Procedure for NTI	115

2.6.4	Procedure for APT	116
<b>3</b>	<b>Hydraulic load models: generic set-up</b>	<b>117</b>
3.1	Introduction	117
3.2	Hydraulic load models in probabilistic failure computations	117
3.2.1	Introduction	117
3.2.2	Hydraulic load models in probabilistic failure computations for single components	118
3.2.3	Generic, modular set-up of load models in Hydra-Ring	120
3.3	Statistical distribution functions of random variables	121
3.3.1	Generic description of the application of distribution functions	121
3.3.2	Standard distribution functions and parameter values	127
3.3.3	Temporal scaling of statistical distribution functions	127
3.3.4	Statistics of durations	129
3.4	Correlation models	131
3.4.1	Introduction	131
3.4.2	Correlations between random variables	132
3.4.3	Correlation in time and space	137
3.5	Hydrodynamic models	138
3.5.1	Simulating water levels and waves in an entire water system	138
3.5.2	Models for quantification of local (corrections on) hydraulic loads	139
3.5.3	Transformation module for a breakwater	143
3.5.4	Spatial interpolation methods	146
<b>4</b>	<b>Hydraulic load models for primary water systems in the Netherlands</b>	<b>147</b>
4.1	Introduction	147
4.2	System characteristics	149
4.2.1	Upper rivers	149
4.2.2	Rhine Meuse delta	149
4.2.3	Lakes	149
4.2.4	Lake delta	149
4.2.5	Coastal zone; dikes	150
4.2.6	Oosterschelde	150
4.2.7	Coastal zone; dunes	150
4.3	Random variables	150
4.4	Load model for the upper rivers and tidal rivers	154
4.4.1	Distribution functions of random variables	154
4.4.2	Hydrodynamic models	160
4.5	Load model for the lakes	160
4.5.1	Distribution functions of random variables	160
4.5.2	Hydrodynamic models	162
4.6	Load model for the lake delta	162
4.6.1	Distribution functions of random variables	162
4.6.2	Hydrodynamic models	165
4.7	Load model for the coastal dikes	165
4.7.1	Distribution functions of random variables	165
4.7.2	Hydrodynamic models	172
4.8	Load model for the Oosterschelde	180
4.8.1	Distribution functions of random variables	180
4.8.2	Hydrodynamic models	181
4.9	Load model for the coastal dunes	183

4.9.1	Distribution functions of random variables	183
4.9.2	Hydrodynamic model	189
4.10	Spatial correlations between load models of different regions	192
<b>5</b>	<b>Failure mechanisms</b>	<b>198</b>
5.1	Introduction	198
5.2	Models of failure mechanisms in probabilistic failure computations	199
5.2.1	Introduction	199
5.3	Overtopping and Overflow	199
5.3.1	Limit State Function Overflow	201
5.3.2	Limit State Function Overtopping	201
5.3.3	Introduction to the overtopping discharge models	202
5.3.4	Overtopping: PC-overslag approach	203
5.3.5	Overtopping: PC-Ring approach	214
5.3.6	Critical overtopping discharge (CIRIA model)	216
5.3.7	Relevant variables for overtopping and overflow failure mechanisms	218
5.4	Failure mechanism 'Macrostability'	220
5.4.1	Introduction	220
5.4.2	Deelmechanisme Stab: Afschuiving MProStab	220
5.4.3	Deelmechanisme StPf: Kans op falen na afschuiving	223
5.4.4	Variabelen voor het mechanisme afschuiving	224
5.5	Pipsing and heave'	224
5.5.1	Mechanismebeschrijving	224
5.5.2	Deelmechanisme ZfOPBA: Opbarsten	225
5.5.3	deelmechanisme ZfPIPE: Piping	226
5.5.4	Variabelen voor het mechanisme opbarsten/piping	228
<b>6</b>	<b>Statistical distribution functions</b>	<b>230</b>
6.1	Introduction	230
6.2	General method to avoid influence of rounding errors	230
6.3	Uniform distribution	230
6.4	Normal distribution	232
6.5	(Shifted) Lognormal distribution	233
6.6	(Shifted) Exponential distribution	234
6.7	Gumbel distribution	235
6.8	Weibull distribution	237
6.9	Rayleigh distribution	237
6.10	Pareto distribution	237
6.11	Triangular distribution	237
6.12	Linear and log-linear interpolation	237
6.13	Conditional Weibull distribution	238
6.14	Modified Gumbel	239
6.15	Truncated Gumbel	241
<b>7</b>	<b>References</b>	<b>243</b>
<b>Appendices</b>		
<b>A</b>	<b>Dutch Hydraulic Boundary Conditions – version TMR2006</b>	<b>1</b>
A.1	Overview of stochastic input variables and regions	1

A.1.1 TMR2006 parameters

3

# 1 Introduction

## 1.1 Background

In compliance with the Dutch Water Act ('Waterwet, 2009'), the safety of the Dutch primary sea and flood defences must be assessed periodically for the required level of protection. In the Netherlands, the safety assessment of flood defences is carried out by evaluating dike sections, dune transects and structures individually, using representative hydraulic load conditions in combination with mathematical descriptions of failure mechanisms. From 2017 onwards, the evaluation will be done for entire 'dike rings', using the probability of flooding of the protected area as the main criterion. A dike ring area is an area that is protected from flooding by a dike ring, i.e. a closed system of flood defences (dikes, dunes, hydraulic structures, barriers). The differences with respect to the present assessment are twofold:

- Safety assessment per dike-ring in stead of per dike stretch;
- Quantification of the probability of failure instead of an assessment 'dike stretch is safe or unsafe'.

The advantage of the new approach is that it provides valuable insights in the weaker spots in the flood defence chain and, consequently, on the efficiency of investments of improvement works in terms of "risk reduction per invested euro".

The computation of failure probabilities for dike rings is a complex matter. Over the last decades, significant research was carried out to explore the possibility carrying out a safety assessment based on failure probabilities for all dike ring areas in the Netherlands. The research resulted, among others, in a probabilistic computation tool called PC-Ring. PC-Ring computes failure probabilities for flood defence systems composed of dikes, dunes and hydraulic structures. It is suitable for research purposes and is currently being used in several research projects, in The Netherlands and abroad. However, it is not suitable for use in an operational, legal and financial setting. Therefore, the Ministry of Public Works and the Environment has asked Deltares to develop a new operational tool for the safety assessment of entire systems of flood defences. The computational algorithms of this new tool, called Hydra-Ring, will be largely based on the algorithms of PC-Ring. The objective is to make Hydra-Ring more user-friendly, robust, efficient, accurate, better documented, better tested, more flexible and easier to maintain.

## 1.2 Purpose and scope of this document

The purpose of this document is to compile the scientific background of Hydra-Ring into one document. The information in this document is provided for future users of Hydra-Ring as well as developers. The scientific document is complementary to the Hydra-Ring design document, in which the technical design of Hydra-Ring is described.

## 1.3 Global setup of Hydra-Ring

This document provides the details of the computational methods that are implemented in Hydra-Ring to compute the probability of failure of a *system* of flood defences, which can be composed of a number of segments. Segments can be e.g. dike segments, dune segments or hydraulic structures. Failure of the flood defence system occurs if any segment within the system fails. Failure of a flood defence can occur due to a number of failure mechanisms. Failure at a segment occurs if the hydraulic load exceeds the resistance of the flood defence. The hydraulic load typically consists of the combination of water levels and waves (and in

some cases currents). In order to determine failure probabilities, the relevant statistical features of the water levels and waves are required. Unfortunately, for most dike sections there are no measurements available of these characteristics directly in front of the flood defence. Statistics of water levels and waves therefore need to be estimated from other sources. For this purpose, a hydraulic load model is required.

Hydra-Ring consists of three major building blocks:

- 1 Hydraulic load models, describing the probability of occurrence of the relevant range of hydraulic loads (water levels, waves and currents) at various locations in different water systems;
- 2 Failure mechanism models, describing under which load conditions a flood defence will fail;
- 3 Probabilistic computation techniques, to quantify the probability of failure of a system of flood defences, i.e. the probability that somewhere in the system the hydraulic load exceeds the resistance of the flood defence.

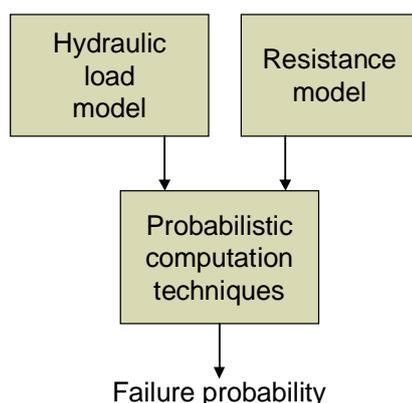


Figure 1.1 The three main building blocks of the computational core of Hydra-Ring

The load and resistance are both described by statistical models to capture the uncertainties that are involved. Uncertainties in the loads reflect the natural variability of the loads in combination with incomplete knowledge of the system characteristics. Uncertainties in the strength model mainly reflect incomplete knowledge of the system characteristics.

Figure 1.2 presents a schematic view of the stochastic nature of hydraulic loads and resistance of the flood defence. Both graphs in this figure represent probability density functions. It demonstrates that, in general, the resistance is higher than the load. However, there is also some overlap between the two graphs, indicating there is a probability that the load can exceed the resistance, in which case failure occurs. The probabilistic computation techniques in Hydra-Ring serve to quantify this probability. In order to do so, the (statistical) information of the load and resistance model is used as input. In practical terms this means that during a probabilistic computation the load model and resistance model are evaluated a (large) number of times, in order to quantify probabilities for the relevant range of load and resistance values.

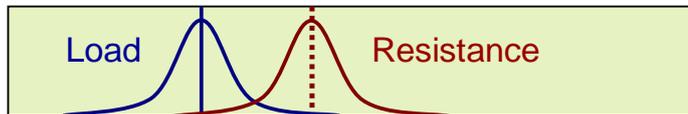


Figure 1.2 Schematic view of the stochastic nature of hydraulic loads and resistance of the flood defence. Both graphs represent probability density functions

The three major building blocks of Hydra-Ring are all designed in a modular way. For instance, a library of resistance models is implemented to consider all relevant failure mechanisms, such as:

- overtopping and overflow;
- piping;
- sliding of the inner slope;
- revetment erosion and subsequent internal erosion;
- dune erosion;
- closing failures of hydraulic structures;
- structural failure of hydraulic structures;
- etc.

Due to the modular setup, additional failure mechanisms can be easily included. Similarly, a library of load models is implemented to describe the hydraulic load of a variety of water systems, such as:

- tidal rivers;
- non-tidal rivers;
- lakes;
- estuaries;
- seas.

The probabilistic model is also designed in a modular way. A library is created for [a] probabilistic techniques to compute the probability of failure of one dike segment and [b] probabilistic techniques to compute the failure probability of an entire system of flood defences. The need for a set of probabilistic computation techniques stems from the fact that each technique has its advantages and disadvantages with respect to criteria like computation time, robustness and accuracy. The choice for the 'best' computation technique is therefore dependent on the characteristics of the load model and resistance model under consideration.

#### 1.4 The role of Hydra-Ring in the safety assessment procedure of flood defences in the Netherlands

Hydra-Ring will support the process of safety assessment of flood defence segments in the Netherlands. For this purpose, Hydra-Ring will be incorporated in the software package *Ringtoets*, the user-interface that is developed for the purpose of the safety assessment procedure. Figure 1.3 shows the various tasks to be carried out in the safety assessment procedure. Hydra-Ring is primarily designed to support the execution of tasks C1 en C2: the fully probabilistic safety assessment of flood defences. Furthermore, Hydra-Ring will be used to generate input of tasks B3 and B4: the semi-probabilistic safety assesment of flood defences. Semi-probabilistic in this case means that failure mechanisms are subsequently evaluated in a deterministic manner, using representative hydraulic boundary conditions that

have been derived with a probabilistic computation procedure. The semi-probabilistic approach is required for assessments for which a fully probabilistic procedure is [a] too time-consuming or [b] requires input data that is not available. The semi-probabilistic procedure is also useful as verification for the results of the fully probabilistic approach. Hydra-Ring will be used to provide the hydraulic boundary conditions for the semi-probabilistic approach.

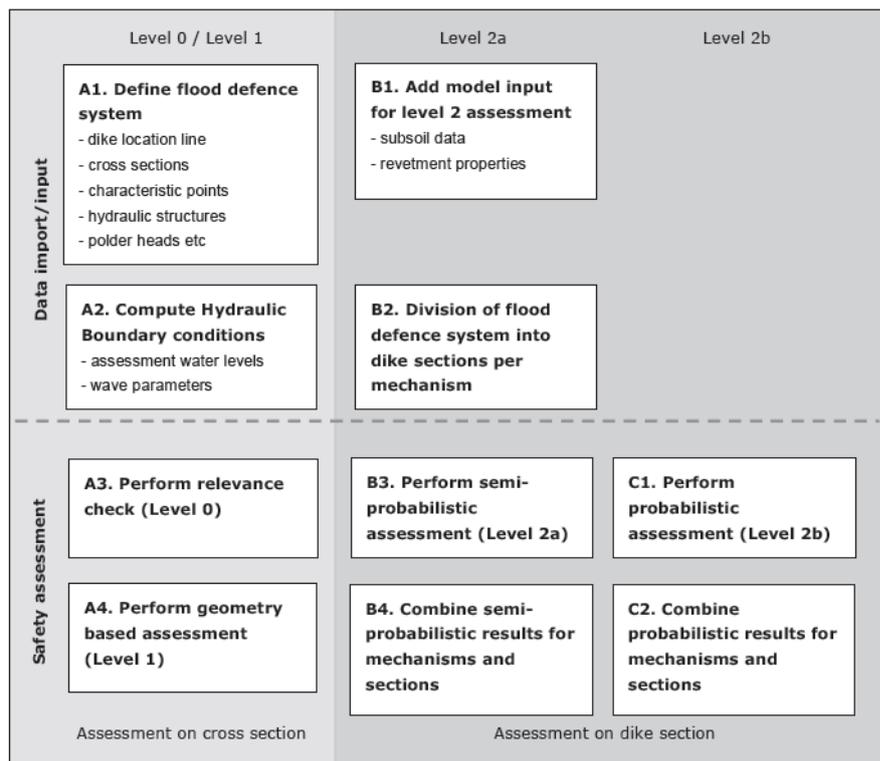


Figure 1.3 The tasks to be carried out in the safety assessment of primary flood defences in the Netherlands

## 1.5 Predecessors of Hydra-Ring

The design and implementation of Hydra-Ring is largely based on experiences with a number of related probabilistic models for flood risk assessment, the most relevant of which are:

- 1 PC-Ring
- 2 The 'Hydra-models'

PC-Ring can be considered to be the most direct predecessor of Hydra-Ring. As mentioned in section 1.1, PC-Ring is the probabilistic computation tool that is one of the main results of the research on the safety assessment of dike ring areas, that was carried out in the Netherlands in the last decades. PC-Ring computes failure probabilities for flood defence systems composed of dikes, dunes and hydraulic structures. Hydra-Ring can be considered an improved version of PC-Ring, i.e. Hydra-Ring is more user-friendly, robust, efficient, accurate, better documented, better tested, more flexible and easier to maintain.

The 'Hydra-models' have been developed within the framework of the six-yearly safety assessment of primary flood defences in the Netherlands. The Hydra-models serve to derive representative hydraulic boundary conditions (HBC) for all primary flood defences in the

Netherlands. Representative means that a flood defence is approved if it is considered to be able to withstand the HBC. Different Hydra-models were implemented for different water systems, to take into account the differences in characteristics of the water systems: Hydra-R for non-tidal rivers, Hydra-B for the tidal Rhine-Meuse delta, Hydra-M for the lakes area, Hydra-VIJ, for the IJssel-Vecht delta and Hydra-K for the coastal systems. Recently, the Hydra-models for the inland water systems (Hydra-R, Hydra-B, Hydra-M and Hydra-VIJ) have been merged into one model: Hydra-Zoet.



Figure 1.4 Different areas in the Netherlands for which hydra-models have been developed

Like Hydra-Ring and PC-Ring, the Hydra-models are probabilistic computation models in which the loads are compared with the resistance of the flood defence. An important difference is that the resistance in the 'Hydra-models' is modeled in a deterministic manner, whereas in Hydra-Ring and PC-Ring the uncertainties in the resistance can be taken into account. Furthermore, the Hydra-models only consider a small number of failure mechanisms compared to PC-Ring and Hydra-Ring. Another difference is that the Hydra-models only consider the probability of failure per cross section, whereas PC-Ring and Hydra-Ring consider entire systems of flood defences.

Despite the above mentioned differences between Hydra-Ring and the other Hydra-models, there are also a lot of similarities. For instance the load models of the Hydra-models served as the blueprint for the load models in Hydra-Ring (and also PC-Ring). Hydraulic loads as computed with Hydra-Ring should therefore be (almost) the same as loads computed with the Hydra-models. It is for this reason that the Hydra-models will be used as a vital reference in this document, especially in describing the load models.

## 1.6 Outline

As described in section 1.3, the three major building block of Hydra-Ring are (libraries of) probabilistic computation techniques, load models and resistance models. Chapter 2 describes the probabilistic computation techniques. Chapters 3 and 4 describe the load models. Chapter 3 describes the generic setup of a load model in Hydra-Ring whereas chapter 4 describes the load models of the primary water systems in The Netherlands, as implemented in Hydra-Ring. Chapter 5 describes the resistance models and chapter 6 describes the library of statistical distribution functions that can be used to describe statistical characteristics of random variables in both the load and the resistance models.

## 2 Probabilistic computation techniques for system reliability

### 2.1 Introduction

This chapter describes the probabilistic computation techniques that are used in Hydra-Ring for the computation of failure probabilities. Hydra-Ring computes the failure probability of a flood defence *system*, which can be composed of a number of segments. Segments can be e.g. dike segments, dune segments or hydraulic structures. Failure of the flood defence system occurs if any segment within the system fails, and failure can occur due to a number of failure mechanisms (see Chapter 5). Mechanisms and segments will be referred to as components of the system. Hydra-Ring thus computes the failure of a system that consists of a (large) number of components. The procedure is designed as follows:

- 1 determine the probability of all the individual components of the system, and
- 2 combine the failure probabilities of the components to derive the failure probability of the entire system.

Section 2.3 describes probabilistic computation methods for the failure probability of the individual components. Each of these methods has been implemented in Hydra-Ring in order to increase the flexibility of the program. To support the description of the probabilistic computation techniques in section 2.3, section 2.2 first describes some relevant terminology in combination with a number of “basic” statistical computation techniques. Section 2.4 describes generic probabilistic techniques for deriving the failure probability of a system. Section 2.5 describes the specific choices and methods for deriving and combining failure probabilities as implemented in Hydra-Ring. Section 2.6 presents an overview of the whole procedure that is applied in Hydra-Ring to derive the failure probability of a flood defence system.

Note that the material in this chapter requires a basic knowledge of statistics and probability.

### 2.2 Basic techniques and terminology

#### 2.2.1 Systems and components

As described in the previous section, Hydra-Ring computes the failure probability of a flood defence system, composed of a number of dike segments, dune segments, and hydraulic structures. In order to derive the failure probability of the system, the failure probabilities of the individual components need to be combined. As will be described in sections 2.4 and 2.5, this is a process that is far more complex than simple additions and/or multiplications of the probabilities of individual segments. Furthermore, combining failure probabilities is not just required for the individual segments of the defence system, but also for:

- *Failure mechanisms*: A flood defence system can fail due to different failure mechanisms.
- *Time periods*: failure probabilities are first computed for relatively small time scales (<1 day) in which the temporal variation of the relevant hydraulic variables (water level, wind etc) is small enough to be assumed constant. However, the required output of Hydra-Ring is an annual failure probability, which means failure probabilities of the smaller time scales need to be combined (scaled-up) into an annual failure probability.
- *Cross sections*: The probability of failure is first derived for a single location (cross-section) of the flood defence segment. This probability is subsequently scaled up to a failure probability of the entire flood defence segment.

- *Closure scenarios of tidal barriers and wind directions:* These are random variables that cannot be included in the probabilistic computation of a single component, for reasons to be explained in section 2.2.3. Therefore, the computed failure probabilities per wind direction and per closure scenario need to be combined as well.

A single component in Hydra-Ring therefore refers to a selected combination of one cross section, one failure mechanism, one wind direction, one closure scenario and one relatively small (<1 day) time interval.

The probabilities of failure of the components are combined to derive the probability of failure of the whole system. This is being referred to as *system analysis*. System analysis generally deals with parallel systems, series systems or combinations of both. A parallel system refers to a system in which failure only occurs if all components fail. A series system refers to a system where failure occurs if at least one of the components fails. This concept is schematically depicted in Figure 2.1.

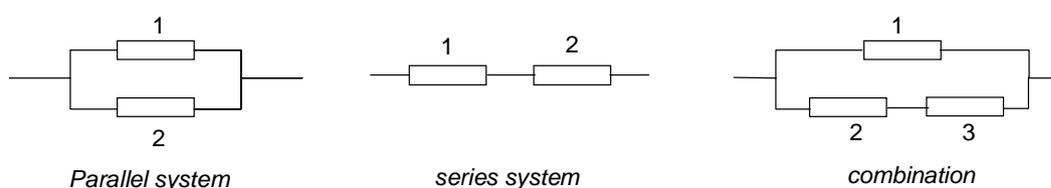


Figure 2.1 Schematic view of a parallel system, series system and a combination of both.

In mathematical descriptions of system analysis, the symbols for ‘AND’,  $\cap$ , and ‘OR’,  $\cup$  are used as follows:

Parallel system:  $P[\text{failure}] = P[\text{failure component 1} \cap \text{failure component 2}]$

Series system:  $P[\text{failure}] = P[\text{failure component 1} \cup \text{failure component 2}]$

Where P stands for probability. Flood risk analysis generally deals with series systems. A system of flood defences protects an area and failure (flooding) occurs if one or more components (flood defences) fails. Nevertheless, parallel (sub)systems can occur as well. For instance the failure mechanism ‘piping and heave’ only occurs if the two submechanisms ‘piping’ and ‘heave’ both occur.

## 2.2.2 Formulation of the probability of failure of a single component

The probability of failure for a single component can be written formally as follows:

$$P_f = P(Z(\mathbf{X}) < 0) \tag{2.1}$$

where:

$P_f$  = failure probability

$Z$  = limit state function

$X$  = vector of random variables

The limit state function,  $Z$ , defines failure in terms of load and strength variables such that  $Z < 0$  represents failure. Function  $Z$  is often denoted:

$$Z = R - S, \quad (2.2)$$

where:

R = strength, or resistance

S = Load

In case of a flood defence, the load typically consists of the combination of water levels and waves and in some cases currents. The strength is a combination of dike characteristics that reflect the ability of the dike to resist high loads. A simple example is overflow of a river levee in a river (see Figure 2.2). The load in this simple example is the water level in the river,  $h_{river}$ , and the strength of the levee is captured by its height,  $h_{levee}$ . The limit state function for this example is simply:

$$Z = R - S = h_{levee} - h_{river} \quad (2.3)$$



Figure 2.2 Illustration of a river levee in danger of overflowing

The limit state function,  $Z$ , is a function of a number of random variables representing both load and strength variables. The probability of failure can be written as follows:

$$P_f = \int \int_{Z < 0} \dots \int f_{X_1, \dots, X_n}(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n = \int_{Z < 0} f_X(\mathbf{x}) d\mathbf{x}, \quad (2.4)$$

where  $f$  is the joint probability distribution of the random variables and  $X$  is the vector of  $X$ -variables:  $X = (X_1, \dots, X_n)$ . Note that random variables are typically denoted with a capital letter ( $X$ ), while potential realizations of the random variables are denoted with lower case letters ( $x$ ).

While an analytical solution to equation (2.4) would be ideal, it is typically not possible because the  $Z$ -function is too complex. Therefore, the probability of failure needs to be estimated with probabilistic computation techniques. Different techniques are available for this purpose. These type of techniques will be described in detail in the section 2.3.

2.2.3 Conversion between standard normal variables (U-domain) and real variables (X-domain)  
 In practice, it is often advantageous to carry out probabilistic analyses in a standardized space, in which each of the variables are independent. This independence can help simplify the probabilistic techniques. The dependence between variables is reintroduced when the standardized variables are transformed back to the real variables. In Hydra-Ring the standardized space that is used is the standard normal space. This means each of the transformed random variables are normally distributed with mean 0, and standard deviation 1. Each of the random variables in X can be transformed to independent standard normal variables  $U = (U_1, \dots, U_n)$ . The transformation is done such that the non-exceedance probabilities are set equal. If all the X-variables are mutually independent, this transformation can be described as follows:

$$\Phi(u_i) = F_i(x_i) \Rightarrow u_i = \Phi^{-1}[F_i(x_i)] \Rightarrow x_i = F_i^{-1}(\Phi(u_i)), \quad (2.5)$$

where:

- $\Phi$  = **Standard normal distribution function**
- $F_i$  = Distribution function (cdf) of  $X_i$ , the  $i^{\text{th}}$  variable in X
- $\Phi^{-1}$  = Inverse standard normal distribution function
- $F_i^{-1}$  = Inverse distribution function of  $X_i$ , the  $i^{\text{th}}$  variable in X
- $x_i$  = Realization of the  $i^{\text{th}}$  variable in X
- $u_i$  = Realisation of the  $i^{\text{th}}$  variable in U corresponding to the  $i^{\text{th}}$  variable in X

The concept of this transformation is schematically depicted in Figure 2.3.

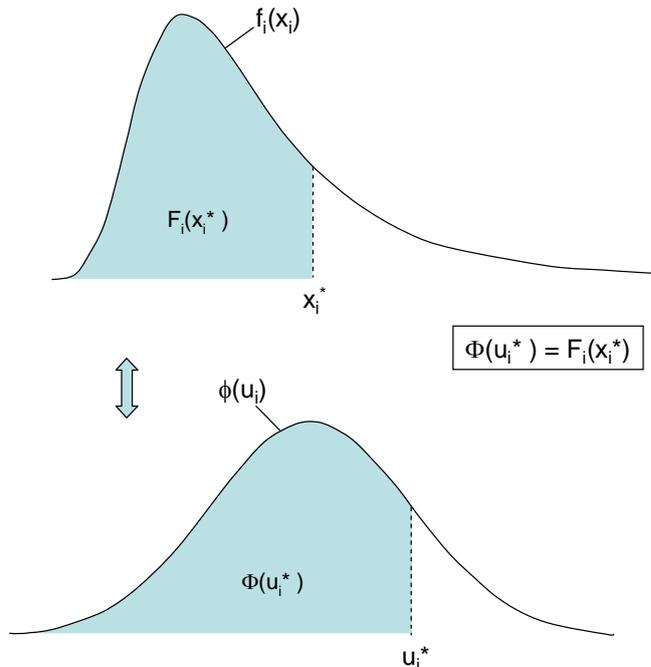


Figure 2.3 Schematic view of the transformation between standard normal variable  $u_i$  and “real world” variable X by means of equal probability of (non-) exceedance.

If the X-variables are not mutually independent, i.e. they are correlated, the transformation becomes more complex. In that case the transformation is done with the following conditional probability functions:

$$\begin{aligned}
 \Phi(u_1) &= F_1(x_1) \\
 \Phi(u_2) &= F_2(x_2 | x_1) \\
 &\vdots \\
 &\vdots \\
 \Phi(u_n) &= F_n(x_n | x_1, \dots, x_{n-1})
 \end{aligned}
 \tag{2.6}$$

This transformation is referred to as the Rosenblatt transformation (Rosenblatt, 1952). The conditional probabilities need to be derived from:

$$f_i(x_i | x_1, \dots, x_{i-1}) = \frac{f_{x_1, \dots, x_i}(x_1, \dots, x_i)}{f_{x_1, \dots, x_{i-1}}(x_1, \dots, x_{i-1})}
 \tag{2.7}$$

In which  $f_{x_i}(x_1, \dots, x_i)$  is a probability density function that is obtained from:

$$f_{x_i}(x_1, \dots, x_i) = \int \dots \int f_{\mathbf{X}}(x_1, \dots, x_n) dx_{i+1} \dots dx_n
 \tag{2.8}$$

The combination of equations (2.7) and (2.8) potentially requires a cumbersome (numerical) integration procedure. In Hydra-Ring, this process is far less complex as correlation is only limited to pairs of variables. Generally, the pairs are split into a dependent variable and an independent variable, and the dependent variable is written as a function of the independent variable. The function consists of a fully deterministic dependent part and a probabilistic independent part:

$$F_{x_2}(x_2 | x_1) = G(x_1) + F_2(x_2^*)
 \tag{2.9}$$

Where G is a deterministic function,  $F_2$  a probability distribution function and  $x_2^*$  a newly introduced random variable that represents the part of variable  $x_2$  that is independent of  $x_1$ . This type of correlation models is schematically depicted in Figure 2.4. More details on this type of correlation models are described in section 0. For now, it is important to note that with a correlation model as described by the general form of equation (2.9), the transformation from and to the u-space for variable  $X_2$  is done for the independent part:

$$\Phi(u_2) = F_2(x_2^*),
 \tag{2.10}$$

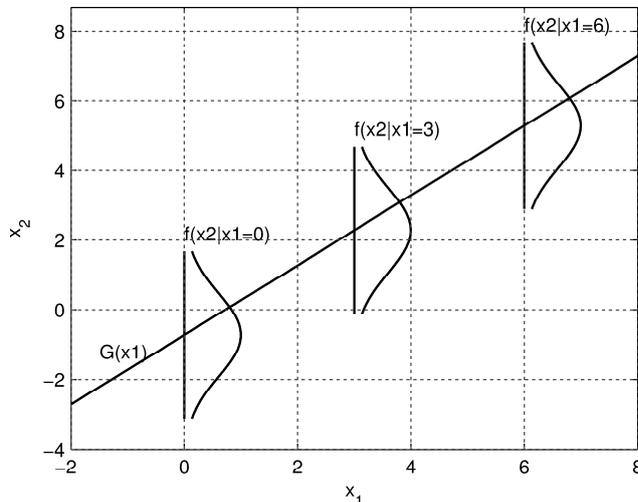


Figure 2.4 Schematic view of correlated variables  $x_1$  and  $x_2$

Not all random variables can be represented in a meaningful way by u-variables. This is for instance the case for “cyclic variables” like wind direction. The reason is that transformation (2.5) only makes sense if large (extreme) u-values are associated with large (extreme) x-values. Potential outcomes of cyclic variables are not “ordered” from small to large. The domain for these variables is 1-360 degrees, but 360 degrees is not “larger” than 1 degree. Actually, 1 degree is almost the same as 360 degrees. As will be discussed later on (section 2.5), cyclic variables in Hydra-Ring will be treated different from variables like wind speed, river discharge and sea water level.

## 2.2.4 Reliability index $\beta$

The reliability index,  $\beta$ , is a measure for the reliability of a system, i.e. a measure for the probability of failure of the system. Similar to the probability of failure,  $\beta$  is often defined for a given period of time, e.g. a year. The reliability index is best explained by an example in which the resistance,  $R$ , and load,  $S$ , of the system are both described as the sum of independent normally distributed random variables. The sum of a set of independent normally distributed random variables is also a random variable (see, e.g. Grimmitt and Sirzaker, 1982). This means in this case,  $R$  and  $S$  are also normally distributed variables and the same can be stated about the Z-function ( $Z=R-S$ ). Define  $\mu_R$ ,  $\mu_S$  and  $\mu_Z$  as the respective mean values of  $R$ ,  $S$  and  $Z$  and  $\sigma_R$ ,  $\sigma_S$  and  $\sigma_Z$  as the respective standard deviations of  $R$ ,  $S$  and  $Z$ . The following relations hold:

$$\begin{aligned} m_Z &= m_R - m_S \\ s_Z &= \sqrt{s_R^2 + s_S^2} \end{aligned} \quad (2.11)$$

Figure 2.5 shows an example with  $\mu_R=6$ ,  $\mu_S=2$ ,  $\sigma_R = \sigma_S=1$  and consequently  $\mu_Z = 4$  and  $\sigma_Z=\sqrt{2}$ . In reliability analysis,  $\mu_Z$  is generally a positive value because otherwise, failure ( $Z<0$ ) would occur even during “average” conditions. Since  $Z$  is normally distributed,  $P[Z<0]$  is equal to:

$$P(Z < 0) = F\left(\frac{m_Z - 0}{s_Z}\right) = 1 - F\left(\frac{0 - m_Z}{s_Z}\right) \quad (2.12)$$

Where  $\Phi$  is the standard normal distribution function. Based on this equation, the reliability index,  $\beta$ , is defined as:

$$b = \frac{m_Z}{S_Z} \quad (2.13)$$

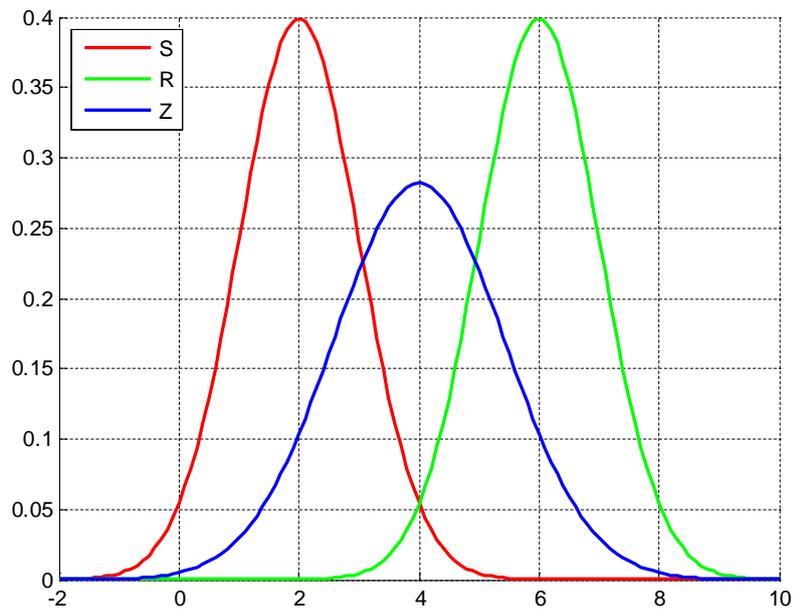


Figure 2.5 Example of normally distributed S, R and Z functions with  $\mu_R=2$ ,  $\mu_S=6$ ,  $\sigma_R=1$ ,  $\sigma_S=1$ ,  $\mu_Z=4$  and  $\sigma_Z=\sqrt{2}$ .

For the specific case where the Z-function is normally distributed, the relation between  $\beta$  and the probability of failure is:

$$P(Z < 0) = F(-b) = 1 - F(b) \quad (2.14)$$

Or inversely:

$$b = F^{-1}(1 - P(Z < 0)) \quad (2.15)$$

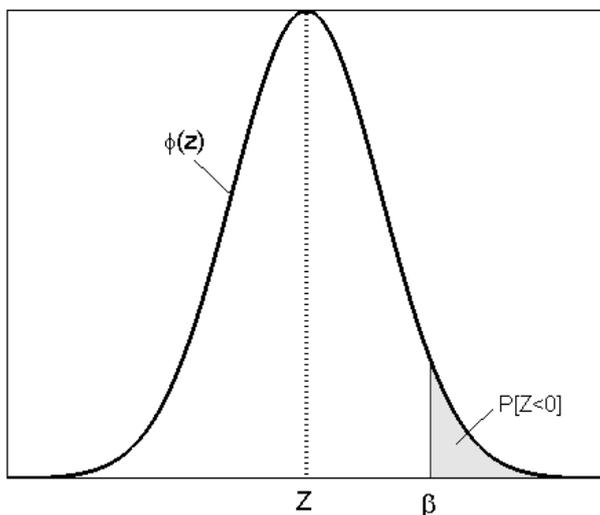


Figure 2.6 Schematic view of the relation between the reliability index  $\beta$  and the probability of failure,  $P[Z < 0]$ .  
Function  $\phi$  is the standard normal density function

This shows why the reliability index  $\beta$  is a measure for reliability. Note that if the Z-function is not normally distributed, equation (2.14) does not necessarily hold. This is why  $\beta$  is a *measure* for reliability, not an exact representative of the probability of failure. Table 2.1 shows a range of  $\beta$ -values and associated probabilities of exceedance.

Table 2.1 Values of reliability index  $\beta$  and associated probability of failure.

$\beta$	p
1.0	1.59E-01
1.5	6.68E-02
2.0	2.28E-02
2.5	6.21E-03
3.0	1.35E-03
3.5	2.33E-04
4.0	3.17E-05
4.5	3.40E-06
5.0	2.87E-07

Small values of  $\beta$  indicate large probabilities of failure, large values of  $\beta$  indicate small probabilities of failure. This can be easily explained with some examples in which, for the sake of simplicity, Z is assumed to be normally distributed. If  $\beta=1$ , failure occurs if a random sample of Z is more than  $1 \cdot \sigma_Z$  lower than the mean. The probability for this to happen is equal to  $\Phi(-1) \approx 0.16$ . On the other hand, if  $\beta=4$ , failure occurs if a random sample of Z is more than  $4 \cdot \sigma_Z$  lower than the mean. The probability for this to happen is equal to  $\Phi(-4) \approx 3.2 \cdot 10^{-5}$ . In other words: larger values of  $\beta$  indicate that more extreme events are required for failure to occur, hence a lower probability of failure, hence a larger reliability of the system.

### 2.2.5 Linearisation of Z-functions

Z-functions in flood risk analysis generally describe a combination of hydrodynamics and geotechnical processes. Due to the complexity of the Z-function it is sometimes practical to use linear approximations. The linearization can result in significant reductions of the computation time. This is for instance the case with the probabilistic computation method FORM, (see section 2.3.6) which is generally much faster than other probabilistic computation techniques such as Monte Carlo (see section 2.3.3).

Another advantage of the linearization is that it enables (semi-)analytical approaches to complex system analysis, that otherwise would not have been possible. Such an approach is for instance used in the 'Hohenbichler method' (see section 2.4.2) which is applied to compute the total probability of failure of a system of components. In case of Hydra-Ring the Hohenbichler method is applied to combine failure probabilities of dike sections in a dike ring and to combine failure probabilities of different failure mechanisms.

The disadvantage of the linearization is of course the fact that it is an approximation of the Z-function, which means an error is likely to be introduced in the estimate of the failure probability. As long as this error is small compared to other modeling errors and uncertainties this poses no real problem, but this needs to be verified as much as possible. First, the linearization process is described.

The linearization of the Z-function is generally applied in the U-space, in which the U-variables are independent standard normally distributed random variables. In other words: the function  $Z(u)$  is linearized, where  $u$  is the vector of realisations,  $u_1, \dots, u_n$ , of the standard normally distributed variables. The linear approximation of the Z-function has the following form:

$$Z_L = B + A_1 u_1 + \dots + A_n u_n \quad (2.16)$$

The linearization is done by taking the tangent of the Z-function in a selected location  $u_d$ . This means the A-values are chosen as follows:

$$A_i = \frac{\partial Z}{\partial u_i}(u_d) \quad ; i = 1..n \quad (2.17)$$

This linearization process is depicted in Figure 2.7 and Figure 2.8.

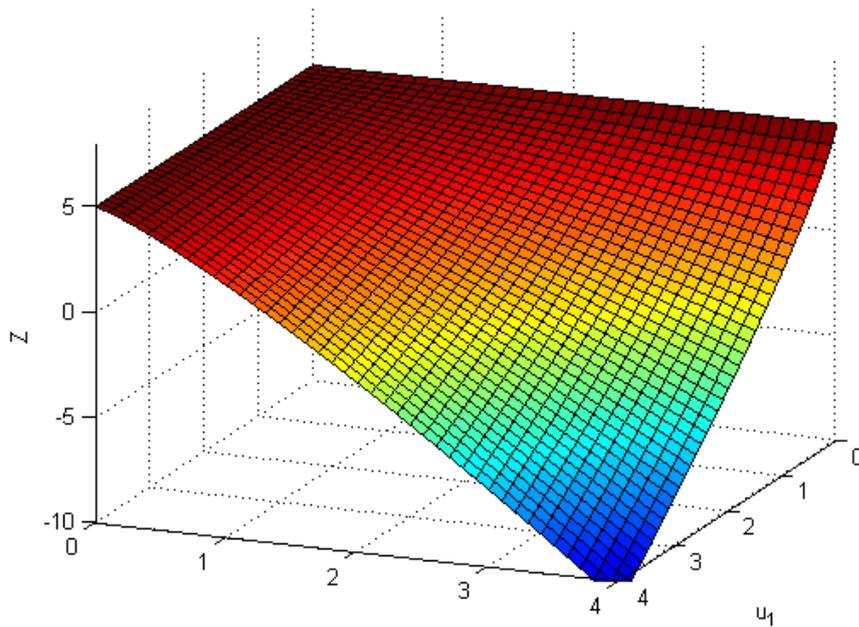


Figure 2.7 Example of function  $Z(u_1, u_2)$

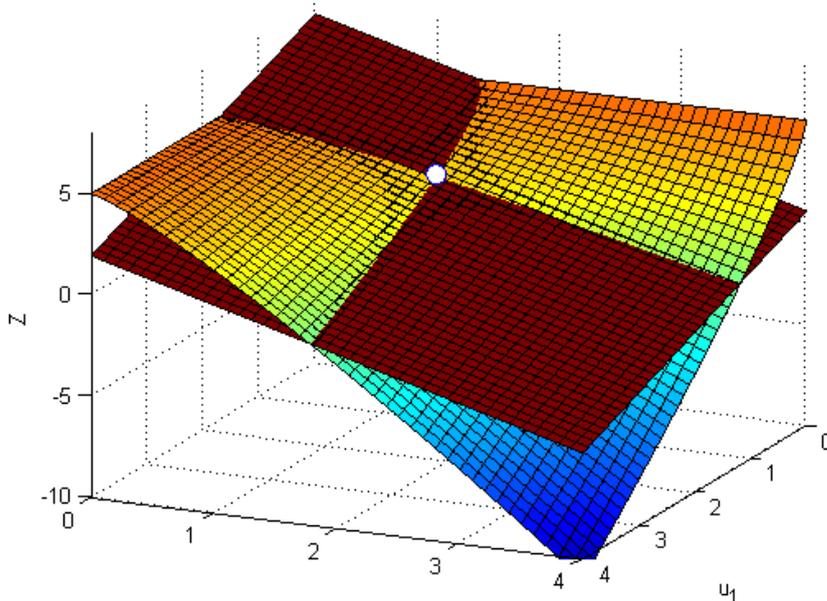


Figure 2.8 Linearisation of the  $Z$ -function of Figure 2.7 in a selected location (white dot).

Clearly, the linearised  $Z$ -function is different from the actual  $Z$ -function. This means an error will be introduced in the estimation of the probability of failure,  $P(Z < 0)$ . To reduce this error as much as possible, the linearization is generally done in the design point,  $u_d$ . This is the

location on the hyperplane  $Z=0$  with the highest probability density. The method FORM, as described in section 2.3.6 is based on this principle.

Generally, the main objective is to compute the probability of failure, i.e.  $P(Z<0)$ . In that case, the right hand side of equation (2.16) can be multiplied or divided by a constant. If this constant is taken to be the norm of vector  $A = (A_1, \dots, A_n)$  the linear Z-function has the following form:

$$Z_L = b + a_1 u_1 + \dots + a_n u_n = b + \sum_{i=1}^n a_i u_i \quad (2.18)$$

In which:

$$b = \frac{B}{\|A\|}; \quad a_i = \frac{A_i}{\|A\|}, i = 1..n; \quad \|A\| = \sqrt{\sum_{i=1}^n A_i^2} \quad (2.19)$$

The norm of vector  $\alpha=(\alpha_1, \dots, \alpha_n)$  is then equal to 1:

$$\sqrt{\sum_{i=1}^n \alpha_i^2} = 1 \quad (2.20)$$

This means the linearised Z-function has been “normalised”. Since the u-values are independent standard normally distributed values, this means:

$$\sum_{i=1}^n \alpha_i u_i \sim N(0,1) \quad (2.21)$$

In other words: the sum of the product of  $\alpha$ -values and u-values,  $\sum \alpha_i u_i$ , is standard normally distributed. This means that in order to compute  $P(Z_L < 0)$ ,  $\sum \alpha_i u_i$  can be replaced by a single standard normally distributed value  $u^*$ :

$$Z_L = b + u^* \quad (2.22)$$

Note that, since the density function of  $U^*$  is symmetric around  $u^*=0$ ,  $Z_L$  can also be described as:

$$Z_L = b - u^* \quad (2.23)$$

In equation (2.22), failure occurs if  $Z_L < 0$ , i.e. if  $u^* < -\beta$ . The probability that this occurs is equal to  $\Phi(-\beta)$ , where  $\Phi$  is the standard normal distribution function and  $\beta$  is the reliability index which was introduced in section 2.2.4. While  $\beta$  is an indicator for the probability of failure, the  $\alpha$ -values are indicators for the relative importance of the associated random variables, as will be shown below. From equation (2.18) it can be seen that:

$$P(Z_L < 0) = P\left\{\beta + \sum_{i=1}^n a_i u_i < 0\right\} \quad (2.24)$$

If we increase the mean of variable  $u_i$  ( $\bar{u}_i$ ) with a small value  $\varepsilon_i$  this will have an effect on the probability of failure. The magnitude of this effect is an indicator of the relative importance of variable  $u_i$ . For this purpose, define the random variable  $u_i'$  as follows:

$$u_i' = u_i + \varepsilon_i \quad (2.25)$$

Since  $u_i$  is standard normally distributed,  $u_i'$  is normally distributed with mean  $\varepsilon_i$  and standard deviation 1. Now we replace  $u_i$  in equation (2.18) by  $u_i'$  and consequently obtain a new Z-function  $Z_L'$ :

$$\begin{aligned} Z_L' &= \beta + \alpha_1 u_1 + \dots + \alpha_i u_i' + \dots + \alpha_n u_n \\ &= \beta + \alpha_1 u_1 + \dots + \alpha_i (u_i + \varepsilon_i) + \dots + \alpha_n u_n \\ &= (\beta + \alpha_i \varepsilon_i) + \alpha_1 u_1 + \dots + \alpha_i u_i + \dots + \alpha_n u_n \end{aligned} \quad (2.26)$$

So the perturbation of the mean of variable  $u_i$  results in a new Z-function with reliability index  $\beta'$  instead of  $\beta$ , with:

$$\beta' = \beta + \alpha_i \varepsilon_i \quad (2.27)$$

This means:

$$\frac{\partial \beta}{\partial \bar{u}_i} = \frac{\partial \beta}{\partial \varepsilon_i} = \frac{\beta' - \beta}{\varepsilon_i} = a_i \quad (2.28)$$

In other words:  $\alpha_i$  is a measure of the sensitivity of reliability index  $\beta$  to changes in the mean value of variable  $u_i$ . This also means  $\alpha_i$  is a measure of the sensitivity of the probability of failure to changes in the mean value of variable  $u_i$ . This characteristic is used in the Hohenbichler method for combining probabilities of components in a system (see section 2.4.2).

As stated before, linearized Z-functions are the basis for various computation techniques that are explained in the following sections. A full understanding of this linearization process and the meaning of  $\alpha$ -values and  $\beta$  is therefore essential for further reading of the report.

## 2.3 Failure probability for a single component

### 2.3.1 Introduction

Equation (2.4) describes the general formulation of the probability of failure of a single component. While an analytical solution to equation (2.4) would be ideal, it is typically not possible because the Z-function is too complex. Therefore, the probability of failure needs to be estimated with probabilistic computation techniques. The computation techniques available within Hydra Ring are summarized in Table 2.2. Each of these techniques will be described in detail in the current section. The reason to implement a set of probabilistic

computation techniques in Hydra-Ring is that each technique has its (dis)advantages with respect to criteria such as robustness, accuracy and required computation time. The “best” technique to be applied therefore depends on the problem under consideration. Section 2.3.8 discusses the (dis)advantages of each of these techniques. First, the techniques are explained individually.

Table 2.2 Computation techniques available in Hydra-Ring for the computation of failure probability of a cross section of a longitudinal segment.

Method	Variant
Numerical integration	--
Monte Carlo	Crude Importance sampling Directional sampling
FORM (First order reliability method)	--

### 2.3.2 Numerical Integration

Numerical integration solves equation (2.4) by discretizing the random variables  $X_1 \dots X_n$ . Each variable is discretized over a range that is relevant for failure, and subsequently each combination of discretized values of the X-variables is used to compute the limit state function. The probabilities of all the combinations that lead to  $Z < 0$  are summed, which provides the estimate of the overall probability of failure. This summation can be written as follows:

$$\hat{P}_f = \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \dots \sum_{i_n=1}^{m_n} 1_{[Z < 0]} f_X(x_{0,1} + (i_1 - 0.5)\Delta x_1, x_{0,2} + (i_2 - 0.5)\Delta x_2, \dots, x_{0,n} + (i_n - 0.5)\Delta x_n) \quad \text{Displa}$$

,

where

$\hat{P}_f$  = Estimated probability of failure

$1_{[Z < 0]}$  = Indicator function, equal to 1 for  $Z < 0$ , equal to 0 for  $Z \geq 0$

$x_{0,k}$  = Lower range limit for the  $k^{\text{th}}$  variable

$\Delta x_k$  = Interval width of the  $k^{\text{th}}$  variable

$m_k$  = Upper bound of  $k$  such that  $x_{0,k} + m_k \Delta x_k$  is the upper bound of the  $k^{\text{th}}$  variable

In equation (2.29), for each variable  $X_k$  an equidistant grid with step size  $\Delta x_k$  is used, but non-equidistant grids can also be used in numerical integration. Figure 2.9 presents a schematic view of the method, for an example of two random variables  $X_1$  and  $X_2$ . A 2-dimensional grid is defined and the Z-function is evaluated at the centre of the grid cells. Red grid points indicate failure ( $Z < 0$ ), green grid points indicate no failure ( $Z \geq 0$ ). The total probability of failure (see equation (2.29)) is estimated as follows: multiply the probability density of the grid cells in the failure domain (red dots) with the size of the grid cells ( $\Delta x_1 \times \Delta x_2$ ) and take the sum of these probabilities.

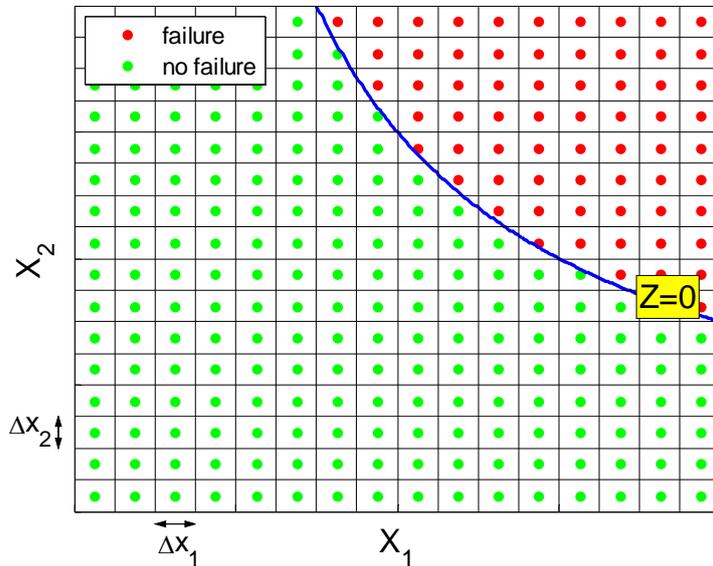


Figure 2.9 Schematic view of the method of numerical integration for an example of two random variables. A 2-dimensional grid is defined and the Z-function is evaluated at the centre of the grid cells. Red grid points indicate failure ( $Z < 0$ ), green grid point indicate no failure ( $Z \geq 0$ ).

Like every probabilistic estimation technique, the result of the numerical integration procedure will be an approximation of the actual probability of failure. The errors that are introduced in this method are caused by the following assumptions and approximations:

- 1 Each grid cell is assumed to be entirely situated in the failure domain or entirely situated outside the domain of failure domain. In reality, grid cells can be partly in the failure domain as can be seen in Figure 2.9.
- 2 The probability density is assumed to be constant over the entire grid cell.
- 4 The domain of potential outcomes of the random variables may not be entirely covered.

In the implementation of the procedure, it may be beneficial to transform the X-variables to standard normally distributed U-variables (see section 2.2.3). One of the benefits of working in the U-space is that the U-variables are independent, which simplifies equation (2.29) as follows:

$$\hat{P}_f = \sum_{i_1=1}^{m_1} \dots \sum_{i_n=1}^{m_n} 1_{[Z < 0]} \phi(u_{0,1} + (i_1 - 0.5)\Delta u_1) \cdot \dots \cdot \phi(u_{0,n} + (i_n - 0.5)\Delta u_n) \Delta u_1 \dots \Delta u_n \quad (2.30)$$

where

- $\hat{P}_f$  = Estimated probability of failure
- $\phi$  = standard normal density function
- $u_{0,k}$  = Lower range limit for the  $k^{\text{th}}$  variable, in the  $u$ -space
- $\Delta u_k$  = Interval width of the  $k^{\text{th}}$  variable, in the  $u$ -space

$m_k$  = Upper bound of  $k$  such that  $u_{0,k} + m_k \cdot \Delta u_k$  is the upper bound of the  $k^{\text{th}}$  variable

### 2.3.3 Crude Monte Carlo

Crude Monte Carlo sampling refers to the repeated sampling the variables from the multivariate probability distribution function  $f_{\mathbf{x}}(\mathbf{x})$  (or, if the variables are mutually independent, sampling from the respective distribution functions  $f_{x_1}(x_1), \dots, f_{x_n}(x_n)$ ). A single sample  $\mathbf{x}^i$  refers to a vector of length  $n$ , where  $n$  is the number of random variables. For each sample  $\mathbf{x}^i$ , the resulting value of limit state function  $Z(\mathbf{x}^i)$  is computed. The probability of failure is estimated as the ratio of samples for which  $Z(\mathbf{x}^i) < 0$ ,  $N_f$ , to the total number of samples,  $N$ :

$$\hat{P}_f = \frac{N_f}{N} = \frac{\sum_{i=1}^N I(Z(\mathbf{x}^i))}{N} \quad (2.31)$$

Where  $I$  is the indicator function, which is equal to unity when  $Z < 0$ , equal to zero when  $Z \geq 0$ . Figure 2.10 shows a schematic view of the procedure for an example with two random variables  $X_1$  and  $X_2$ . Each dot represents a sampled pair  $(x_1, x_2)$ . Red grid points indicate failure ( $Z < 0$ ), green grid point indicate no failure ( $Z \geq 0$ ). The estimated probability of failure is equal to the number of the red dots divided by the total number of dots.

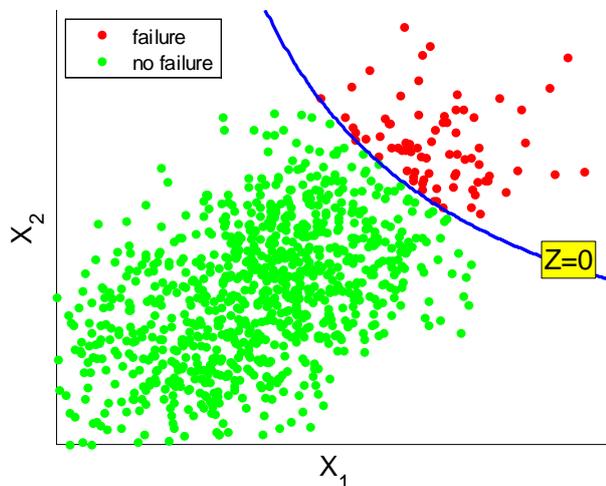


Figure 2.10 Schematic view of Monte Carlo sampling for an example with two random variables. Each dot represents a sampled pair  $(x_1, x_2)$ . Red dots indicate failure ( $Z < 0$ ), green dots indicate no failure ( $Z \geq 0$ ).

The required number of samples,  $N$ , to provide a reliable estimate of the probability of failure depends on the actual failure probability  $P_f$  and on the acceptable error in the estimate of  $P_f$ . Additionally, it depends on the acceptable probability that the real error is within the accepted range. This is because even though taking a large number of samples will most likely result in small errors (law of large numbers), it can never be fully guaranteed due to the random character of the Monte Carlo sampling. However, it is possible to take  $N$  large enough to guarantee with e.g. 95% or 99% certainty that the error in the estimate is within the acceptable range. This probability,  $p_k$ , can be expressed as:

$$p_k = \Phi(k) - \Phi(-k) \quad ; k > 0. \quad (2.32)$$

where  $\Phi$  represents the standard normal distribution function, and  $k$  represents a sort of reliability index that the error is within the accepted range. The relation between  $k$  and  $p_k$  is schematically depicted in Figure 2.11.

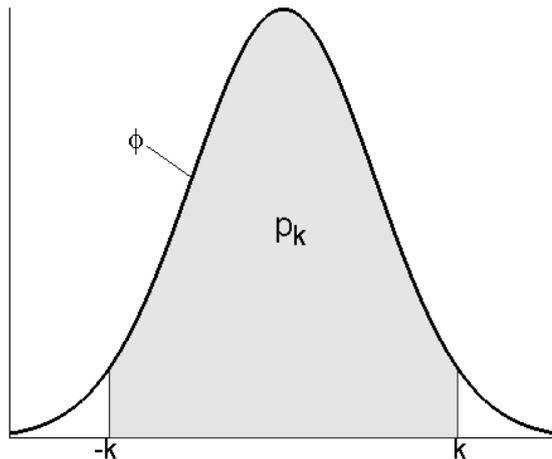


Figure 2.11 Relation between  $k$  and  $p_k$  according to equation (2.33). Function  $\phi$  is the standard normal density function

For example, a probability of 95% would correspond with a  $k$  value of 1.96, because 95% of samples from a standard normal distribution function have a value between -1.96 and 1.96. In general terms,  $k$  is defined as:

$$k = \Phi^{-1}\left(\frac{1 + p_k}{2}\right). \quad (2.33)$$

Where  $p_k$  is the desired probability that the actual error is within the defined acceptable range. The required number of samples  $N$  can be estimated with the following formula (Melchers, 2002):

$$N = \frac{k^2}{\varepsilon^2} \left( \frac{1 - P_f}{P_f} \right). \quad (2.34)$$

Where  $\varepsilon$  is the acceptable *relative error* in the estimate of  $P_f$ :

$$\varepsilon = \left( \frac{|\hat{P}_f - P_f|}{P_f} \right) \quad (2.35)$$

Where  $\hat{P}_f$  is the Monte Carlo estimator of failure probability  $P_f$ .

Note that the required number of samples depends on the failure probability, which is not known in advance. Therefore, an estimate of the order of magnitude of the failure probability must be assumed, which can subsequently be revised during the Monte Carlo sampling procedure.

Table 2.3 shows the required number of samples for combinations of  $\varepsilon$  and  $p_f$ . The value of  $k$  in this example is taken equal to 1.96. The numbers from this table show that, taking into account that each sample involves an evaluation of the Z-function, crude Monte Carlo is a rather inefficient method (i.e. a large number of Z-function evaluations is required) especially for estimating small failure probabilities.

Table 2.3 Required number of samples with crude Monte Carlo for combinations of the acceptable relative error  $\varepsilon$  and actual probability of failure  $p_f$ . The value of  $k$  in equation (2.33) is taken equal to 1.96.

	$\varepsilon$		
$P_f$	0.10	0.05	0.01
1E-02	4E+04	2E+05	2E+07
1E-03	4E+05	2E+06	2E+08
1E-04	4E+06	2E+07	2E+09
1E-05	4E+07	2E+08	2E+10

### 2.3.4 Monte Carlo Importance Sampling

#### 2.3.4.1 General description

Importance sampling is a method to increase the efficiency of the crude Monte Carlo method; that is, to decrease the number of samples and Z-function evaluations required to produce a reliable estimate of the failure probability. This is done by replacing the initial probability density,  $f_x$ , of the input variables by a more efficient one,  $h_x$ , in which “efficient” refers to the proportion of the samples which will result in failure. An increasing percentage of samples in the failure domain results in a reduction in the variance of the estimator of the failure probability, hence a smaller number of samples is required for a reliable estimate.

There are a number of ways in which importance sampling can be applied; two of these are described in this section. The first increases the variance of the density function, resulting in a higher likelihood that failure events are sampled. The second essentially shifts the density function towards the failure domain so that, again, the likelihood of a failure sample increases. These two methods are illustrated in Figure 2.12; the left-hand side illustrates the concept of a shifting of the density function, and the right-hand side illustrates the concept of increased variance.

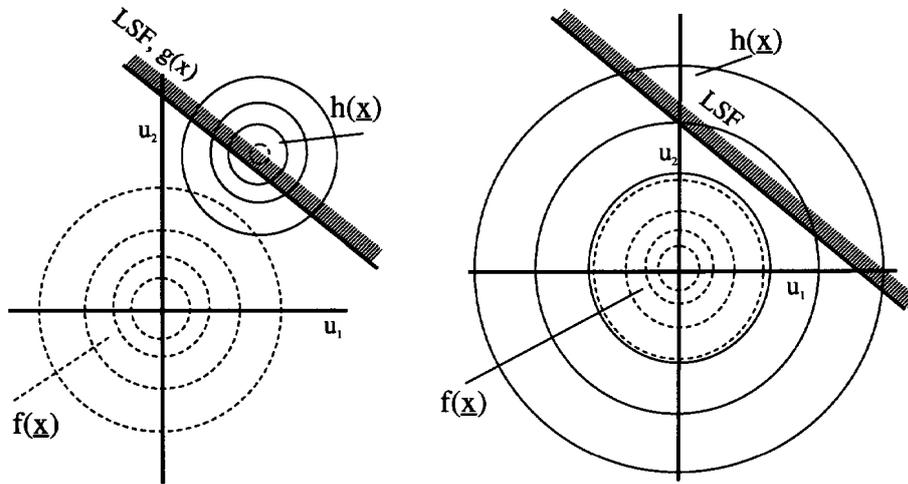


Figure 2.12 – Concept of importance sampling; left shows the concept of density shifting, right shows the concept of increased variance. The limit state function (LSF) is illustrated. Note that the LSF's are linear in this figure, but this is by no means a requirement for the applicability of importance sampling.

Because the sampling hasn't taken place from the initial distribution, the typical estimator of the failure probability (see equation (2.31)) needs to be corrected for this fact. This is done via the following formula:

$$\hat{P}_f = \frac{\sum_{i=1}^N I(Z(\mathbf{x}^i)) \frac{f_x(\mathbf{x}^i)}{h_x(\mathbf{x}^i)}}{N} \tag{2.36}$$

where  $P_f$  is the estimated probability of failure,  $I$  is the indicator function (equal to unity when  $Z < 0$ , equal to zero when  $Z \geq 0$ ),  $N$  is the total number of samples taken,  $f_x$  is the density function of  $x$  and  $h_x$  is the importance sampling density function.

Equation (2.36) can be explained by comparing it with equation (2.31), which describes the crude Monte Carlo method. In both equations, the indicator  $I$  is equal to one if the sampled vector  $x^i$  is in the failure domain and equal to 0 if  $x^i$  is outside the failure domain. In the *crude* Monte Carlo method, each sampled failure event “scores” a point and the more points scored, the higher the estimated probability of failure. In the importance sampling method it needs to be taken into account that the sampling of vector  $x^i$  was influenced by the fact that the density function was changed:  $h_x(x)$  was applied instead of the real density function  $f_x(x)$ . This means the probability of sampling  $x^i$  was increased by a factor  $c = h_x(x^i)/f_x(x^i)$ . This manipulation in the density function needs to be compensated for in the “scoring”. Therefore, a sampled event  $x^i$  in the failure domain does not score a full point, but “only”  $1/c = f_x(x^i)/h_x(x^i)$ .

So, the difference between equation (2.36) and (2.31) is the correction term  $f/h$ . This correction is necessary to make the estimate of  $P_f$  unbiased (provided  $h$  is well chosen) and accordingly that the error in the estimate can be made as small as desired by taking a sufficiently large number of samples,  $N$ . For importance sampling there is no simple generic error estimate like equation (2.34) for Crude Monte Carlo sampling, because the error estimate depends on the choice of  $h_x(x)$ . The efficiency of importance sampling therefore also

depends strongly on the choice of  $h_X(x)$ . Prior knowledge of the problem under consideration is therefore very valuable to be able to define an efficient importance sampling method. Without such knowledge, there is even the potential danger that the important area for the limit state function (LSF) will be missed.

#### 2.3.4.2 Example: implementation of the method of increased variance

The implementation of this general formula in Hydra-Ring for the case of increased variance will now be described. Equations (2.37) through (2.42) show how the formula programmed in Hydra-Ring can be derived from the general expression in equation (2.36). A random sampling of standard normal variables is first done; let's refer to these variables as  $U_1$ , where  $U_1$  is the vector containing all the variables. Subsequently, each  $u_1$ -value is multiplied by a constant factor,  $a$ , to obtain a sample  $u = a \cdot u_1$ . Thus, the initial distribution of each  $u_1$ -value is standard normal. The new set of variables,  $U$ , then have a normal distribution with a mean value equal to zero and a standard deviation equal to  $a$ . The general form of the normal distribution is given below for reference.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (2.37)$$

The ratio  $f(u)/h(u)$  is then derived as follows, shown for the case of one variable:

$$f(u) = \frac{1}{\sqrt{2\pi(1)^2}} \exp\left[-\frac{(u-0)^2}{2(1)^2}\right] = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right] \quad (2.38)$$

$$h(u) = \frac{1}{\sqrt{2\pi(a)^2}} \exp\left[-\frac{(u-0)^2}{2(a)^2}\right] = \frac{1}{\sqrt{2\pi a^2}} \exp\left[-\frac{u^2}{2a^2}\right] \quad (2.39)$$

$$\frac{f(u)}{h(u)} = \frac{\frac{1}{\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right]}{\frac{1}{\sqrt{2\pi a^2}} \exp\left[-\frac{u^2}{2a^2}\right]} = a \frac{\exp\left[-\frac{u^2}{2}\right]}{\exp\left[-\frac{u^2}{2a^2}\right]} \quad (2.40)$$

Writing  $u$  in terms of  $u_1$ , the equation becomes:

$$\frac{f(u)}{h(u)} = a \frac{\exp\left[-\frac{u^2}{2}\right]}{\exp\left[-\frac{(au_1)^2}{2a^2}\right]} = a \frac{\exp\left[-\frac{u^2}{2}\right]}{\exp\left[-\frac{u_1^2}{2}\right]} \quad (2.41)$$

Note that the one-variable example can easily be expanded to more variables, using the property that the  $u$ -values are independent, and hence their probability densities can be multiplied for the multi-variate case. The multi-variate form of equation (2.41) is:

$$\frac{f(u)}{h(u)} = a^k \frac{\exp\left[-\frac{1}{2} \sum_k u_k^2\right]}{\exp\left[-\frac{1}{2} \sum_k u_{1k}^2\right]} \quad (2.42)$$

where  $k$  is the total number of random variables. Equation (2.42) is the form in which the ratio is programmed in Hydra-Ring.

### 2.3.5 Monte Carlo directional sampling

Directional Sampling (see, e.g. Bjerager P [1988] and Ditlevsen, Melchers and Gluwer [1990]), also referred to as Directional Simulation, is a method in which “directions” are sampled. Like importance sampling it is a type of Monte Carlo method which aims to (strongly) reduce the number of samples in comparison with the Crude Monte Carlo method.

The concept of this method is schematically depicted in Figure 2.13. For each sampled direction, it is evaluated where along this line the limit state function  $Z$  equals zero. This procedure is repeated for a number of directions that is large enough to have enough information on the limit state ( $Z=0$ ) to provide a proper estimate of the probability of failure.

For convenience, this method is generally applied in the standard normal space, using equations (2.5) to transform the standard normal variables  $U_1 \dots U_n$  to their corresponding “real world” variables  $X_1 \dots X_n$ , and vice versa. The sampling of a direction in the  $U$ -space is very straightforward. First, a vector  $u = u_1, \dots, u_n$  is sampled, where each value  $u_i, i=1 \dots n$  is taken from the standard normal distribution function. Subsequently, define a vector  $v$  as:

$$v = \frac{u}{\|u\|} \quad (2.43)$$

Clearly,  $v$  is a vector of unit length and with this procedure, each vector of unit length has an equal probability density. For each of the sampled directions,  $v$ , the distance,  $\beta$ , from the origin to the limit state ( $Z=0$ ) is derived. Subsequently, the probability of failure,  $P(Z<0)$ , is estimated as follows:

$$\hat{P}_f = \frac{1}{N} \sum_{i=1}^N P_i ; \quad \text{where: } P_i = 1 - \chi_n^2(\beta_i^2) \quad (2.44)$$

- $\hat{P}_f$  = Estimated probability of failure
- $\chi_n^2$  = Chi-square distribution with  $n$  degrees of freedom ( $n$ =number of variables)
- $P_i$  = Contribution to the failure estimate of the  $i^{\text{th}}$  sampled direction
- $\beta_i$  = Distance to the limit state ( $Z=0$ ) in the  $i^{\text{th}}$  sampled direction
- $N$  = Total number of sampled directions

Note that the Chi-square distribution with  $n$  degrees of freedom is the distribution function of the sum of the squares of  $n$  independent standard normal random variables. This means that

the probability that  $\|u\| \leq \beta_i$  is described by this distribution function, which explains equation (2.44).

The search procedure for the location on the line where  $Z=0$  is often referred to as the “line search algorithm”. This is an iterative procedure that is schematically depicted in Figure 2.13 and Figure 2.14. In both figures, the numbers 0..4 refer to successive Z-function evaluations along one direction. Figure 2.13 shows 3 evaluated directions in the  $u$ -space, whereas Figure 2.14 shows the (fictional) value of the Z-function along one direction (where  $\|u\|$  is the distance from the origin along the line). As long as the Z-function is “smooth”, generally only a couple of Z-function evaluations are required to reach the location where  $Z$  is acceptably close to 0. For instance, after evaluation 3, this location can be estimated by linear interpolation between the results of evaluations 2 and 3. Sometimes a better estimate can be provided by applying quadratic interpolation, using three previous Z-function evaluations.

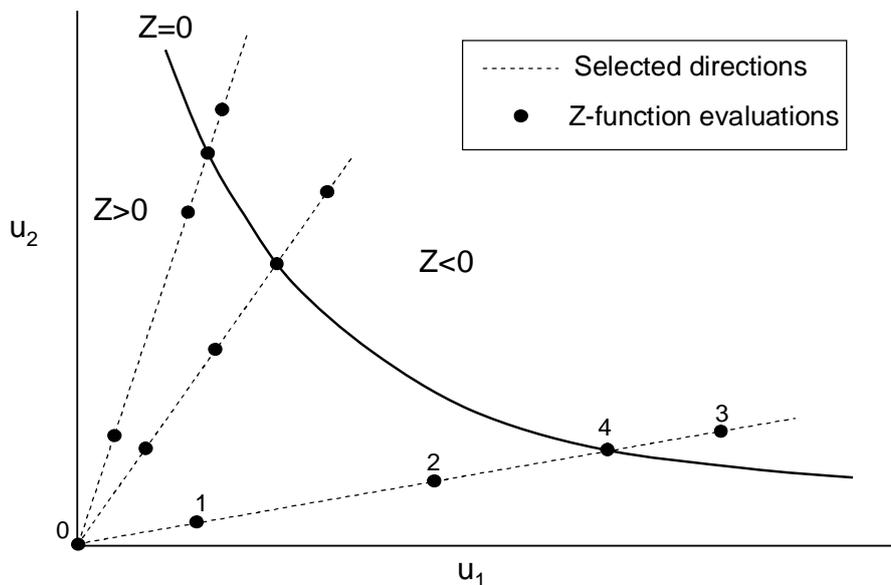


Figure 2.13 Schematic view of the directional sampling method in the  $u$ -space. The numbers 0 ..4 show successive Z-function evaluations along one direction.

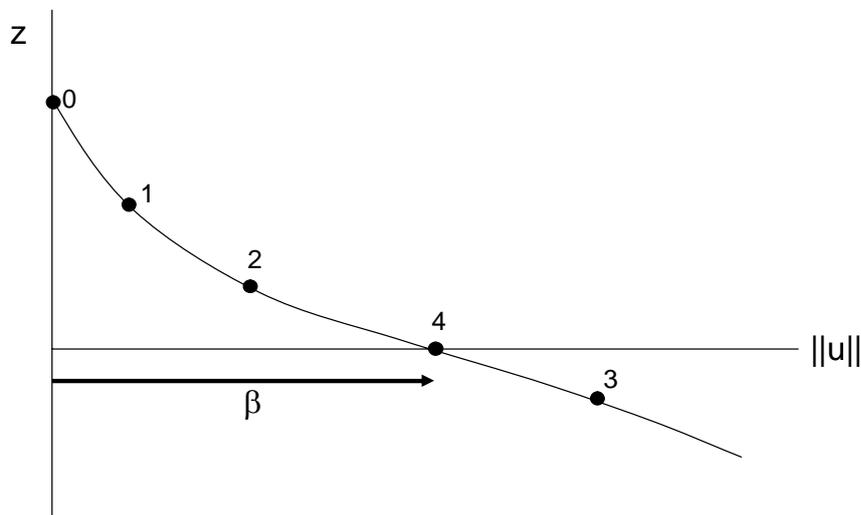


Figure 2.14 Schematic view of the line search algorithm.  $\beta$  is the distance from the origin in the  $u$ -space along one direction. The numbers 0 ..4 correspond to the ones shown in Figure 2.14.

Similar to other Monte Carlo methods, the outcome of the estimated probability of failure is a random variable and the error in the estimate can be made as small as possible by taking a sufficient number of samples. For directional sampling, the standard deviation,  $\sigma$ , of the estimated probability of failure can be quantified as follows (see, e.g. Grooteman, 2011, Melchers, 2002, pp 84):

$$\sigma_{\hat{p}_f} = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (P_i - \hat{P}_f)^2} \quad (2.45)$$

From this equation, relative errors and confidence intervals can be estimated. Note that the error in the estimated failure probability is a random variable that approaches a normal distribution function, as  $N$  increases. This follows from the central limit theorem (see, e.g. Grimmet and Stirzaker [1982]). As can be seen, the error in the estimated probability of failure will decrease with increasing number of sampled directions. Equation (2.45) can be used to determine the number of sampled directions that is required for a reliable estimate of the failure probability.

Directional sampling has the advantage that, if the line search algorithm is implemented well, it is a robust method and in many applications also very efficient (i.e. a low number of  $Z$ -function evaluations). The method is less efficient if a large number of random variables are involved.

### 2.3.6 First-order reliability method (FORM)

The term first-order refers to the linearization of the limit state function, as previously described in 2.2.5. This linearization takes place at a location referred to as the design point. A "location" in this case refers to a specific realization  $x_1, \dots, x_n$  of the  $X$ -variables, or  $u_1, \dots, u_n$  of the  $U$ -variables. The design point is the location along the limit state ( $Z=0$ ) where the probability density is maximal. This location is not known in advance and is determined via an iterative procedure, which will be explained in this section.

The FORM procedure is generally executed in the standard normal space (U-variables). The standard normally distributed variables have by definition a mean value of zero and a standard deviation of 1, and are mutually independent. The advantage of working in the standard normal space is that in this space the design point has a clear interpretation. Namely, in the standard normal space, the design point is the location along the limit state ( $Z=0$ ) which is closest to the origin (see Figure 2.15 for an illustration). This can be easily explained by the fact that for standard normally distributed variables the density is highest for  $u=0$  and decreases with increasing value of  $|u|$ . So in the  $u$  space the density decreases with increasing distance from the origin. Therefore, the design point is the point along the limit state that is closest to the origin.

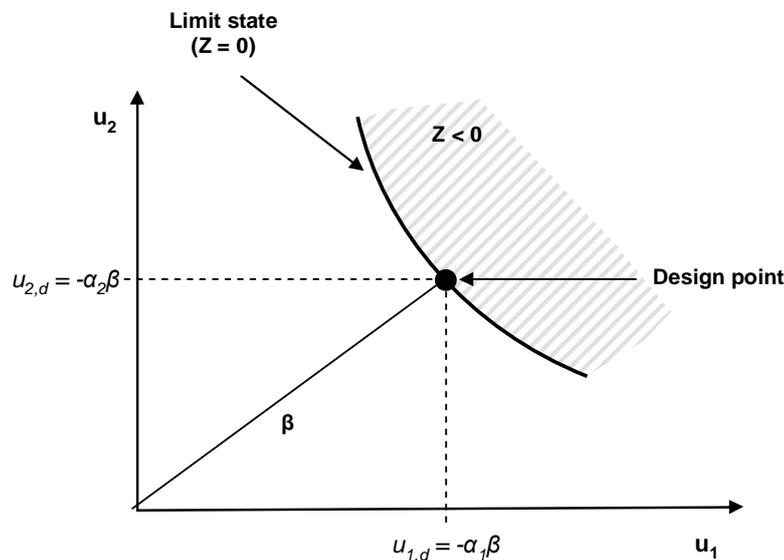


Figure 2.15 – Illustration of the design point in the  $u$ -space

The distance from the origin to the design point is equal to the reliability index  $\beta$  that was introduced in section 2.2.4. This means if the location of the design point is known, the reliability index,  $\beta$ , is known and hence the estimated probability of failure can be derived from equation (2.14). This is how the probability of failure is estimated in the FORM method. Note that this is an *estimate*, and not the precise probability of failure. This is because the limit state function ( $Z=0$ ) was linearized to provide the estimate. The error in the estimate therefore depends on the extent in which the real limit state function is non-linear. This is illustrated in Figure 2.16 below, where the solid line indicates the true limit state and the dashed line represents the FORM approximation. The shaded area represents the true failure domain, the area to the upper right of the dashed line is the assumed failure domain.

The reason why FORM in general provides good estimates of the failure probability is the fact that the linearization is done in the design point, which means in the vicinity of the design point the linear  $Z$ -function is a good approximation of the real  $Z$ -function. The design point is the location on the limit state with the highest probability density. This means the failure events with the highest probability of occurrence will generally be in the vicinity of the design point. So, the linearised  $Z$ -function is generally a good approximation for the areas that matter the most in terms of probability of failure.

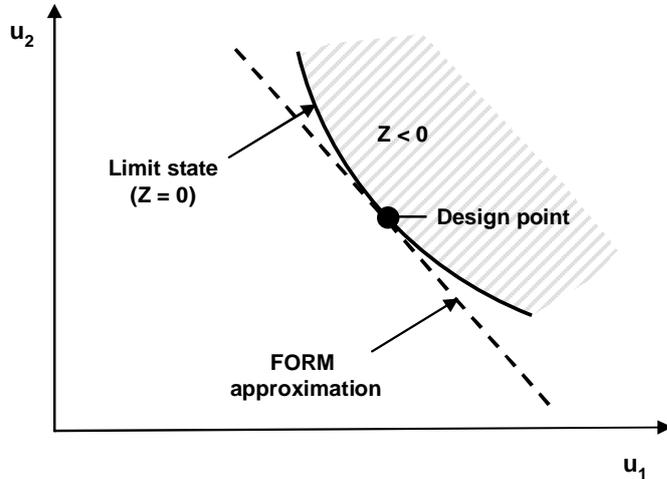


Figure 2.16 – Illustration of the FORM approximation

As demonstrated in section 2.2.5, the linearized limit state function  $Z_L$  is essentially a hyper-plane and is described mathematically as follows:

$$Z_L = b + a_1 u_1 + \dots + a_n u_n = b + \sum_{i=1}^n a_i u_i \quad (2.46)$$

Note that the  $\alpha$ -values have been normalized, as described in section 2.2.5. This means that the sum of the squares of the  $\alpha$ -values is equal to 1. In the remainder of this report,  $\alpha$ -values are always assumed to be normalized in case they are used as coefficients in an equation with  $u$ -variables (unless mentioned otherwise). Given that the design point is the point along the  $Z = 0$  line that lies closest to the origin, the coordinates of the design point can be determined using geometry as follows:

$$u_{d,i} = -\alpha_i \beta \quad ; i = 1..n \quad (2.47)$$

Where  $u_{d,i}$  is the value of the  $i^{\text{th}}$  random variable in the design point. As demonstrated in section 2.2.4, the probability of failure can be estimated directly from the value of  $\beta$ :

$$P(Z < 0) = F(-b) = 1 - F(b) \quad (2.48)$$

Where  $\Phi$  is the standard normal distribution function. In other words, once the design point is known, the  $Z$ -function can be linearised as described in section 2.2.5, and subsequently the probability of failure can be estimated from equation (2.48). The challenge in the FORM procedure is therefore not to compute the failure probability but to locate the design point.

The procedure to locate the design point is described below. The procedure and formulas will be presented in general form. Furthermore, the procedure will be clarified in a number of figures for the following example  $Z$ -function:

$$Z = 5 - u_1^{0.8} u_2^{1.2}$$

Where  $U_1$  and  $U_2$  are standard normal random variables. Figure 2.17 shows the contourlines of this Z-function. Generally, these contour lines are not known in advance, otherwise the search for the design point would be straightforward. Therefore, an iterative search procedure is required. The procedure starts at a “user defined” starting location in the u-space and jumps to a selected location in each following iteration step. In other words: in each iteration step the location in the u-space is determined that will serve as the starting point for the next iteration step. The procedure ends when the design point is found. Each iteration step consists of the following five sub-steps:

*The five steps in a FORM iteration.*

- [1] Linearisation of the Z-function in  $u^t$ , where  $u^t$  is the starting location of iteration t;
- [2] Normalisation of the linearised Z-function in  $u^t$ ;
- [3] Estimation of the location of the design point, based on the Z-function of step 2;
- [4] Selection of location  $u^{t+1}$ , which will serve as the starting location of iteration t+1;
- [5] Verification if the iteration procedure has converged.

These five steps are described in more detail below. Note that location  $u^t$  refers to a vector of u values:  $u^t = (u_1^t, \dots, u_n^t)$ .

[1] The starting location of each iteration, t, is determined in the previous iteration, t-1. The starting location in the first iteration step can either be selected “arbitrarily” or by more advanced methods in which the U-space is partially explored in advance of the FORM procedure. In the current example the starting location in the first iteration step is chosen to be  $u_i=1; i=1\dots n$  (red dot in Figure 2.17).

In each iteration, first the Z-function is determined for the selected location at the beginning of the iteration, i.e.  $Z(u_1, \dots, u_n)$  is quantified. Subsequently, the Z-function is linearised in the current location. For this purpose, the partial derivatives of Z to the individual U-variables are quantified. Generally, the Z-function is too complex to have an analytical expression of the partial derivatives, which means a numerical estimation technique is required. For this purpose, the Z-function is evaluated for small perturbations ( $\Delta u$ ) of the u-values as shown in Figure 2.18. The partial derivatives can then be estimated as follows:

$$\frac{\partial Z}{\partial u_i}(u_1, \dots, u_n) \approx \frac{Z(u_1, \dots, u_i + Du_i, \dots, u_n) - Z(u_1, \dots, u_n)}{Du_i} ; i = 1..n \quad (2.49)$$

Note that equation (2.49) describes a one-sided discretisation method. Hydra-Ring actually uses a two-sided method, in which also a negative perturbation is applied on  $u_i$ :

$$\frac{\partial Z}{\partial u_i}(u_1, \dots, u_n) \approx \frac{Z(u_1, \dots, u_i + 0.5Du_i, \dots, u_n) - Z(u_1, \dots, u_i - 0.5Du_i, \dots, u_n)}{Du_i} ; i = 1..n \quad (2.50)$$

A two-sided method is generally more robust, but requires approximately twice as much computation time. The linearised Z-function is described by:

$$Z_L = B + A_1 u_1 + \dots + A_n u_n \quad (2.51)$$

In which the A-values are the partial derivatives as derived in equation (2.49) and B is derived by substituting the known Z-value in the current location ( $u_1, \dots, u_n$ ):

$$B = Z(u_1, \dots, u_n) - A_1 u_1 - \dots - A_n u_n \quad (2.52)$$

The linearised function is (temporarily) assumed to be valid for the entire U-space. This results in linear contour lines as shown in Figure 2.19.

[2] Subsequently, the linearised Z-function is normalized by dividing equation (2.51) by  $\|A\|$ , i.e. the norm of the A-vector (as earlier described in section 2.2.5). The normalized linear Z-function is described as:

$$Z_L = b + a_1 u_1 + \dots + a_n u_n \quad (2.53)$$

In which:

$$b = \frac{B}{\|A\|}; \quad a_i = \frac{A_i}{\|A\|}, i = 1..n; \quad \|A\| = \sqrt{\sum_{i=1}^n A_i^2} \quad (2.54)$$

The normalization changes the contour lines of the linearised Z-function (compare Figure 2.19 with Figure 2.20). The orientation of the lines is still the same, but the distances between the contour lines have changed. The location of the contour line  $Z=0$ , however, remains the same.

[3] From the linear contour lines it is easy to estimate the location of the design point. This is done by drawing the line through the origin that is perpendicular to the contour line  $Z_L=0$  (see Figure 2.21). In formula this means the estimated location of the design point is as described in equation (2.47). The values of  $\alpha$  and  $\beta$  in equation (2.47) are set equal to the ones derived from equation (2.54).

[4] The estimated location of the design point can be chosen as the next location in the iteration procedure. Note, however, that this is most likely not the actual location of the design point, since it was derived from the linearised Z-function and not from the real Z-function. This is the reason why the design point will not be located straight away, i.e. a number of iteration steps are required. This is also the reason why in practical applications generally a relaxation parameter,  $r$ , is used in each iteration step:

$$u^{t+1} = r u_d^t + (1 - r) u^t \quad (2.55)$$

In which:

- $t$  = iteration step
- $r$  = the relaxation parameter ( $0 \leq r \leq 1$ )
- $u^t$  = the selected location at the beginning iteration step  $t$
- $u^{t+1}$  = the selected location at the beginning iteration step  $t+1$
- $u_d^t$  = the estimated location of the design point in iteration step  $t$

The functionality of the relaxation parameter can be explained as follows: in each iteration step, the Z-function is linearised in location  $u^t$ . The linearised function,  $Z_L$  is the tangent of the

actual Z-function at location  $u^t$ . In the vicinity of  $u^t$ ,  $Z_L$  is generally a good approximation of  $Z$ . However, with increasing distance from  $u^t$ , differences between  $Z$  and  $Z_L$  may increase, as can be seen from e.g. Figure 2.8. Since the estimated location of the design point,  $u_d^t$ , is based on  $Z_L$ , this estimate may be unreliable if the distance between  $u^t$  and  $u_d^t$  is large. This might even lead to non-convergence of the iteration procedure. It is therefore better to prevent that the distance between two subsequent iteration steps becomes too large, and for this reason the relaxation parameter is used. The relaxation parameter helps making the iterative procedure more robust.

Figure 2.22 demonstrates the application of the relaxation parameter. It shows the location at the beginning of the iteration,  $u^t$ , (red dot), the estimated location,  $u_d^t$ , of the design point (yellow dot) and the location at the beginning of the next iteration,  $u^{t+1}$  (green dot). Location  $u^{t+1}$  is chosen somewhere on the line between the current location and the estimated location of the design point. For values of  $r < 0.5$ ,  $u^{t+1}$  will be closer to  $u^t$  for values of  $r > 0.5$ ,  $u^{t+1}$  will be closer to  $u_d^t$ .

[5] Figure 2.23 shows the resulting iteration steps of the example problem. The iteration procedure continues until location  $u_d^t$  satisfies the following 2 criteria:

$$\frac{|Z(u_d^t)|}{\|A\|} = |Z_L(u_d^t)| < e_1 \quad (2.56)$$

$$b^t - e_2 < \|u_d^t\| < b^t + e_2 \quad (2.57)$$

Where

$\varepsilon_{1,2}$  = Small numbers, quantifying convergence criteria

$\beta^t$  = Estimate of reliability index  $\beta$  in iteration step 2

Criterion (2.56) guarantees that the Z-function is sufficiently close to 0, i.e. that  $u_d^t$  is on (or in the neighbourhood of) the limit state  $Z=0$ . Note that for this purpose the value of  $Z$  is normalized by dividing it by the norm of the vector of A-values. The second criterion guarantees that the distance from  $u_d^t$  to the origin is (approximately) equal to the estimated reliability index beta, which makes  $u_d^t$  the point on  $Z=0$  with the highest probability density.

The FORM procedure has the advantage that it requires relatively little computation time, i.e. a relatively small number of Z-function evaluations. The disadvantage of this method is that the iterative algorithm sometimes does not converge and results are not always reliable. This is especially the case if the Z-function is highly non-linear.

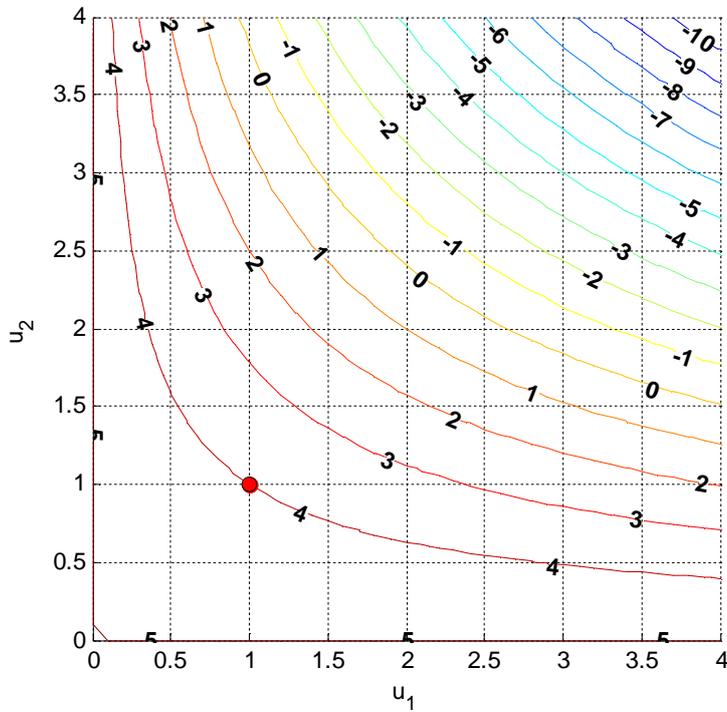


Figure 2.17 Contour lines of the example Z-function and the starting location (red dot) of the FORM procedure

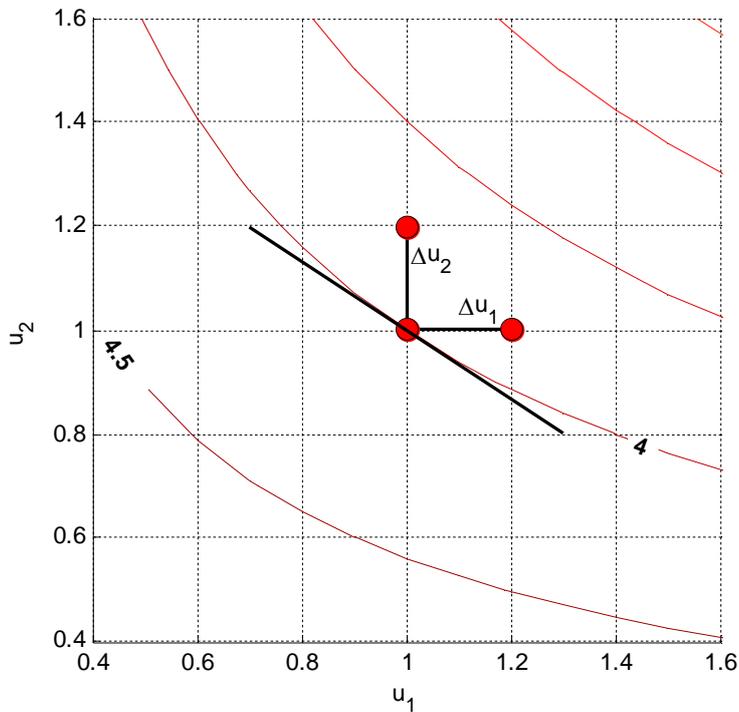


Figure 2.18 Sampling the Z-function in all directions to estimate the derivative of Z to all u-variables.

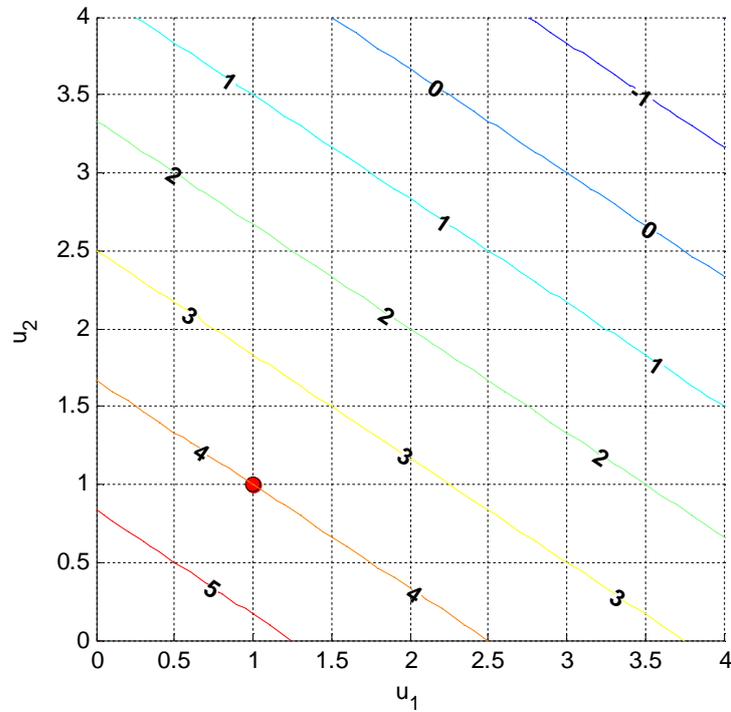


Figure 2.19 Contour lines of the linearised Z-function

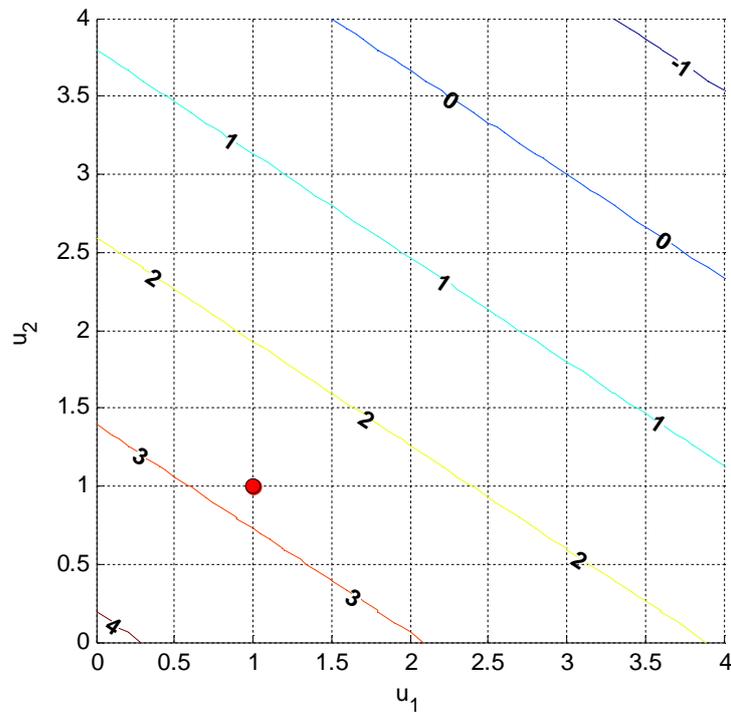


Figure 2.20 Contour lines of the normalised linearised Z-function

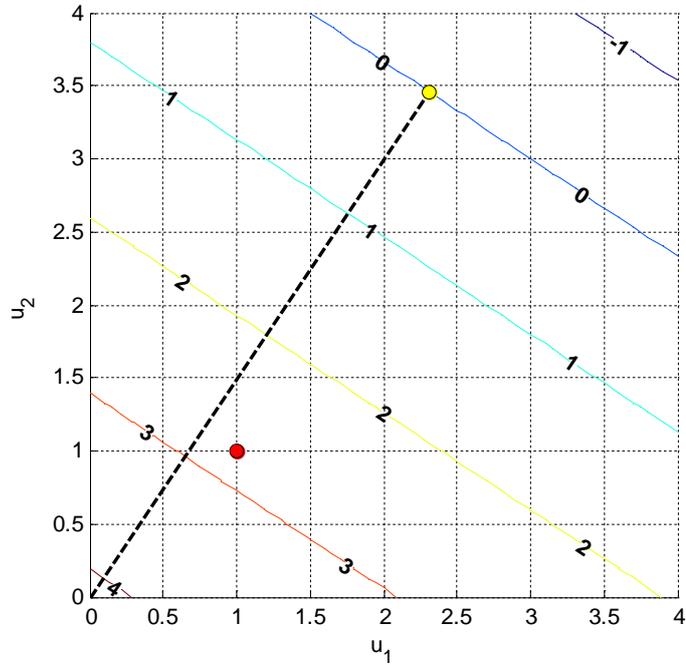


Figure 2.21 Estimated location of the design point based on the normalised linearised Z-function

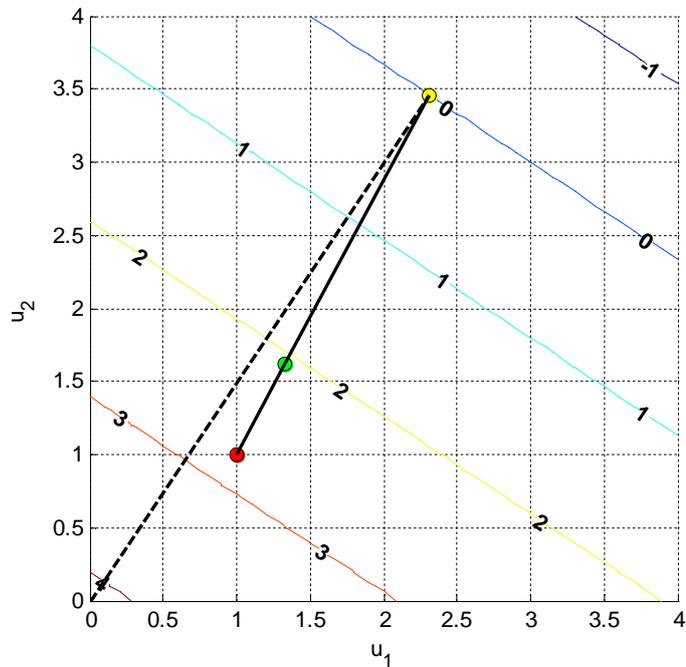


Figure 2.22 Proces of relaxation. The red dot is the location of the current iteration step, the yellow dot is the estimated location of the design point based on the normalised linearised Z-function. The green dot shows

the selected location of the next iteration, which is somewhere on the line between the red dot and the yellow dot.

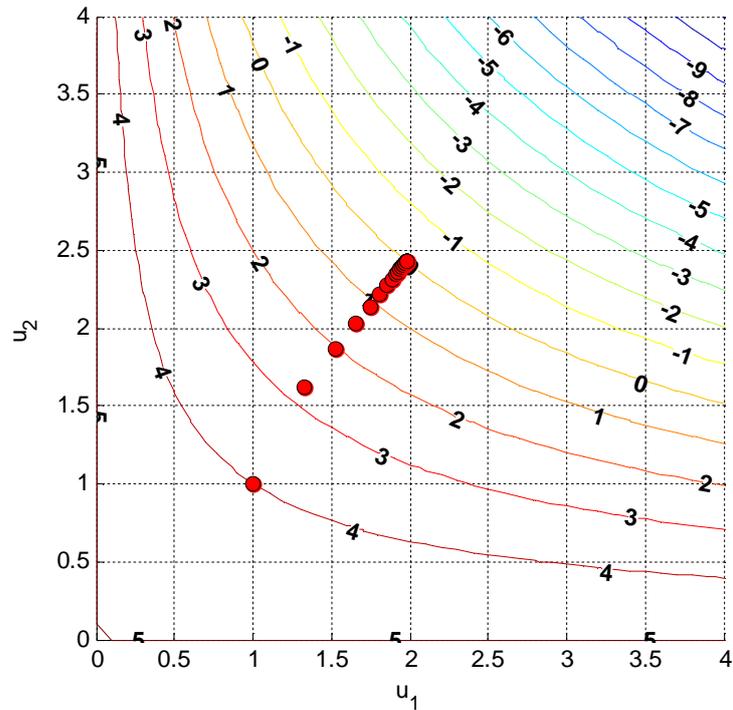


Figure 2.23 Resulting steps in the iteration procedure.

### 2.3.7 Computing $\alpha$ values for other methods than FORM

In the previous section it was demonstrated that the FORM procedure not only provides an estimate of the probability of failure, but also a design point with a set of associated  $\alpha$ -variables that provide information on the relative influence of the random variables on the reliability index  $\beta$  (see also section 2.2.5 on the meaning of  $\alpha$ -variables). In section 2.4 it will be demonstrated that these  $\alpha$ -variables are very practical for estimating failure probabilities of systems that consist of a set of components.

The other probabilistic techniques that were described in the previous sections, i.e. numerical integration and the various Monte Carlo techniques, do not provide  $\alpha$ -variables as output. Nevertheless, there are methods available to estimate  $\alpha$ -variables for Monte Carlo methods and numerical integration (see, e.g., Van Gelder [2002]). For example, for Monte Carlo, the following methods can be applied.

- 1 centre of gravity
- 2 method of angles
- 3 nearest to the mean.

The third method can also be used for numerical integration, as will be explained below. These methods all take into account the fact that the design point is the location in the failure domain that is closest to the origin (in the U-space). For all methods, the quality of the estimates increases with increasing number of samples. The methods are explained below:

### Centre of gravity

Suppose a crude Monte Carlo run is done with  $N$  samples of which  $M$  lead to failure. This means there are  $M$  sampled combinations of  $u_1, \dots, u_n$ , for which  $Z(u_1, \dots, u_n) < 0$ . For each of the  $n$  random variables the mean value over the  $M$  failure-events is derived as follows:

$$\bar{u}_j = \frac{1}{M} \sum_{i=1}^M u_{ij} \quad ; j = 1..n \quad (2.58)$$

where  $u_{ij}$  is the  $i^{\text{th}}$  "failure" sample of the  $j^{\text{th}}$  random variable. The resulting point  $\bar{u} = (\bar{u}_1, \dots, \bar{u}_n)$  is the "centre of gravity" in the failure domain in the  $u$ -space. This is an estimate of the "probability weighted mean" of the locations in the failure domain. Note that the actual probability weighted mean is equal to:

$$\bar{u}_j = \frac{1}{P[Z < 0]} \int_{Z < 0} f_U(u) u_j du \quad (2.59)$$

From the estimated centre of gravity this location a line is drawn towards the origin in the  $u$ -space (see Figure 2.24). The location where this line crosses the limit state ( $Z=0$ ) is the estimated location of the design point. This guarantees that the first characteristic of the design point, i.e. that it is located on the limit state  $Z=0$ , is taken care of. The second characteristic, that it is the location on the limit state with the highest density is not guaranteed. However, the use of the centre of gravity makes that the estimated location of the design point is likely to be close to the real design point. The likelihood increases with increasing number of samples.

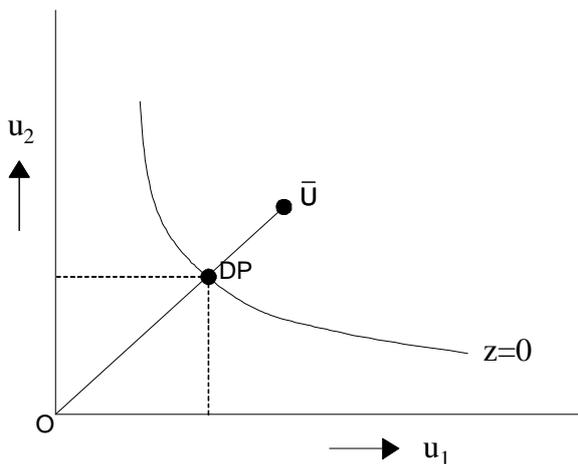


Figure 2.24 Schematic view of the method "centre of gravity".

If importance sampling has been applied in the sampling procedure, the method needs to be corrected for:

$$\bar{u}_j = \frac{1}{M} \sum_{i=1}^M u_{ij} \frac{f_U(u_{i1}, \dots, u_{in})}{h_U(u_{i1}, \dots, u_{in})} \quad ; i = 1..n \quad (2.60)$$

Where  $f_U$  is the probability density function of vector  $U$  and  $h_U$  is the applied density function of the importance sampling method.

#### *Method of angles*

The method of angles is similar to the method of centre of gravity. For each of the  $M$  samples that lead to failure the angle with the origin in the  $u$ -space is derived. After completion of the MC-procedure the "mean angle" of all  $M$  samples is derived. Note that the mean is derived with respect to the sine and cosine of the angles of all samples.

#### *Method "nearest to the mean"*

In the method of "nearest to the mean" the distance to the origin of all samples in the failure domain is derived:

$$|u_i| = \sqrt{\sum_{j=1}^n u_{ij}^2} \quad ; i = 1..M \quad (2.61)$$

The sample with the smallest distance to the origin is taken to be the design point. This method can also be applied for other Monte Carlo techniques (directional sampling, importance sampling) and even for numerical integration. In the latter case, the design point is taken equal to the grid point in the failure domain that is closest to the origin in  $u$ -space.

### 2.3.8 Rationale

In Hydra-Ring, a variety of probabilistic techniques has been implemented, including first- and second-order reliability methods (FORM and SORM), various Monte-Carlo techniques (crude, directional sampling, importance sampling) and numerical integration. Each of these techniques requires a considerable number of evaluations of the Z-function at (randomly) selected  $x$ -values. The choice of the most suitable probabilistic computation technique depends on the problem under consideration.

If the computation time of one Z-function evaluation is significant, crude Monte Carlo and numerical integration are generally not the preferred candidates because both methods generally require a large number of Z-function evaluations. For crude Monte Carlo, the required number of Z-function evaluations is inversely proportional to the failure probability. This is because a small probability of failure means it takes a large amount of samples to obtain even a single failure event and it takes more than one failure event to obtain a reliable estimate of the failure probability. For numerical integration, the number of Z-function evaluations is defined by the number of random variables and the number of grids for each random variable. Generally, numerical integration is too time-consuming if more than just a few random variables are involved. In theory, Monte Carlo and numerical integration are exact methods but in practice some error can be expected because the number of Z-function evaluations is limited.

Directional Sampling is a more advanced Monte Carlo method in comparison with crude Monte Carlo. For most practical problems it reduces the amount of Z-function evaluations in comparison with crude Monte Carlo. For a large number of random variables, the efficiency of directional sampling decreases (see e.g. Waarts, 2000, pp 73). Importance sampling is another efficient Monte Carlo variant. The efficiency of importance sampling is

accommodated by prescience of the location of the limit state function. Without prescience, the performance of importance sampling techniques is volatile.

FORM has the advantage that it requires relatively little computation time. The disadvantage of this method is that the iterative algorithm to find the design point sometimes does not converge or converges to a “local” design point. Furthermore, the Z-function is linearised in the method, which means errors are introduced if the actual Z-function is highly non-linear.

Combining two different probabilistic methods may result in the combined advantage of the underlying methods. For instance, a relatively fast method like FORM can be applied first to locate the design point and subsequently a more precise method like importance sampling can be applied to derive the probability of failure by sampling in the vicinity of the design point. Or, vice versa, Monte Carlo sampling can be applied to provide a starting point for FORM in the neighbourhood of the design point, to increase the chance that FORM converges to the correct design point. This increases the robustness of the FORM procedure.

## 2.4 Combining failure probabilities for components - generic methods

### 2.4.1 Introduction

In section 2.2.1, the concept of system analysis was explained, with special emphasis on parallel systems (the system fails only if all components of the system fail) and series systems (the system fails if one or more components of the system fail). The general formulations of failure probabilities for parallel and series systems of  $k$  components are as follows:

$$\text{Series: } P_f = P[Z_1 < 0 \cup \dots \cup Z_k < 0] = P\left[\bigcup_{i=1}^k Z_i < 0\right] = 1 - P\left[\bigcap_{i=1}^k Z_i \geq 0\right] \quad (2.62)$$

$$\text{Parallel: } P_f = P[Z_1 < 0 \cap \dots \cap Z_k < 0] = P\left[\bigcap_{i=1}^k Z_i < 0\right] = 1 - P\left[\bigcup_{i=1}^k Z_i \geq 0\right] \quad (2.63)$$

If the events  $[Z_i < 0]$ ,  $i=1..k$  are mutually independent, this can be simplified to:

$$\text{Series: } P_f = 1 - \prod_{i=1}^k \{1 - P[Z_i < 0]\} \quad (2.64)$$

$$\text{Parallel: } P_f = \prod_{i=1}^k P[Z_i < 0] \quad (2.65)$$

The failure probabilities,  $P[Z_i < 0]$ , for the individual components are determined by the probabilistic computation techniques as described in section 2.3. System analysis for mutually independent components is therefore a relatively straightforward procedure. However, if the components are mutually correlated, the complexity of the system analysis increases. The correlations need to be taken into account as it increases the probability of failure of parallel systems and decreases the probability of failure of series systems. The following sections describe various general techniques that can be applied to carry out system analysis for systems with mutually correlated components.

### 2.4.2 Combining $n$ components: the Hohenbichler method

#### 2.4.2.1 Introduction

The Hohenbichler method initially is a method for computing conditional probabilities of two  $Z$ -functions:  $P(Z_2 < 0 | Z_1 < 0)$ , taking into account the mutual correlation between these two  $Z$ -functions. The application of this method can be extended to compute failure probabilities of:

- [1] A parallel system of two components;
- [2] A series system of two components;
- [3] Parallel and series systems of multiple components.

This is explained as follows:

[1] A parallel system with two components refers to a system in which *both* components must fail in order for failure to occur (keyword: AND). That is, the probability of failure is given as follows:

$$P(F) = P(Z_1 < 0 \cap Z_2 < 0). \tag{2.66}$$

A parallel system and the schematization of the associated failure probability ( $P(Z_1 < 0 \cap Z_2 < 0)$ ) is schematically depicted in Figure 2.25 below.

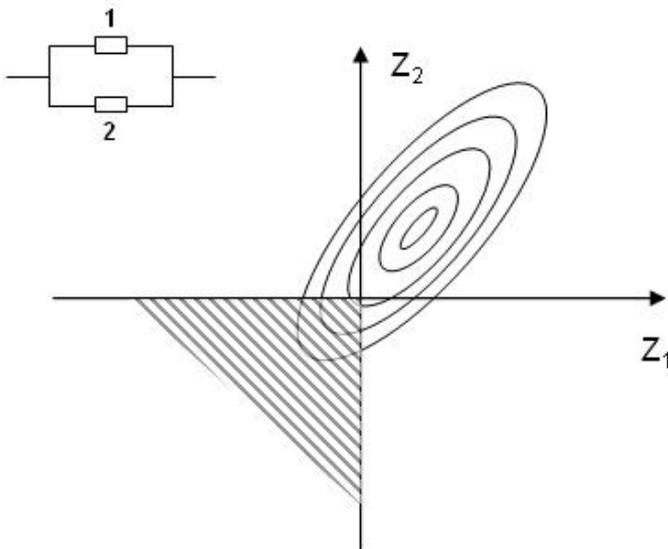


Figure 2.25 Failure domain for a parallel system of two components – the shaded area indicates the area that contributes to the failure probability.

The failure probability of this system can be written equivalently as the product of a probability and a conditional probability:

$$P(F) = P(Z_1 < 0) \cdot P(Z_2 < 0 | Z_1 < 0). \tag{2.67}$$

The first term,  $P(Z_1 < 0)$  can be computed with the methods as described in section 2.3. The second term,  $P(Z_2 < 0 | Z_1 < 0)$ , can be determined with the Hohenbichler method, as will be demonstrated in subsequent sections. This shows that the Hohenbichler method can also be applied to compute the failure probability of a parallel system of two components.

[2] A series system with two components refers to a system in which *at least one* component must fail in order for failure to occur (keyword: OR):

$$P(F) = P(Z_1 < 0 \cup Z_2 < 0) \tag{2.68}$$

This probability can be rewritten as follows:

$$P(F) = P(Z_1 < 0 \cup Z_2 < 0) = P(Z_1 < 0) + P(Z_2 < 0) - P(Z_1 < 0 \cap Z_2 < 0). \quad (2.69)$$

The first two terms on the right hand side of equation (2.69) describe failure probabilities of single components, which can be derived with the techniques that were described in section 2.3. The last term describes a parallel system of two components, for which the computational method was described in [1]. This shows that the Hohenbichler method can also be applied to compute the failure probability of a series system of two components.

**[3]** Consider a series system of  $n$  components. The failure probability for the system is given by:

$$P(F) = P(Z_1 < 0 \cup Z_2 < 0 \cup Z_3 < 0 \cup \dots \cup Z_n < 0) \quad (2.70)$$

If we define  $Z_{12} = Z_1 \cup Z_2$ , this equation can be rewritten as an arbitrary system of  $n-1$  components:

$$P(F) = P(Z_{12} < 0 \cup Z_3 < 0 \cup \dots \cup Z_n < 0) \quad (2.71)$$

Repeating this procedure  $n-1$  times will result in a system of one component. So, the probability of failure for a system of  $n$  components can be derived by the successive combining of combinations of two components. The method of Hohenbichler can be used to combine two components, as demonstrated above, and therefore it can also be used to combine  $n$  components of a series system. A similar approach can be applied for a parallel system of  $n$  components. In other words: successive application of the method can then be used to compute the probability of failure of a system of  $n$  components, which is what is required in Hydra-Ring for computing the probability of failure for a dike ring system with multiple dike sections and failure mechanisms.

The basic principles of the Hohenbichler method for computing the conditional probability of failure of two components is described in section 2.4.2.2. The follow-up sections elaborate on some of the finer details.

Note: the Hohenbichler method makes use of linearization of the Z-functions, as described in section 2.2.5. The probability of failure for a system as derived by the method of Hohenbichler is therefore an approximation of the real probability of failure. Errors made in the approximation will depend on the system under consideration.

#### 2.4.2.2 Probability of failure for a parallel system of two components

As stated in the previous section, the Hohenbichler method initially is a method for computing conditional probabilities of two Z-functions:  $P(Z_2 < 0 | Z_1 < 0)$ , taking into account the mutual correlation between these two Z-functions. Suppose the following information is available from the single component probabilistic analysis as described in section 2.3:

- The reliability index  $\beta_1$  of  $Z_1$
- The reliability index  $\beta_2$  of  $Z_2$
- The influence variables,  $\alpha$ , for each random variable involved.

This means we can formulate the two Z-functions in the standard linearised form (see section 2.2.5):

$$Z_1 = \beta_1 + \alpha_{11}u_{11} + \dots + \alpha_{1n}u_{1n}. \quad (2.72)$$

$$Z_2 = \beta_2 + \alpha_{21}u_{21} + \dots + \alpha_{2n}u_{2n}$$

Where  $u_{ij}$  refers to the  $j^{\text{th}}$  random variable of the  $i^{\text{th}}$  Z-function. The U-variables can be different for the different Z-functions. However, for the sake of simplicity, we assume for the moment that they are the same:

$$u_{1k} = u_{2k} \quad ; k = 1..n. \quad (2.73)$$

Later on, in section 2.4.2.4 the slightly more complex case will be dealt with in which  $u_{1k} \neq u_{2k}$ . Since  $u_{1j}$  and  $u_{1k}$ ,  $j \neq k$ , are mutually independent, it can easily be verified that the correlation between  $Z_1$  and  $Z_2$  is equal to:

$$\rho(Z_1, Z_2) = \sum_{j=1}^n \alpha_{1j} \cdot \alpha_{2j} \quad (2.74)$$

The linearised Z-functions can be written (see section 2.2.5) as follows:

$$Z_1 = \beta_1 - u_1 \quad (2.75)$$

$$Z_2 = \beta_2 - u_2$$

Where  $u_1$  and  $u_2$  are potential realizations of two newly defined standard normally distributed variables  $U_1$  and  $U_2$ . Because the  $\beta$ -values in equation (2.75) are constant, the correlation between the components  $Z_1$  and  $Z_2$  is equivalent to the correlation between the variables  $U_1$  and  $U_2$ :

$$\rho(Z_1, Z_2) = \rho(U_1, U_2) = \rho \quad (2.76)$$

In other words, equation (2.75) is only valid in this case if  $U_1$  and  $U_2$  are mutually correlated with correlation coefficient  $\rho$ . To assure that this is the case,  $u_2$  is written as a function of  $u_1$ :

$$u_2 = \rho u_1 + u_2^* \sqrt{1 - \rho^2}. \quad (2.77)$$

In this equation,  $u_2^*$  is also standard normally distributed and independent of  $u_1$ . The first term in this equation represents the dependent part of  $u_2$  and the second term represents the independent part. Note in equation (2.77) that if  $\rho = 1$ , then  $u_2 = u_1$  (100% correlated), and if  $\rho = 0$ , then  $u_2 = u_2^*$  (100% uncorrelated).

To verify the applicability of equation (2.77) it needs to be shown that [1]  $u_2$  is standard normally distributed and [2] that  $u_1$  and  $u_2$  have a mutual correlation coefficient that is equal to  $\rho$ . To prove [1], we apply the following general rule (see, e.g. Grimmett and Sirzaker, 1982): If  $X$  and  $Y$  normally distributed random variables, then  $aX+bY$  is also normally distributed with a mean,  $\mu$ , and standard deviation,  $\sigma$ , equal to:

$$\begin{aligned}\mu &= a\mu_x + b\mu_y \\ \sigma &= \sqrt{a^2\sigma_x^2 + b^2\sigma_y^2}\end{aligned}\quad (2.78)$$

Application of this rule on equation (2.77), where  $u_1$  and  $u_2^*$  are both normally distributed with mean 0 and standard deviation 1, gives:

$$\begin{aligned}\mu_{u_2} &= \rho \cdot 0 + \sqrt{1-\rho^2} \cdot 0 = 0 \\ \sigma_{u_2} &= \sqrt{\rho^2 \cdot 1 + (1-\rho^2) \cdot 1} = 1\end{aligned}\quad (2.79)$$

Which proves that  $u_2$  is standard normally distributed. To prove [2], the correlation coefficient between  $u_1$  and  $u_2$  is derived. The correlation coefficient of  $u_1$  and  $u_2$  is defined as:

$$\rho(u_1, u_2) = \frac{\text{COV}(u_1, u_2)}{[\sigma(u_1)\sigma(u_2)]} = \frac{\text{COV}(u_1, u_2)}{[1 \cdot 1]} = \text{COV}(u_1, u_2) \quad (2.80)$$

The covariance of  $u_1$  and  $u_2$  is equal to:

$$\begin{aligned}\text{COV}(u_1, u_2) &= E[u_1 u_2 - \mu(u_1)\mu(u_2)] = E[u_1 u_2] \\ &= E\left[u_1 \left(\rho u_1 + u_2^* \sqrt{1-\rho^2}\right)\right] \\ &= E\left[\rho u_1^2 + u_1 u_2^* \sqrt{1-\rho^2}\right] = E[\rho u_1^2] = \rho E[u_1^2] = \rho\end{aligned}\quad (2.81)$$

Which proves that the application of equation (2.77) preserves the correlation between  $U_1$  and  $U_2$  and hence the correlation between  $Z_1$  and  $Z_2$ . The combination of equations (2.75) and (2.77) provides the following description for  $Z_2$ :

$$Z_2 = \beta_2 - \rho \cdot u_1 - u_2^* \sqrt{1-\rho^2}. \quad (2.82)$$

This expression represents a line in the  $Z_2 = 0$  plane. The hatched area in Figure 2.26 indicates the area in the  $u$ -space that contributes to the failure probability

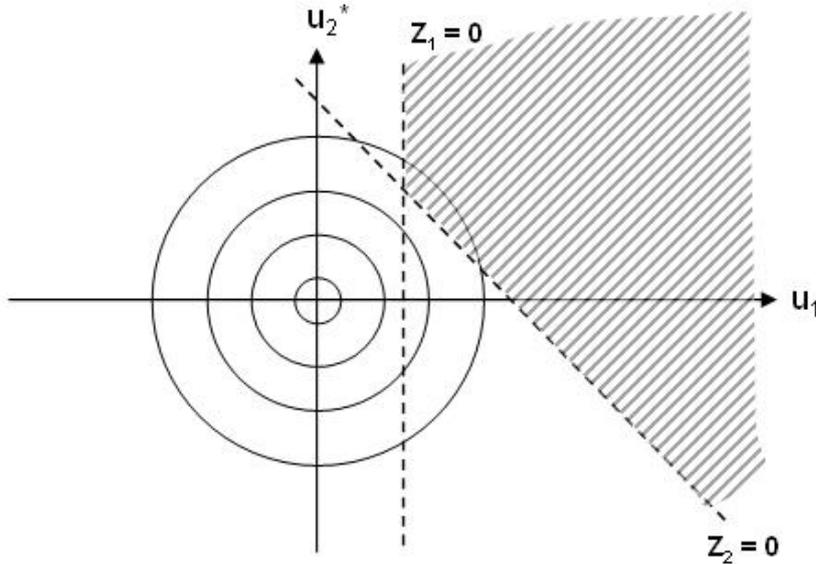


Figure 2.26 The  $Z_1 = 0$  and  $Z_2 = 0$  contours in the  $u$ -space; hatched area indicates the area that contributes to the failure probability.

Using the expressions for  $Z_1$  and  $Z_2$  (equations (2.75) and (2.82)), the conditional probability (i.e. the second term in equation (2.67)) can be rewritten as follows:

$$P(Z_2 < 0 | Z_1 < 0) = P\left(\beta_2 - \rho \cdot u_1 - u_2^* \sqrt{1 - \rho^2} < 0 | \beta_1 - u_1 < 0\right). \quad (2.83)$$

So with the assumption of linearity of functions  $Z_1$  and  $Z_2$  of two components that are combined, the conditional probability of failure,  $P(Z_2 < 0 | Z_1 < 0)$  can be written in the form of equation (2.83). At first sight this does not seem to be a simplification, but nevertheless it helps to process the conditional part as will be explained below.

Consider a realisation,  $u_1$ , from a standard normal distribution, for which:  $\beta_1 - u_1 > 0$ , i.e.  $u_1 > \beta_1$ . Even though  $u_1$  comes from a standard normal distribution, the information that  $u_1 > \beta_1$  changes the probability distribution, i.e.  $P[U_1 < u_1 | U_1 > \beta_1] \neq P[U_1 < u_1]$ . To show in which way the information that  $u_1 > \beta_1$  changes the probability distribution, consider the basic law for conditional probability:

$$P(A \cap B) = P(A | B)P(B) \Leftrightarrow P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (2.84)$$

If  $A$  is the event that  $U_1 > u_1$  and  $B$  is the event that  $U_1 > \beta_1$ , then:

$$P(U_1 > u_1 | U_1 > \beta_1) = \frac{P(U_1 > u_1 \cap U_1 > \beta_1)}{P(U_1 > \beta_1)} \quad (2.85)$$

The numerator in this equation is equal to:

$$P(U_1 > u_1 \cap U_1 > \beta_1) = \begin{cases} P(U_1 > \beta_1) & ; u_1 \leq \beta_1 \\ P(U_1 > u_1) & ; u_1 > \beta_1 \end{cases} \quad (2.86)$$

Since  $\beta_1$  is known, we can define the following constant,  $p$ :

$$p = P(U_1 > \beta_1) \Leftrightarrow \beta_1 = \Phi^{-1}(1 - p), \quad (2.87)$$

Where  $\Phi$  is the standard normal distribution function. Substitution of equations (2.87) and (2.86) in equation (2.85) gives:

$$P(U_1 > u_1 | U_1 > \beta_1) = \begin{cases} 1 & ; u_1 \leq \beta_1 \\ \frac{P(U_1 > u_1)}{p} = \frac{1 - \Phi(u_1)}{p} & ; u_1 > \beta_1 \end{cases} \quad (2.88)$$

The probability of the complement then becomes:

$$P(U_1 \leq u_1 | U_1 > \beta_1) = \begin{cases} 0 & ; u_1 \leq \beta_1 \\ 1 - \frac{1 - \Phi(u_1)}{p} = \frac{p - 1 + \Phi(u_1)}{p} & ; u_1 > \beta_1 \end{cases} \quad (2.89)$$

Taking the derivative to  $u_1$  in equation (2.89) gives the probability density function:

$$f_{U_1|\beta_1}(U_1 \leq u_1 | U_1 > \beta_1) = \begin{cases} 0 & ; u_1 \leq \beta_1 \\ \frac{\phi(u_1)}{p} & ; u_1 > \beta_1 \end{cases} \quad (2.90)$$

Where  $\phi$  is the standard normal density function. This equation shows the influence of the information that  $u_1 > \beta_1$  on the probability of the outcome of  $u_1$ . The density function for values of  $u_1 \leq \beta_1$  becomes 0, which is obvious, since we know that  $u_1 > \beta_1$ . For values of  $u_1 > \beta_1$  the original probability density is increased by a factor  $1/p$ . The concept of this change in density function is clarified in Figure 2.27. The top panel shows the original probability density of  $u_1$ , the bottom figure shows the probability density of a newly defined variable  $u_1'$ , which has a density function as described in equation (2.90). For the variable  $u_1'$ , the probability density for values lower than  $\beta_1$  is zero, since the condition  $u_1 > \beta_1$  has been incorporated. Essentially the probability density of  $u_1'$  is the tail of the distribution of the variable  $u_1$  above the threshold  $\beta_1$ , where the values have been scaled such that the area under the curve is equal to 1.

The concept of the Hohenbichler method is to take a random sample  $u_1'$  from the density function (or associated distribution function) as described in equation (2.90). The sample  $u_1'$ , which is by definition greater than  $\beta_1$ , replaces the value of  $u_1$  in equation (2.83). By this replacement, it is guaranteed that the sample will have a value that is higher than  $\beta_1$ . This means that the conditional statement in equation (2.83) can be left out. Equation (2.83) thus changes into:

$$P(Z_2 < 0 | Z_1 < 0) = P(Z'_2 < 0) = P(\beta_2 - \rho \cdot u'_1 - u_2^* \sqrt{1 - \rho^2} < 0) \quad (2.91)$$

The conditional probability with two Z-functions has now been replaced by a description with a single Z-function  $Z'$ . Note that in order to derive equation (2.83), the only approximation that was applied is the assumption that  $Z_1$  and  $Z_2$  are both linear.

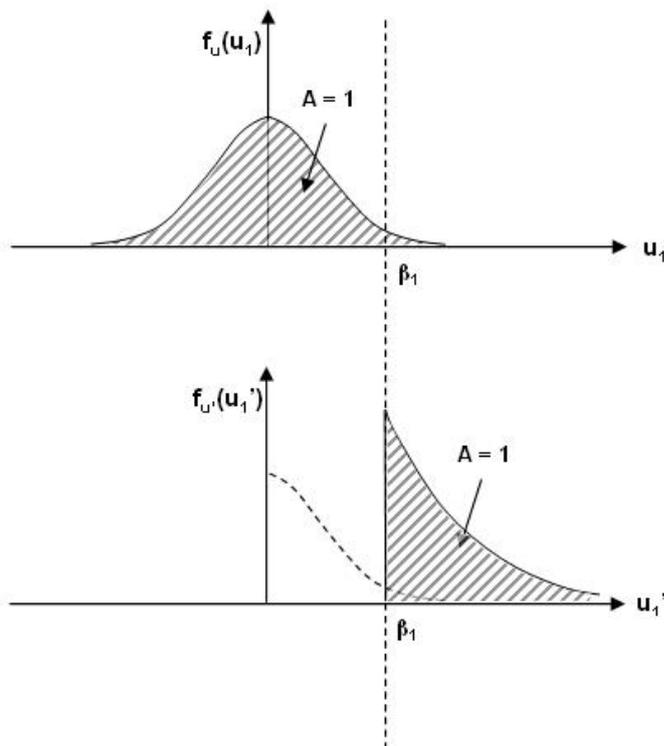


Figure 2.27 Processing the conditional part of failure probability: 'given  $u_1 > \beta_1$ '.

In order to simulate a sample  $u_1'$  from the probability density as shown in the lower panel of Figure 2.27 and as described in equation (2.90), the following sampling method can be used:

$$u_1' = \Phi^{-1}(1 - p\Phi(v)), \quad (2.92)$$

Where  $\Phi$  is the standard normal distribution function and  $v$  is a sample from the standard normal distribution function. This can be proven as follows: Variable  $v$  is standard normally distributed, in notation:  $v \sim N(0,1)$ . As a consequence,  $\Phi(v)$  is standard uniformly distributed:  $\Phi(v) \sim U(0,1)$ . This means that  $p\Phi(v) \sim U(0,p)$  and  $1-p\Phi(v) \sim U(1-p,1)$ . In other words,  $u'$  can be written as:

$$u_1' = \Phi^{-1}(p') \quad (2.93)$$

Where  $p'$  is a sample from a uniform distribution over the interval  $[1-p, 1]$ :  $p' \sim U(1-p, 1)$ . This clearly shows that  $u_1'$  is a sample from the right tail of the standard normal distribution function. From equation (2.87) we know:

$$\beta_1 = \Phi^{-1}(1-p) \quad (2.94)$$

Which shows that  $\beta_1$  is the lowest possible outcome of  $u_1'$ . This proves that  $u_1'$  is a sample from the probability density as shown in the lower panel of Figure 2.27. Equation (2.92) can be used in combination with equation (2.91) to express the new function  $Z_2'$ :

$$Z_2' = \beta_2 - \rho \cdot \Phi^{-1}(1-p \cdot \Phi(v)) - u_2^* \sqrt{1-\rho^2} \quad (2.95)$$

Thus the probability of failure of a parallel system (see equation (2.67)) reduces to:

$$P(F) = P(Z_1 < 0 \cap Z_2 < 0) = P(Z_1 < 0) \cdot P(Z_2 < 0 | Z_1 < 0) = P(Z_1 < 0) \cdot P(Z_2' < 0) \quad (2.96)$$

Where  $Z_2'$  is given in equation (2.95). This equation is a multiplication of failure probabilities of two single components  $Z_1$  and  $Z_2'$ . These probabilities can both be determined with the techniques as described in section 2.3. Note that  $Z_2'$  consists of two independent standard normally distributed random variables  $v$  and  $u_2^*$ . The probability,  $p(Z_2' < 0)$  is therefore relatively easy to compute and can be evaluated with all techniques of section 2.3 with little computation time.

So the essence of the Hohenbichler method is to write a conditional probability of failure as a product of failure probabilities of two single components. The method is designed to compute failure probabilities of parallel systems:  $P(F) = P(Z_1 < 0 \cap Z_2 < 0)$ . However, through application of equation (2.69) it can also be used to compute failure probabilities of series systems:  $P(F) = P(Z_1 < 0 \cup Z_2 < 0)$ .

The only approximation that is required to apply the Hohenbichler method is the assumption that  $Z_1$  and  $Z_2$  are both linear functions. Naturally this introduces some error if  $Z_1$  and  $Z_2$  are non-linear. Furthermore, errors will be introduced in computing the individual failure probabilities of  $Z_1$  and  $Z_2'$ . However these errors can be made as small as desired by using for instance a Monte Carlo approach with sufficient number of samples (see section 2.3.3 for reference).

Another option is to use numerical integration to compute the probability of failure of  $Z_2$ , given failure of  $Z_1$ . The required formulas for this approach are described below. First, consider the following general formulation for conditional failure probability:

$$P[Z_2 < 0 | Z_1 < 0] = \frac{P[Z_2 < 0 \cap Z_1 < 0]}{P[Z_1 < 0]} \quad (2.97)$$

From equation (2.75) it can be seen that  $Z_1 < 0$  if and only if  $u_1 > \beta_1$ . This means:

$$P[Z_2 < 0 | Z_1 < 0] = \frac{P[Z_2 < 0 \cap u_1 > \beta_1]}{P[u_1 > \beta_1]} = \frac{P[Z_2 < 0 \cap u_1 > \beta_1]}{\Phi(-\beta_1)} \quad (2.98)$$

Substitution of equation (2.82) in equation (2.98) gives:

$$P[Z_2 < 0 | Z_1 < 0] = \frac{P\left[\beta_1 - \rho u_1 - \sqrt{1 - \rho^2} u_2^* < 0 \cap u_1 > \beta_1\right]}{\Phi(-\beta_1)} \quad (2.99)$$

The numerator can be computed by integration over all potential realisations of  $U_1 > \beta_1$ :

$$P[Z_2 < 0 | Z_1 < 0] = \frac{\int_{\beta_1}^{\infty} P\left[\beta_1 - \rho u_1 - \sqrt{1 - \rho^2} u_2^* < 0\right] \phi(u_1) du_1}{\Phi(-\beta_1)} \quad (2.100)$$

The probability in the numerator can be rewritten as follows:

$$\begin{aligned} P\left[\beta_1 - \rho u_1 - \sqrt{1 - \rho^2} u_2^* < 0\right] &= P\left[u_2^* > \frac{\beta_1 - \rho u_1}{\sqrt{1 - \rho^2}}\right] \\ &= P\left[u_2^* < -\frac{\beta_1 - \rho u_1}{\sqrt{1 - \rho^2}}\right] = \Phi\left(-\frac{\beta_1 - \rho u_1}{\sqrt{1 - \rho^2}}\right) \end{aligned} \quad (2.101)$$

Substitution of equation (2.101) in equation (2.100) gives:

$$P[Z_2 < 0 | Z_1 < 0] = \frac{\int_{\beta_1}^{\infty} \Phi\left(-\frac{\beta_1 - \rho u_1}{\sqrt{1 - \rho^2}}\right) \phi(u_1) du_1}{\Phi(-\beta_1)} \quad (2.102)$$

Note that this can also be written as:

$$P[Z_2 < 0 | Z_1 < 0] = \frac{\int_{-\infty}^{-\beta_1} \Phi\left(-\frac{\beta_1 + \rho u_1}{\sqrt{1 - \rho^2}}\right) \phi(u_1) du_1}{\Phi(-\beta_1)} \quad (2.103)$$

Both equations (2.102) and (2.103) can be solved by numerical integration.

### 2.4.2.3 Derivation of equivalent $\alpha$ -values

The previous section describes the Hohenbichler method for combining the probability of failure of two single components. The result is the combined probability of failure. The goal of a systems approach to failure probability is to combine the failure probabilities of all the contributing components to determine the failure probability of the whole system. This combining of probabilities takes place in a sequential fashion. That is, first two components are combined into one component, and this new component is then combined with an additional component, and so on, until only one component (the entire system) remains.

The combining of failure probabilities over components relies on  $\alpha$ -values, see equation (2.74). This means we require  $\alpha$ -values for  $Z_1 \cup Z_2$ , in order to be able to quantify the correlation with a third component,  $Z_3$ . The required  $\alpha$ -values are referred to as equivalent  $\alpha$ -values. Basically, we require a Z-function description,  $Z^e$ , that represents the combined components  $Z_1$  and  $Z_2$ . This Z-function needs to have the same probability of failure as  $Z_1$  and  $Z_2$ :

$$P(Z^e < 0) = P(Z_1 < 0 \cup Z_2 < 0) \quad (2.104)$$

The equivalent  $\alpha$ -values should be such that they describe this Z-function in the standard linearised way:

$$Z^e = \beta^e + \alpha_1^e u_1 + \dots + \alpha_n^e u_n. \quad (2.105)$$

The superscript "e" in this equation refers to the fact that these are equivalent values and functions. The equivalent value  $\beta^e$  is the reliability index that is derived with the Hohenbichler method as described in section 2.4.2.2. In order to derive the  $\alpha$ -values of function  $Z^e$ , recall from section 2.2.5 that the  $\alpha$ -values of a Z-function are related to the reliability index  $\beta$  as follows:

$$\frac{\partial \beta^e}{\partial \bar{u}_i} = \frac{\partial \beta^e}{\partial e_i} = a_i^e \quad (2.106)$$

Where  $\varepsilon_i$  represents a small change in  $\bar{u}_i$  (=the mean value of variable  $u_i$ ). Recall from equation (2.15) that:

$$\beta^e(e) = F^{-1} \left( 1 - P(Z^e(e) < 0) \right) = F^{-1} \left( 1 - P(Z_1(e) < 0 \cup Z_2(e) < 0) \right) \quad (2.107)$$

Note that  $\beta^e$  and all Z-functions are written as a function of  $\varepsilon$  to show that they depend on the choice of  $\varepsilon$ . The equivalent  $\alpha$ -value,  $\alpha_i^e$ , can be derived from:

$$a_i^e = \frac{\partial \beta^e}{\partial e_i} = \frac{\partial}{\partial e_i} \left( F^{-1} \left( 1 - P(Z_1(e_i) < 0 \cup Z_2(e_i) < 0) \right) \right) \quad (2.108)$$

Generally, Z-functions are too complex to derive equation (2.108) analytically. Therefore a numerical approach is required in which the mean of variable  $u_i$  is perturbed by a small value  $\varepsilon_i$  and subsequently the change in the value of  $\beta^e$  is quantified. Consider the linear Z-function of the 2 components:

$$Z_k = \beta_k + \alpha_{k1} u_1 + \dots + \alpha_{ki} u_i + \dots + \alpha_{kn} u_n \quad ; k = 1, 2 \quad (2.109)$$

Now we increase the mean of variable  $u_i$  with a small value of  $\varepsilon_i$ . For this purpose, define the random variable  $u_i'$  as follows:

$$u_i' = u_i + \varepsilon_i \quad (2.110)$$

Since  $u_i$  is standard normally distributed,  $u_i'$  is normally distributed with mean  $\varepsilon_i$  and standard deviation 1. Now  $u_i$  in equation (2.109) is replaced by  $u_i'$ , and consequently two new Z-functions  $Z_k(\varepsilon_i)$ ,  $k=1..2$  are obtained:

$$\begin{aligned} Z_k(\varepsilon_i) &= \beta_k + \alpha_{k1}u_1 + \dots + \alpha_{ki}u_i' + \dots + \alpha_{kn}u_n && ; k = 1, 2 \\ &= \beta_k + \alpha_{k1}u_1 + \dots + \alpha_{ki}(u_i + \varepsilon_i) + \dots + \alpha_{kn}u_n && ; k = 1, 2 \\ &= (\beta_k + \alpha_{ki}\varepsilon_i) + \alpha_{k1}u_1 + \dots + \alpha_{ki}u_i + \dots + \alpha_{kn}u_n && ; k = 1, 2 \end{aligned} \quad (2.111)$$

The change in  $\beta^e$  as a result of the change in the mean of  $u_i$  is quantified by substituting  $Z_1$  and  $Z_2$  by  $Z_1(\varepsilon_i)$  and  $Z_2(\varepsilon_i)$  in equation (2.107). From equations (2.109) and (2.111) it follows:

$$Z_k(\varepsilon_i) = Z_k + \alpha_{ki}\varepsilon_i \quad ; k = 1, 2 \quad (2.112)$$

and therefore:

$$P(Z_k(\varepsilon_i) < 0) = P(Z_k < -\alpha_{ki}\varepsilon_i) \quad ; k = 1, 2 \quad (2.113)$$

Substituting  $Z_1$  and  $Z_2$  by  $Z_1(\varepsilon_i)$  and  $Z_2(\varepsilon_i)$  in equation (2.107) therefore results in:

$$b^e(\varepsilon_i) = F^{-1}(1 - P(Z_1 < -a_i\varepsilon_i \text{ \& } Z_2 < -a_i\varepsilon_i)) \quad (2.114)$$

Where  $\beta^e(\varepsilon_i)$  is the resulting equivalent reliability index as a function of a small perturbation  $\varepsilon_i$  in the mean value of  $u_i$ . So, to derive the change in the value of  $\beta^e$  as a result of a change in the mean of variable  $u_i$  we need to quantify the following probability:

$$P(Z_1 < -a_i\varepsilon_i \text{ \& } Z_2 < -a_i\varepsilon_i) \quad (2.115)$$

This probability can be quantified again by the Hohenbichler method (or alternative methods for combining two components). The equivalent  $\alpha$ -value of variable  $u_i$  is subsequently estimated from:

$$a_i^e = \frac{b^e(\varepsilon_i) - b^e}{\varepsilon_i} \quad (2.116)$$

This procedure needs to be repeated  $n$  times, to derive the  $n$  equivalent  $\alpha$ -values. The result of this procedure is an equivalent Z-function, as described in equation (2.105), that represents the two Z-functions  $Z_1$  and  $Z_2$ .

Figure 2.28 shows an example of two Z-functions,  $Z_1$  and  $Z_2$ , and the equivalent Z-function,  $Z^e$ . The variables of the Z-functions are displayed in Table 2.4. The failure probability  $P(Z_1 < 0 \cup Z_2 < 0)$  as derived with the Hohenbichler method in this example is equal to  $6.120 \cdot 10^{-3}$ , whereas the failure probability based on Monte Carlo sampling with  $10^8$  samples (i.e. a very accurate method) is equal to  $6.121 \cdot 10^{-3}$ . Note that in this example the Hohenbichler method was also evaluated with a very accurate procedure (Monte Carlo with  $10^5$  samples) in order to

compute  $P(Z' < 0)$ , with  $Z'$  according to equation (2.95). As a consequence, the error in the Hohenbichler method in this case is less than 0.02%. This error can be made even smaller by taking more Monte Carlo samples.

This shows that the Hohenbichler method in principle is an exact method for combining 2 linear  $Z$ -functions (linear as a function of the  $u$ -variables). However, errors will be introduced [a] if the real  $Z$ -functions are non-linear or [b] when combining more than 2 components (because then the equivalent  $\alpha$ -values get involved) or [c] when FORM is used instead of Monte Carlo to compute  $P(Z' < 0)$ , with  $Z'$  according to equation (2.96).

Table 2.4 Description of the  $Z$ -functions of Figure 2.28.

Function	Before normalization			After normalisation		
	$\alpha_1$	$\alpha_2$	$\beta$	$\alpha_1$	$\alpha_2$	$\beta$
$Z_1$	-2	-1	6	-0.89	-0.45	2.68
$Z_2$	-1	-2	6	-0.45	-0.89	2.68

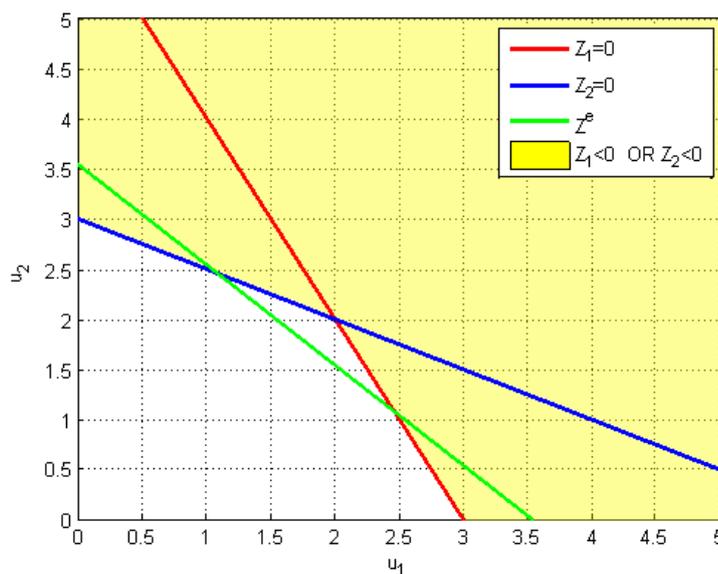


Figure 2.28 Replacement of " $Z_1 < 0 \cup Z_2 < 0$ " with " $Z^e < 0$ ". Failure domains of the three  $Z$ -functions are on the upper right side of the respective lines.

#### 2.4.2.4 Hohenbichler method for the general case of partial correlation

In the previous sections,  $Z_1$  and  $Z_2$  were functions of the same variables  $U_1, \dots, U_n$ . In the current section the more general case is considered:

$$Z_1 = \beta_1 + \alpha_{11}u_{11} + \dots + \alpha_{1n}u_{1n} \tag{2.117}$$

$$Z_2 = \beta_2 + \alpha_{21}u_{21} + \dots + \alpha_{2n}u_{2n}$$

Where  $u_{ik}$  refers to the  $k^{\text{th}}$  random variable of the  $i^{\text{th}}$   $Z$ -function. In this case, variables  $U_{1k}$  and  $U_{2k}$  are partially correlated. All other correlations are equal to zero:

$$\rho(u_{1j}, u_{2k}) = \begin{cases} \rho_{12k} & ; j = k \\ 0 & ; j \neq k \end{cases} \quad (2.118)$$

For components 1 and 2, the  $k^{\text{th}}$  random variable (i.e.  $u_{1k}$  and  $u_{2k}$ ) in principle refer to the same load or strength variable, but the sampled values can be different because they refer to different components. For instance, the  $k^{\text{th}}$  variable may refer to the thickness of a clay layer. This thickness will be different for different dike segments. The correlation between  $Z_1$  and  $Z_2$  now becomes:

$$\rho(Z_1, Z_2) = \sum_{k=1}^n \alpha_{1k} \alpha_{2k} \rho_{12k} \quad (2.119)$$

The system as described with equations (2.117) and (2.118) represents for instance the case where two neighbouring dike segments are combined. In that case the variables  $u_{1k}$  and  $u_{2k}$  both describe the same load or strength variable, but the realizations of the two samples are not necessarily the same. Note that equation (2.117) also covers the case in which the two  $Z$ -functions depend on different sets of random variables, for instance because two dike segments are combined that are situated along different water systems, or if one component is a dike segment and the other one is a dune segment. In that case, the combined sets of random variables will be used in equation (2.117) and some of the  $\alpha$ -values will be equal to zero.

In order to determine the probability of failure  $P(Z_1 < 0 \cup Z_2 < 0)$ , the exact same method as described in section 2.4.2.2 is used. So, the combined set of equations (2.96) and (2.95) are solved, using the correlation coefficient of equation (2.119) as input for equation (2.95). In order to derive the equivalent  $\alpha$ -values, the procedure becomes somewhat more complicated than the procedure with full correlation that was described in section 2.4.2.3. The approach is to describe the  $U$ -variables of  $Z_2$  as a function of the  $u$ -variables of  $Z_1$ :

$$u_{2k} = u_{1k} \rho_{12k} + u_{2k}^* \sqrt{(1 - \rho_{12k}^2)} \quad ; k = 1..n \quad (2.120)$$

The first term in this function represents the part of  $u_{2k}$  that is fully correlated to  $u_{1k}$ , the second term describes the part that is fully uncorrelated to  $u_{1k}$ . Variable  $U_{2k}^*$  is standard normally distributed and independent of variable  $U_{1k}$ . To verify the applicability of equation (2.120) it needs to be shown that [1]  $U_{2k}$  is standard normally distributed and [2] that  $U_{1k}$  and  $U_{2k}$  have a mutual correlation coefficient that is equal to  $\rho_{12k}$ . Note that this was proven earlier in the follow-up of equation (2.77). Inserting the expression for  $u_{2k}$ , given by equation (2.120), into the formula for  $Z_2$  (equation (2.117)) gives:

$$Z_2 = \beta_2 + \alpha_{21} \left( u_{11} \rho_{121} + u_{21}^* \sqrt{(1 - \rho_{121}^2)} \right) + \dots + \alpha_{2n} \left( u_{1n} \rho_{12n} + u_{2n}^* \sqrt{(1 - \rho_{12n}^2)} \right) \quad (2.121)$$

which can be written more compactly as follows:

$$Z_2 = \beta_2 + \sum_{k=1}^n \alpha_{2k} \left( u_{1k} \rho_{12k} + u_{2k}^* \sqrt{(1 - \rho_{12k}^2)} \right) \quad (2.122)$$

Note that the expression for  $Z_1$  in equation (2.117) can also be written more compactly as follows:

$$Z_1 = \beta_1 + \sum_{k=1}^n \alpha_{1k} u_{1k} \quad (2.123)$$

The procedure for determining the equivalent  $\alpha$ -values is similar to that introduced for the case of full correlation (section 2.4.2.3). So again, an equivalent Z-function of the following form is derived:

$$Z^e = \beta^e + \alpha_1^e u_1 + \dots + \alpha_n^e u_n = \beta^e + \sum_{k=1}^n \alpha_k^e u_k \quad (2.124)$$

The major difference with section 2.4.2.3 is that the random variable  $u_k$  in equation (2.124) will represent the two variables  $u_{1k}$  and  $u_{2k}$  of the functions  $Z_1$  and  $Z_2$ . So, essentially an equivalent  $\alpha$ -value is computed for the two random variables  $u_{1k}$  and  $u_{2k}$  and these are subsequently combined into a single equivalent  $\alpha$ -value. Actually, for practical reasons, equivalent  $\alpha$ -values are computed for  $u_{1k}$  and  $u_{2k}^*$  because they are mutually independent.

The first step is to derive the partial derivative of  $\beta^e$  to  $u_{1k}$ . Similar to section 2.4.2.3 this is done by quantifying the change in  $\beta^e$  as a result of a small perturbation in the value of  $u_{1k}$ .  $\beta^e$  is related to the two Z-functions as follows:

$$\beta^e = \Phi^{-1} [1 - P\{Z_1 < 0 \cup Z_2 < 0\}] \quad (2.125)$$

Substituting equations (2.122) and (2.123) into equation (2.125) gives:

$$\beta^e = \Phi^{-1} \left[ 1 - P \left\{ \beta_1 + \sum_{k=1}^n \alpha_{1k} u_{1k} < 0 \cup \beta_2 + \sum_{k=1}^n \alpha_{2k} \left( u_{1k} \rho_{12k} + u_{2k}^* \sqrt{1 - \rho_{12k}^2} \right) < 0 \right\} \right] \quad (2.126)$$

A small perturbation,  $\varepsilon_k^I$ , on the mean value of  $u_{1k}$  will have the following effect on  $\beta^e$ :

$$\beta^e(\varepsilon_k^I) = \Phi^{-1} [1 - P\{Z_1 < -\alpha_{1k} \varepsilon_k^I \cup Z_2 < -\alpha_{2k} \varepsilon_k^I \rho_{12k}\}], \quad (2.127)$$

This value can be computed once again with the Hohenbichler method (or alternative methods for combining two components). The equivalent  $\alpha$ -value for variable  $u_{1k}$  can then be obtained from:

$$\alpha_k^I = \frac{\beta^e(\varepsilon_k^I) - \beta^e}{\varepsilon_k^I} \quad (2.128)$$

Subsequently a similar sensitivity analysis is done for  $u_{2k}^*$ . A small perturbation,  $\varepsilon_k^{II}$ , in the mean value of  $u_{2k}^*$  has the following effect on  $\beta^e$  (see equations(2.125) and (2.126)):

$$\beta^e(\varepsilon_k^{II}) = \Phi^{-1} [1 - P\{Z_1 < 0 \cup Z_2 < -\alpha_{2k} \varepsilon_k^{II} \sqrt{1 - \rho_{12k}^2}\}] \quad (2.129)$$

This value can be computed once again with the Hohenbichler method (or alternative methods for combining two components). The equivalent  $\alpha$ -value for variable  $u_{2k}^*$  can then be obtained from:

$$\alpha_k^{II} = \frac{\beta^e (\varepsilon_k^{II}) - \beta^e}{\varepsilon_k^{II}} \quad (2.130)$$

The two derived  $\alpha$ -values can then be combined as follows:

$$\alpha_k^e = \sqrt{(\alpha_k^I)^2 + (\alpha_k^{II})^2} \quad (2.131)$$

This is the required equivalent  $\alpha$ -value for the  $k^{\text{th}}$  random variable in the combined Z-function  $Z^e$ . Equation (2.131) can be explained as follows. The Z-functions of the two components,  $Z_1$  and  $Z_2$  are a function of mutually independent standard normally distributed variables  $u_{11}, u_{21}^*, \dots, u_{1n}, u_{2n}^*$  (see equations (2.117) and (2.122)). For each of these variables an equivalent  $\alpha$ -value was derived:  $\alpha_1^I, \alpha_1^{II}, \dots, \alpha_n^I, \alpha_n^{II}$ . This means the combined Z-function of the two components can be written as follows:

$$\begin{aligned} Z^e &= \beta^e + \alpha_1^I u_{11} + \alpha_1^{II} u_{21}^* + \dots + \alpha_n^I u_{1n} + \alpha_n^{II} u_{2n}^* \\ &= \beta^e + \sum_{k=1}^n (\alpha_k^I u_{1k} + \alpha_k^{II} u_{2k}^*) \end{aligned} \quad (2.132)$$

If we compare this equation with equation (2.124) it is clear that the “new” random variable  $u_k$  replaces the pair of random variables  $u_{1k}$  and  $u_{2k}^*$  and also that the equivalent  $\alpha$ -value,  $\alpha_k^e$ , replaces  $\alpha_k^I$  and  $\alpha_k^{II}$ . This can only be done if the standard deviation of  $\alpha_k^e u_k$  is equal to the standard deviation of:  $\alpha_k^I u_{1k} + \alpha_k^{II} u_{2k}^*$ . This is the case if we chose  $\alpha_k^e$  according to equation (2.131).

Note 1: application of equation (2.131) will result in a value of  $\alpha_k^e$  that is non-negative. For load variables this is incorrect, as they have negative  $\alpha$ -values. Therefore, for load variables,  $\alpha_k^e$  should be taken equal to:

$$\alpha_k^e = -\sqrt{(\alpha_k^I)^2 + (\alpha_k^{II})^2} \quad (2.133)$$

If it is not known whether the  $k^{\text{th}}$  variable is a load variable, this information can be obtained by reading the sign of  $\alpha_{1k}$  and  $\alpha_{2k}$ .

Note 2: since the equivalent  $\alpha$ -values are derived numerically, the sum of the squares of the equivalent  $\alpha$ -values may differ from 1. In that case an additional normalization step is required:

$$\alpha_{k;final}^e = \frac{\alpha_k^e}{\sqrt{\sum_{j=1}^n (\alpha_j^e)^2}} \quad ; k = 1..n \quad (2.134)$$

#### 2.4.2.5 System with arbitrary number of components

We have just considered the case of probability of failure of a parallel system of two components. In this section we extend the concept to an arbitrary number of components. Suppose we have an arbitrary system of  $n$  components. The failure probability for the system is given by:

$$P(F) = P(Z_1 < 0 \cup Z_2 < 0 \cup \dots \cup Z_n < 0) \quad (2.135)$$

An example is the computation of the failure probability due to the mechanism overtopping over  $n$  defense segments. The function  $Z_i$  is the limit state function for component  $i$  and the occurrence  $Z_i < 0$  indicates failure of component  $i$ .

The procedure of combining is to first combine two components, so that the problem with  $n$  components reduces to a problem with  $n-1$  components. The next step combines two components again so that the problem reduces to one with  $n-2$  components, and continues in this fashion until only one component remains, where this last component represents the entire system.

The order of the combination is important. The determination of equivalent  $\alpha$ -values, discussed in the previous section, is an approximating method, which makes the entire combination procedure an approximating method. The accuracy of the resulting failure probability is influenced by the sequence in which the components are combined. The most accurate results are obtained by combining the most correlated components first. This is clarified by the example below with three Z-functions in Table 2.5 and Figure 2.29.

Table 2.5 Description of the Z-functions of Figure 2.29

Function	Variable		
	$\alpha_1$	$\alpha_2$	$\beta$
$Z_1$	-1	0	2
$Z_2$	0	-1	2
$Z_3$	-1	0	2.5

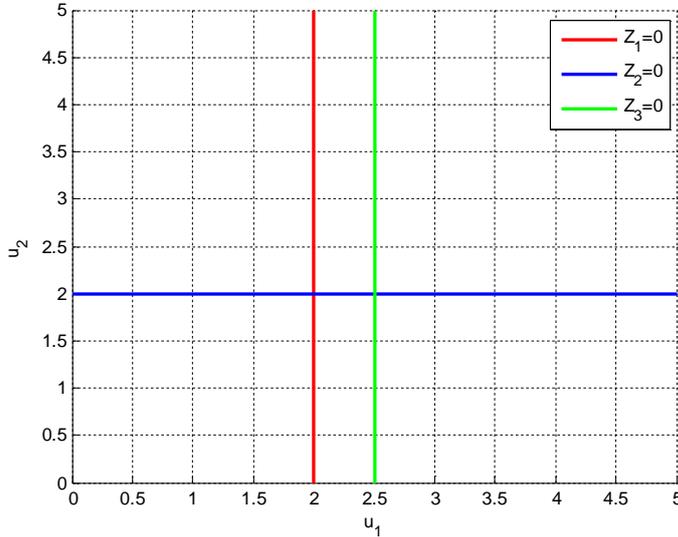


Figure 2.29 Functions  $Z_1$ ,  $Z_2$  and  $Z_3$  of the current example

In this example, functions  $Z_1$  and  $Z_3$  are mutually fully correlated, whereas they are fully uncorrelated with function  $Z_2$ . We will demonstrate that the best strategy is to first combine the two correlated Z-functions ( $Z_1$  and  $Z_3$ ). First of all the exact solution for this relatively easy example is derived. It is clear from Figure 2.29 that if  $Z_3 < 0 \Rightarrow Z_1 < 0$ . This means:

$$P(Z_1 < 0 \cup Z_3 < 0) = P(Z_1 < 0) \quad (2.136)$$

And therefore:

$$P(F) = P(Z_1 < 0 \cup Z_2 < 0 \cup Z_3 < 0) = P(Z_1 < 0 \cup Z_2 < 0) \quad (2.137)$$

Since  $Z_1$  and  $Z_2$  are independent, the failure probability is equal to:

$$P(Z_1 < 0 \cup Z_2 < 0) = 1 - P(Z_1 \geq 0)P(Z_2 \geq 0) = 1 - \Phi(\beta_1)\Phi(\beta_2) = 1 - \Phi(2)^2 \approx 0.045 \quad (2.138)$$

If we combine  $Z_1$  and  $Z_3$  first with the Hohenbichler method, the exact same result is obtained. If we combine  $Z_1$  and  $Z_2$  first, the estimated failure probability is equal to 0.0482, whereas if we combine  $Z_2$  and  $Z_3$  first, the probability of failure is equal to 0.0492. This demonstrates that in this example indeed the best strategy is to first combine the two components with the largest mutual correlation. To understand why this is the case, Figure 2.30 shows the equivalent function,  $Z^e$  of  $Z_1 \cup Z_3$ . This function turns out to be exactly the same as  $Z_1$ . Recall from equation (2.136) that this means that  $Z^e$  is an exact representation of  $Z_1 \cup Z_3$ .

Figure 2.31 shows the equivalent function,  $Z^e$  of  $Z_1 \cup Z_2$ . Clearly, this function is a compromise between  $Z_1 < 0 \cup Z_2 < 0$  and it is clear why this introduces some errors after combining with function  $Z_3 < 0$  (for instance because  $Z_3$  now defines part of the failure domain, whereas in the original problem statement it was redundant).

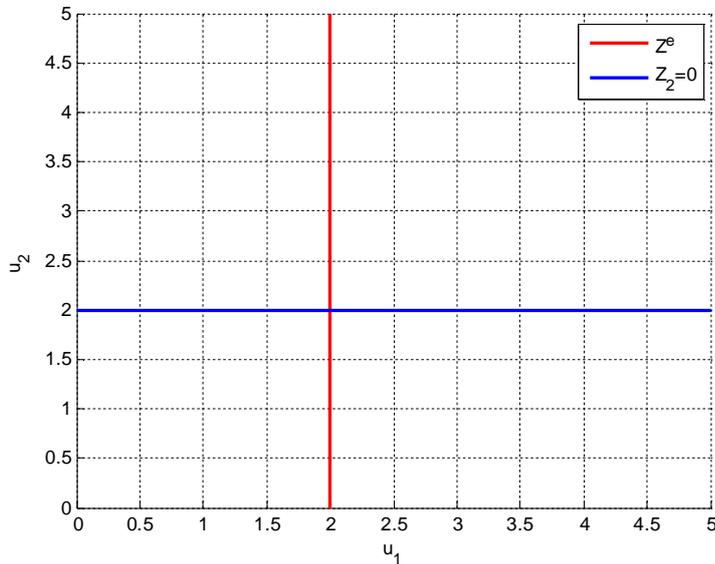


Figure 2.30 Function  $Z_2$  and the equivalent Z-function of  $Z_1 \cup Z_3$ .

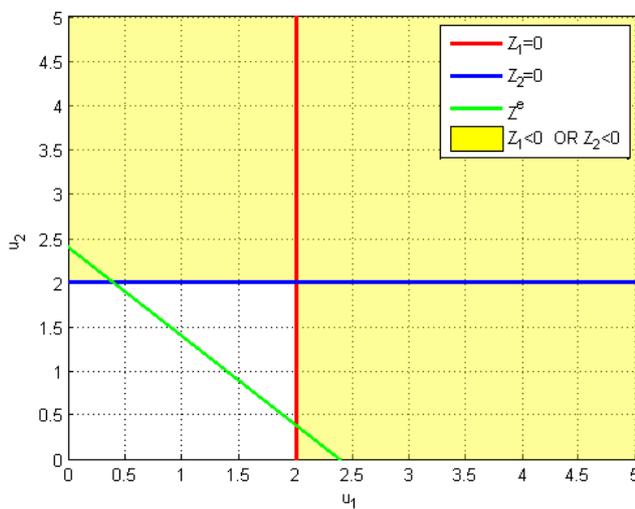


Figure 2.31 Functions  $Z_1=0$ ,  $Z_2=0$  and the equivalent Z-function of  $Z_1 < 0 \cup Z_2 < 0$ .

Figure 2.32 helps illustrate the concept of combining components with the largest mutual correlation, with an example. Shown are four components with reliability functions  $Z_1$ ,  $Z_2$ ,  $Z_3$ , and  $Z_4$ . Let functions  $Z_1$  and  $Z_2$  be the most strongly correlated. These two components are then first combined and replaced by the equivalent reliability function  $Z_2^e$ . For the three remaining components, the correlations between them are again computed. Consider the case where  $Z_3$  and  $Z_4$  are now the most correlated; the following step will be the combination of  $Z_3$  and  $Z_4$ , resulting in the equivalent reliability function  $Z_4^e$ . The final step is the combination of  $Z_2^e$  and  $Z_4^e$ .

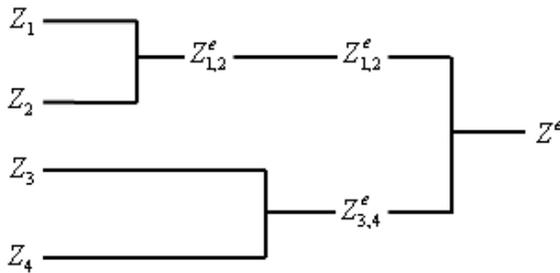


Figure 2.32 Example of combining failure probabilities over four components

### 2.4.3 Upscaling for systems with identical components: numerical integration with constant correlation

#### 2.4.3.1 Computation of failure probability

Upscaling refers to combining failure probabilities over “identical components”. Upscaling is distinguished from the more generic Hohenbichler combination techniques because the components being identical allows for some convenient simplifications. Identical in this case refers to the fact that the components have the same failure probability (i.e. the same reliability index  $\beta$ ) and they are mutually correlated with the same correlation coefficient  $\rho$ :

$$\begin{aligned} \beta(Z_i) &= \beta \quad ; i = 1..n_e \\ \rho(Z_i, Z_j) &= \rho \quad ; i \neq j \end{aligned} \tag{2.139}$$

Where  $n_e$  is the number of components,  $\rho$  is the correlation coefficient and  $Z_i$  is the Z-function of component  $i$ . Note that in general, the components also have in common the underlying set of random variables and the associated  $\alpha$ -values, but that is not a necessary condition for applying the method as described in this section.

Examples of when upscaling may be applied are the combining of failure probabilities at one time scale to a larger time scale and upscaling failure probabilities from a cross section of a defense segment to the longitudinal extent of the segment. Such applications are described in section 2.5. The value of  $\rho$  first needs to be determined based on knowledge of the system. This will be discussed in section 2.5 where applications of the method in Hydra-Ring are discussed. For now,  $\rho$  is assumed to be known.

The upscaling method makes use of linearized approximations of the Z-functions, as described in section 2.2.5. The estimated probability of failure of the system will therefore be an approximation of the actual probability of failure. Errors made in the approximation will depend on the system under consideration. In section 2.2.5 it was shown that linearised Z-functions can be described as follows:

$$Z_i = \beta - \alpha_1 u_{i1} - \dots - \alpha_n u_{in} \quad ; i = 1..n_e \tag{2.140}$$

Furthermore, it was shown that the sum of the product of  $\alpha$ -values and standard normal  $u$ -values can be replaced by a single standard normal  $u$ -value:

$$Z_i = \beta - u_i \quad ; i = 1 \dots n_e \quad (2.141)$$

Where  $u_i$  is a standard normally distributed variable and  $\beta$  is the reliability index. The value of  $\beta$  is considered to be known, i.e. it is determined by the probabilistic computation techniques as described in section 2.3. This means  $\beta$  is a constant in equation (2.141) and the mutual correlation of the Z-functions is therefore entirely determined by the mutual correlation of the U-variables:

$$\rho(Z_i, Z_j) = \rho(U_i, U_j) = \rho \quad ; i \neq j \quad (2.142)$$

To describe a system that satisfies the relation of equation (2.142), variable  $u_i$  is written as a function of two independent standard normal random variables  $u_i^*$  and  $v$ :

$$u_i = u_i^* \sqrt{1-\rho} - v \sqrt{\rho} \quad (2.143)$$

The variables  $U_i^*$ ,  $i=1..n$  are taken to be mutually independent:

$$\rho(u_i^*, u_j^*) = 0 \quad ; i \neq j \quad (2.144)$$

Note 1: there is difference between equation (2.143) and equation (2.77), where  $\rho$  is used instead of  $\sqrt{\rho}$ , even though in both cases the correlation between  $u_1$  and  $u_2$  is equal to  $\rho$ . The reason for this difference is the fact that in equation (2.77), variable  $u_2$  is written as a function of  $u_1$ , whereas in this case,  $u_1$  and  $u_2$  are both written as a function of a separate variable  $v$ .

Note 2: in equation (2.77) it is assumed that  $\rho > 0$ . For  $\rho \leq 0$  the Hohenbichler method together with the outcrossing approach should be used.

To verify the applicability of equation (2.143) it needs to be shown that [1]  $u_i$  is standard normally distributed and [2] that the relation of equation (2.142) holds. To prove [1], we apply the following general rule (see, e.g. Grimmer and Sirzaker, 1982): If X and Y are normally distributed random variables, then  $aX+bY$  is also normally distributed with a mean,  $\mu$ , and standard deviation,  $\sigma$ , equal to:

$$\begin{aligned} \mu &= a\mu_x + b\mu_y \\ \sigma &= \sqrt{a^2\sigma_x^2 + b^2\sigma_y^2} \end{aligned} \quad (2.145)$$

Application of this rule on equation (2.143), where  $u_i^*$  and  $v$  are both normally distributed with mean 0 and standard deviation 1, gives:

$$\begin{aligned} \mu &= \sqrt{1-\rho} \cdot 0 - \sqrt{\rho} \cdot 0 = 0 \\ \sigma &= \sqrt{(1-\rho) \cdot 1 + \rho \cdot 1} = 1 \end{aligned} \quad (2.146)$$

Which proves that  $u_i$  is standard normally distributed. To prove [2], i.e. that the relation of equation (2.142) holds, equations (2.141) and (2.143) are combined:

$$Z_i = \beta - u_i^* \sqrt{1-\rho} - v\sqrt{\rho} \quad ; i = 1 \dots n_e \quad (2.147)$$

The variable  $v$  in equation (2.147) is part of each Z-function. This creates the desired mutual correlation between the functions  $Z_i$ ,  $i=1 \dots n_e$ . To prove this, it needs to be shown that variables  $u_i$  and  $u_j$ ,  $i \neq j$ , have a mutual correlation equal to  $\rho$ . The correlation between  $u_i$  and  $u_j$  is derived as follows:

$$\rho(u_i, u_j) = \frac{\text{cov}(u_i, u_j)}{[\sigma(u_i)\sigma(u_j)]} = \frac{\text{cov}(u_i, u_j)}{[1 \cdot 1]} = \text{cov}(u_i, u_j) \quad (2.148)$$

The covariance of  $u_i$  and  $u_j$  is equal to:

$$\begin{aligned} \text{cov}(u_i, u_j) &= E[u_i u_j - \mu(u_i)\mu(u_j)] \\ &= E[u_i u_j] = E\left[\left(u_i^* \sqrt{1-\rho} - v\sqrt{\rho}\right)\left(u_j^* \sqrt{1-\rho} - v\sqrt{\rho}\right)\right] \\ &= E\left[u_i^* u_j^* (1-\rho) - u_i^* v\sqrt{\rho(1-\rho)} - u_j^* v\sqrt{\rho(1-\rho)} + \rho v^2\right] \\ &= E[0 - 0 - 0 + \rho v^2] = \rho E[v^2] = \rho \end{aligned} \quad (2.149)$$

This proves that equation (2.147) describes a system of  $n_e$  components that are mutually correlated with a correlation coefficient  $\rho$ . The theorem of total probability is used to derive the probability of failure of this system:

$$P(F) = P(Z_1 < 0 \cup \dots \cup Z_n < 0) = \int_v P(Z_1 < 0 \cup \dots \cup Z_n < 0 | v) f(v) dv \quad (2.150)$$

Where  $f(v)$  is the density function of the standard normal distribution. The probability that at least 1 component fails is equal to 1 minus the probability that none of the components fails. Equation (2.150) can therefore be rewritten as:

$$\begin{aligned} P(F) &= \int_v [1 - P(Z_1 \geq 0 \cap \dots \cap Z_n \geq 0 | v)] f(v) dv \\ &= \int_v [1 - P\{(Z_1 \geq 0 | v) \cap \dots \cap (Z_n \geq 0 | v)\}] f(v) dv \end{aligned} \quad (2.151)$$

For a given value of  $v$ , the individual failure probabilities of the Z-functions are mutually independent:

$$P[(Z_i < 0 | v) \cap (Z_j < 0 | v)] = P(Z_i < 0 | v) P(Z_j < 0 | v) \quad ; i \neq j \quad (2.152)$$

This can be easily verified from equation (2.143). If the value of  $v$  is given, equation (2.143) only contains one random variable:  $u_i^*$ . Since the  $u_i^*$ -values are mutually independent (see equation (2.144)), the Z-functions of equation (2.143) are mutually independent as well, which

leads to the equality in equation (2.152). Implementation of equation (2.152) in equation (2.151) gives:

$$P(F) = \int_v \left[ 1 - \prod_{i=1}^{n_e} P(Z_i \geq 0 | v) \right] f(v) dv \quad (2.153)$$

Because all components are identical, the following is true:

$$P(Z_1 \geq 0 | v) = P(Z_2 \geq 0 | v) = \dots = P(Z_n \geq 0 | v) \stackrel{def}{=} P(Z \geq 0 | v) \quad (2.154)$$

This changes equation (2.153) into:

$$P(F) = \int_v \left[ 1 - P(Z \geq 0 | v)^{n_e} \right] f(v) dv \quad (2.155)$$

The conditional probability of the Z-function in the integral is equal to:

$$P(Z \geq 0 | v) = P\left(\beta - u^* \sqrt{1-\rho} - v\sqrt{\rho} \geq 0\right) = P\left(u^* \leq \frac{\beta - v\sqrt{\rho}}{\sqrt{1-\rho}}\right) \quad (2.156)$$

Since v is a given constant, u\* is the only random variable in equation (2.156), and since u\* is standard normally distributed, the conditional probability is equal to:

$$P(Z \geq 0 | v) = \Phi\left(\frac{\beta - v\sqrt{\rho}}{\sqrt{1-\rho}}\right) = \Phi(\beta^*) \quad (2.157)$$

Where  $\Phi$  is the standard normal distribution function and  $\beta^*$  is equal to:

$$\beta^* = \frac{\beta - v\sqrt{\rho}}{\sqrt{1-\rho}} \quad (2.158)$$

Equation (2.155) then changes into:

$$P(F) = \int_v \left[ 1 - \{1 - \Phi(-\beta^*)\}^{n_e} \right] f(v) dv, \quad (2.159)$$

Equation (2.159) can be solved by numerical integration over the standard normal variable v. Since v is the only variable, the grid size of v can be chosen small without requiring significant computation time. The error from the numerical integration of equation (2.159) can therefore be as small as desired. This means the only potentially significant error that is introduced in this method is related to the linearization of the Z-function, which was necessary to derive equation (2.159).

## 2.4.3.2 Equivalent alpha-values

As with the Hohenbichler method, equivalent  $\alpha$ -values can be computed for the component that represents the combination of  $n_e$  identical components. This is necessary in case the resulting component is used in subsequent combining procedures where  $\alpha$ -values are required. A similar approach with perturbed u-values is used as in the Hohenbichler method. However, because in this special case the components are identical, this allows for some convenient simplifications that require less computation time.

Consider, again, the system of  $n_e$  identical components as described in equation (2.140):

$$Z_i = \beta - \alpha_1 u_{i1} - \dots - \alpha_n u_{in} \quad ; i = 1 \dots n_e \quad (2.160)$$

These components are combined according to the method as described in section 2.4.3.1, resulting in a failure probability and associated reliability index  $\beta^e$ . The combined component can be described by (similar to equation (2.120)):

$$Z^e = \beta^e - \alpha_1^e u_1 - \dots - \alpha_n^e u_n \quad (2.161)$$

The superscript “e” in this equation refers to the fact that these are equivalent values and functions. In order to derive the  $\alpha$ -values of function  $Z^e$ , recall from section 2.2.5 that the  $\alpha$ -values of a Z-function are related to the reliability index  $\beta^e$  as follows:

$$\frac{\partial b^e}{\partial \bar{u}_k} = a_k^e \quad ; k = 1..n \quad (2.162)$$

In which  $\bar{u}_i$  is the mean of variable  $u_i$ . Note the value  $\alpha_i^e$  represents the combined effect of variables  $u_{i1} \dots u_{in}$ , and that these variables are mutually correlated. In order to determine  $\alpha_i^e$ , these variables need to be split in an independent and mutually dependent part, similar to the description in section 2.4.3.1. Consider for this purpose equation (2.160). The different u-values within a single component are mutually uncorrelated, whereas corresponding u-values in different components can be correlated. In formula:

$$\begin{aligned} \rho(U_{ij}, U_{ik}) &= 0 \quad ; j \neq k \\ \rho(U_{ik}, U_{\ell k}) &= \rho_k \quad ; i \neq \ell \end{aligned} \quad (2.163)$$

Each u-variable can therefore be split in a correlated and uncorrelated part:

$$u_{ik} = u_{ik}^* \sqrt{1 - \rho_k} + v_k \sqrt{\rho_k} \quad (2.164)$$

In which  $u_{ik}^*$  and  $v_k$  are realizations of mutually independent standard normally distributed variables  $U_{ik}^*$  and  $V_k$ . Furthermore, the variables  $U_{ik}^*$ ,  $i=1..n_e$ ,  $k=1..n$  and  $v_k$ ,  $k=1..n$  are all taken to be mutually independent:

$$\begin{aligned}
\rho(U_{ij}^*, U_{k\ell}^*) &= 0 \quad ; i \neq j \cup k \neq \ell \\
\rho(U_{ij}^*, v_k) &= 0 \\
\rho(v_j, v_k) &= 0 \quad ; j \neq k
\end{aligned} \tag{2.165}$$

With this formulation, variables  $U_{ik}$ ,  $i=1..n_e$ ,  $k=1..n$ , automatically fulfill requirement (2.163), as can be shown in the same manner as shown below equation (2.144) in the previous section. Substituting equation (2.164) in equation (2.160) gives:

$$\begin{aligned}
Z_i &= \beta - \alpha_1 (u_{i1}^* \sqrt{1-\rho_1} + v_1 \sqrt{\rho_1}) - \dots - \alpha_n (u_{in}^* \sqrt{1-\rho_n} + v_n \sqrt{\rho_n}) \quad ; i=1..n_e \\
&= \beta - \sum_{k=1}^n \alpha_k u_{ik}^* \sqrt{1-\rho_k} - \sum_{k=1}^n \alpha_k v_k \sqrt{\rho_k} \quad ; i=1..n_e
\end{aligned} \tag{2.166}$$

This equation can be replaced by:

$$Z_i = \beta - u_i^* \sqrt{1-\rho} - v \sqrt{\rho} \quad ; i=1..n_e \tag{2.167}$$

In which:

$$\begin{aligned}
\rho &= \sum_{k=1}^n (\alpha_k)^2 \rho_k \\
u_i^* &= \frac{1}{\sqrt{1-\rho}} \sum_{k=1}^n \alpha_k u_{ik}^* \sqrt{1-\rho_k} \quad ; i=1..n_e \\
v &= \frac{1}{\sqrt{\rho}} \sum_{k=1}^n \alpha_k v_k \sqrt{\rho_k}
\end{aligned} \tag{2.168}$$

The validity of this replacement can be easily verified by substituting the formulations of  $u_i$  and  $v$  of equation (2.168) into equation (2.167). Equation (2.167) is similar to equation (2.147) *if and only if*  $U_i^*$  and  $V$  are mutually independent standard normally distributed variables. The mutual independence can easily be shown since all components  $U_{ik}^*$ ,  $i=1..n_e$ ,  $k=1..n$ , and  $v_k$ ,  $k=1..n$  are mutually independent (see (2.165)).

To verify if  $U_i^*$  and  $V$  are standard normally distributed we apply the following general rule (see, e.g. Grimmett and Sirzaker, 1982): If  $X$  and  $Y$  are normally distributed random variables, then  $aX+bY$  is also normally distributed with a mean,  $\mu$ , and standard deviation,  $\sigma$ , equal to:

$$\begin{aligned}
\mu &= a\mu_X + b\mu_Y \\
\sigma &= \sqrt{a^2\sigma_X^2 + b^2\sigma_Y^2}
\end{aligned} \tag{2.169}$$

Application of this rule on equation (2.168), where all components  $U_{ik}^*$  and  $V_k$  are normally distributed with mean 0 and standard deviation 1, gives:

$$\begin{aligned}\mu(U_i) &= \frac{1}{\sqrt{\rho}} \sum_{k=1}^n \alpha_k \cdot 0 \cdot \sqrt{1-\rho_k} = 0 \\ \sigma(U_i) &= \sqrt{\left(\frac{1}{\sqrt{1-\rho}}\right)^2 \sum_{k=1}^n (\alpha_k)^2 \cdot (1)^2 \cdot (1-\rho_k)} = \sqrt{\frac{1}{1-\rho} \left[ \sum_{k=1}^n \alpha_k^2 - \sum_{k=1}^n \alpha_k^2 \rho_k \right]} \\ &= \sqrt{\frac{1}{1-\rho} [1-\rho]} = \sqrt{1} = 1\end{aligned}\quad (2.170)$$

This shows that  $U_i^*$  and  $V$  in equation (2.167) are mutually independent standard normally distributed variables. Taking into account the formulation of the Z-function in equation (2.167), The equivalent coefficient  $\alpha_k^e$  can now be derived as follows

$$\alpha_k^e = \frac{\partial \beta^e}{\partial \bar{u}_k} = \frac{\partial \beta^e}{\partial \bar{u}} \frac{\partial \bar{u}}{\partial \bar{u}_k} + \frac{\partial \beta^e}{\partial \bar{v}} \frac{\partial \bar{v}}{\partial \bar{u}_k} \quad (2.171)$$

Where  $\bar{u}$ ,  $\bar{v}$  and  $\bar{u}_k$  are the mean values of variables  $U^*$ ,  $V$  and  $U_k$ . So, the derivation of coefficients  $\alpha_k$   $i=1..n_e$  it comes down now to determining the four partial derivatives of equation (2.171). The first two can be determined directly from equation (2.168):

$$\begin{aligned}\frac{\partial \bar{u}}{\partial \bar{u}_k} &= \frac{\alpha_k \sqrt{1-\rho_k}}{\sqrt{1-\rho}} \\ \frac{\partial \bar{v}}{\partial \bar{u}_k} &= \frac{\alpha_k \sqrt{\rho_k}}{\sqrt{\rho}}\end{aligned}\quad (2.172)$$

The partial derivative of  $\beta^e$  to  $\bar{v}$  is determined numerically:

$$a_v = \frac{\partial \beta^e}{\partial \bar{v}} \gg \frac{b^e(e_v) - b^e}{e_v} \quad (2.173)$$

In which  $\beta^e(\varepsilon_v)$  is the reliability index of the of the upscaled system of  $n_e$  components, after perturbation of  $\bar{v}$  with a small value  $\varepsilon_v$ .

$$b^e(e_v) = F^{-1}\left(1 - P\left(Z^e(e_v) < 0\right)\right) = F^{-1}\left(1 - \prod_{i=1}^{n_e} P\left(Z_i(e_v)\right)\right) \quad (2.174)$$

In which function  $Z_i(\varepsilon_v)$  is as follows:

$$\begin{aligned}Z_i(\varepsilon_v) &= \beta - u_i^* \sqrt{1-\rho} - (v + \varepsilon_v) \sqrt{\rho} \quad ; i = 1 \dots n_e \\ &= (\beta - \sqrt{\rho} \varepsilon_v) - u_i^* \sqrt{1-\rho} - v \sqrt{\rho} \quad ; i = 1 \dots n_e\end{aligned}\quad (2.175)$$

In other words:  $Z_{(\varepsilon_v)}$  is a Z-function with reliability index  $\beta - \varepsilon_v \sqrt{\rho}$ . So  $\beta^e(\varepsilon_v)$  is quantified by substituting equation (2.175) into equation (2.174) and subsequent application of the upscaling procedure of section 2.4.3.1. Subsequently,  $\beta^e(\varepsilon_v)$  is substituted in equation (2.173) in order to derive  $\alpha_v$  the partial derivative of  $\beta^e$  to  $\bar{v}$ . The next step is to derive the partial derivative of  $\beta^e$  to  $\bar{u}$ . This can be derived as follows:

$$\frac{\partial \beta^e}{\partial \bar{u}} = \sqrt{1 - \left( \frac{\partial \beta^e}{\partial \bar{v}} \right)^2} = \sqrt{1 - \alpha_v^2} \quad (2.176)$$

This can be explained as follows: the partial derivative of  $\beta^e$  to  $\bar{v}$  is the resulting  $\alpha$ -value for the *dependent* part of the  $n_e$  components, represented by variable  $V$ . The partial derivative of  $\beta^e$  to  $\bar{u}$  is the resulting  $\alpha$ -value for the *independent* part of the  $n_e$  components, represented by variable  $U^*$ . The sum of the squares of these alpha-values should be equal to 1. Substitution of equations (2.172), (2.173) and (2.176) into equation (2.171) provides the requested equivalent  $\alpha$ -values:

$$\alpha_k^e = \frac{\partial \beta^e}{\partial \bar{u}_k} = \sqrt{1 - \alpha_v^2} \frac{\alpha_k \sqrt{1 - \rho_k}}{\sqrt{1 - \rho}} + \alpha_v \frac{\alpha_k \sqrt{\rho_k}}{\sqrt{\rho}} \quad (2.177)$$

The Z-function of the resulting component from the upscaling procedure (equation (2.161)) needs to have a standard deviation equal to 1. This means the sum of the squares of the equivalent  $\alpha$ -values should be equal to 1. Equation (2.177) guarantees that this is the case if all values of  $\rho_k$  are equal to either 0 or 1, which is generally the case for upscaling in time (i.e. slow varying random load variables and strength variables have an autocorrelation equal to 1, fast varying random variables have an autocorrelation equal to 0). This can be deduced as follows:

$$\begin{aligned} \sum_{k=1}^n (\alpha_k^e)^2 &= \sum_{k=1}^n \left[ (1 - \alpha_v^2) \frac{\alpha_k^2 (1 - \rho_k)}{1 - \rho} + 2\alpha_v \left( \sqrt{1 - \alpha_v^2} \right) \frac{\alpha_k \sqrt{1 - \rho_k}}{\sqrt{1 - \rho}} \frac{\alpha_k \sqrt{\rho_k}}{\sqrt{\rho}} + \alpha_v^2 \frac{\alpha_k^2 \rho_k}{\rho} \right] \\ &= \sum_{k=1}^n \left[ (1 - \alpha_v^2) \frac{\alpha_k^2 (1 - \rho_k)}{1 - \rho} + \alpha_v^2 \frac{\alpha_k^2 \rho_k}{\rho} \right] \\ &= \frac{(1 - \alpha_v^2)}{1 - \rho} \sum_{k=1}^n \alpha_k^2 (1 - \rho_k) + \frac{\alpha_v^2}{\rho} \sum_{k=1}^n \alpha_k^2 \rho_k \\ &= \frac{(1 - \alpha_v^2)}{1 - \rho} (1 - \rho) + \frac{\alpha_v^2}{\rho} \rho = (1 - \alpha_v^2) + \alpha_v^2 = 1 \end{aligned} \quad (2.178)$$

Note: in the second step of this equation, the middle term is removed because it is equal to zero (since either  $\rho_k=0$  or  $1-\rho_k=0$ ). In the fourth step, equation (2.168) is used. If not all values of  $\rho_k$  are equal to either 0 or 1, the sum of the squares of the equivalent  $\alpha$ -values is not necessarily equal to 1. In that case, they have to be normalized:

$$\alpha_k^e = \frac{\alpha_k^e}{\sum_{k=1}^n (\alpha_k^e)^2}; k = 1..n \quad (2.179)$$

To summarise, the method for deriving equivalent  $\alpha$ -values is as follows:

- 1 Apply the upscaling method of section 2.4.3.1 on the  $n_e$  components with reliability index  $\beta$  and mutual correlation  $\rho$  to derive the reliability index  $\beta^e$  of the combined (upscaled) component.
- 2 Apply the upscaling method of section 2.4.3.1 on the  $n_e$  components with reliability index  $\beta - \varepsilon_v \sqrt{\rho}$  and mutual correlation  $\rho$  to derive the reliability index  $\beta^e(\varepsilon_v)$  of the combined (upscaled) component.
- 3 Determine  $\alpha_v$  through application of equation (2.173).
- 4 For all all random variables  $k=1..n$ , determine the value of  $\alpha_k^e$  through application of equation (2.177).
- 5 Normalise the equivalent  $\alpha$ -values

The equivalent  $\alpha$ -values of the  $n$  variables for the combined  $n_e$  components are derived after only two applications of the upscaling method (steps 1 and 2 above). This is very efficient, taking into account that the Hohenbichler method needs to be repeated  $2n+1$  times to derive the equivalent  $\alpha$ -values for combining only two components.

#### 2.4.4 System analysis with Monte Carlo methods or numerical integration

The methods described in sections 2.4.3 and 2.4.2 rely on linear approximations of Z-functions of the individual components, which automatically introduces errors in the estimated failure probability of the system. Errors made in the approximation will depend on the system under consideration. If these errors are considered to be unacceptable, a systems approach based on Monte Carlo or numerical integration techniques is an alternative. However, these alternatives generally require computation times that are unacceptably large, which is why for larger systems often the approximative techniques of the previous sections are applied. This is also the reason why this method has not been implemented in Hydra-Ring to compute the probability of failure of the entire system. Nevertheless it is (briefly) described in the current section for Monte Carlo for the purpose of completeness.

Consider the general formulation of the failure probability for a series system, consisting of  $m$  components:

$$P_f = P[Z_1(u_{11}, \dots, u_{1n}) < 0 \cup \dots \cup Z_m(u_{m1}, \dots, u_{mn}) < 0] = P\left[\bigcup_{i=1}^m Z_i(u_{i1}, \dots, u_{in}) < 0\right] \quad (2.180)$$

In which  $u_{ij}$  is the  $j^{\text{th}}$  variable of the Z-function of the  $i^{\text{th}}$  component. For different components  $i$  and  $k$ , the  $j^{\text{th}}$  random variable (i.e.  $u_{ij}$  and  $u_{kj}$ ) in principle refer to the same load or strength variable, but the sampled values can be different because they refer to different components. For instance, the  $j^{\text{th}}$  variable may refer to the thickness of a clay layer. This thickness will be different for different dike segments.

The crude Monte Carlo approach is to sample  $N$  sets of variables  $U_{ij}$ ,  $i=1..m$ ,  $j=1..n$  and then to quantify the percentage of sets for which failure occurs anywhere in the system:

$$P_f = \frac{\sum_{k=1}^N I_{[Z_k^* < 0]}}{N}; \quad Z_k^* = \bigcup_{i=1}^m Z_i(u_{i1k}, \dots, u_{ink}) < 0 \quad (2.181)$$

In which:

$u_{ijk}$  = the  $k^{\text{th}}$  sample of the  $j^{\text{th}}$  variable of the Z-function of the  $i^{\text{th}}$  component.

$I_{[Z < 0]}$  = 1 if  $Z < 0$ , 0 otherwise.

$Z_k^*$  = the Z-function for the  $k^{\text{th}}$  "event"

The N samples basically represent N "events", although the sampling may also involve variables that represent the uncertainty in the resistance of the flood defence system. This crude Monte Carlo procedure seems straightforward, but in the sampling procedure the mutual correlation of the various components needs to be taken into account. For this purpose, similar techniques as in the previous sections can be applied, i.e. two variables of neighbouring components can be related to each other according to equation (2.147) or equation (2.77).

Numerical integration can be theoretically an alternative to Monte Carlo, but since computation times increase an order of magnitude for each additional random variable, this is no serious option in system analysis.

## 2.4.5 Techniques for time and space dependent processes

### 2.4.5.1 Introduction

The techniques described in the previous sections all deal with system analysis of a discrete number of components which may represent dike sections, wind directions, etc. In some applications, however, Z is a function of space and time, which means in principle the number of components is infinite. This is schematically depicted in Figure 2.33. In the left panel, Z is a time-dependent function and failure potentially can occur at any time. On the right, Z is a function of space, and failure can occur at any location.

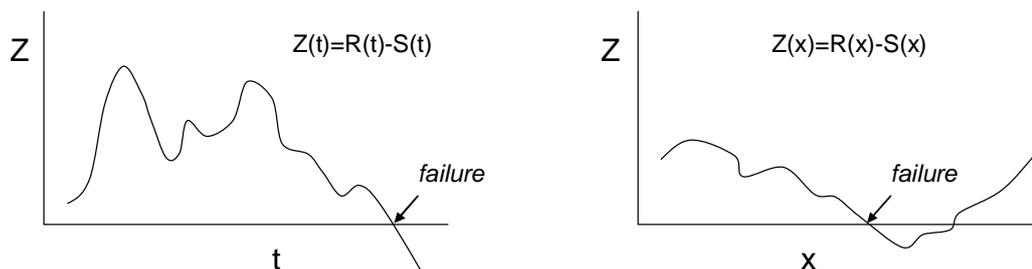


Figure 2.33 Stochastic variation of the Z-function in time (left) and space (right)

This section describes some approaches to deal with these type of continuous descriptions of Z-functions.

### 2.4.5.2 Poisson counting process

The Poisson counting process describes the probability of occurrence of N events, where a single event generally refers to an upcrossing or downcrossing of a threshold value. With respect to failure, the downcrossing of the threshold  $Z=0$  is most relevant. In a Poisson process it is assumed that for small values of  $\Delta t$  [a] the occurrence of an event in a interval [t,

$t+\Delta t$ ] is proportional to  $\Delta t$  and [b] the probability of occurrence of two events occurring in  $[t, t+\Delta t]$  is negligible. This means for small values of  $\Delta t$ , the probability of an event occurring in  $[t, t+\Delta t]$  is approximately equal to:

$$P(1 \text{ event during } [t, t + \Delta t]) \approx \nu \Delta t \quad (2.182)$$

In this equation,  $\nu$  is the 'intensity' of the Poisson process. This is the single parameter that describes the Poisson process. Define  $N(t)$  as the number of events occurring in the time interval  $[0, t]$ . For a Poisson process the probability distribution of  $N(t)$  is:

$$P(N(t) = n) = \frac{(\nu t)^n e^{-\nu t}}{n!} \quad (2.183)$$

The time interval between two subsequent events is also a random variable and it is exponentially distributed. So if  $t_1$  is the time interval between two events, then:

$$P(T_1 \leq t_1) = 1 - e^{-\nu t_1} \quad (2.184)$$

The assumption of a Poisson process is often used to translate exceedance frequencies into exceedance probabilities. Suppose  $\nu$  is expressed as "number of events per year". In that case  $\nu$  is the annual frequency of occurrence. Then, according to equation (2.184), the annual probability of occurrence is equal to:

$$P(T_1 \leq 1) = 1 - e^{-\nu} \quad (2.185)$$

This shows the relation between probability and frequency in case of a Poisson process. An event can be for instance the exceedance of a threshold level  $x$ , for load variable  $X$ . In that case, equation (2.185) can be applied to translate the annual frequency of exceedance of threshold  $x$  into the annual probability of exceedance of threshold  $x$ , or vice versa.

In the description above,  $\nu$  was assumed to be time-independent. If this is not the case, equation (2.183) changes into:

$$P(N(t) = n) = \frac{\left( \int_0^t \nu(\tau) d\tau \right)^n e^{-\int_0^t \nu(\tau) d\tau}}{n!} \quad (2.186)$$

### 2.4.5.3 Outcrossing

If an event refers to failure in a continuous process, i.e. the downcrossing of threshold  $Z=0$  in Figure 2.33, then the outcrossing rate is defined as:

$$\nu = \lim_{\Delta t \downarrow 0} \frac{P[Z(t) \geq 0 \cap Z(t + \Delta t) < 0]}{\Delta t} \quad (2.187)$$

Note that the numerator in this equation is the probability that failure occurs in time interval  $[t, t+\Delta t]$ . The rate  $\nu$  is similar to the one defined in the previous section and can also be time-

dependent:  $v=v(t)$ . Assume for the moment that  $v$  is a constant, i.e. independent of time. The probability that failure occurs in an interval  $(0,T]$ , given the fact that no failure occurs at  $t=0$ , is then equal to:

$$P\left[\min_{t \in (0,T]} \{Z(t)\} < 0 \mid Z(0) \geq 0\right] = 1 - e^{-vT} \quad (2.188)$$

Note that this probability of failure is described by an exponential distribution function. The exponential distribution is by definition the distribution which describes failure probabilities for processes with a constant failure rate (see e.g. Grimmer and Stirzaker, 1982). Figure 2.34 shows an example of an exponential distribution function. In this figure the failure rate,  $v$ , is taken equal to 1, which makes this a standard exponential distribution function.

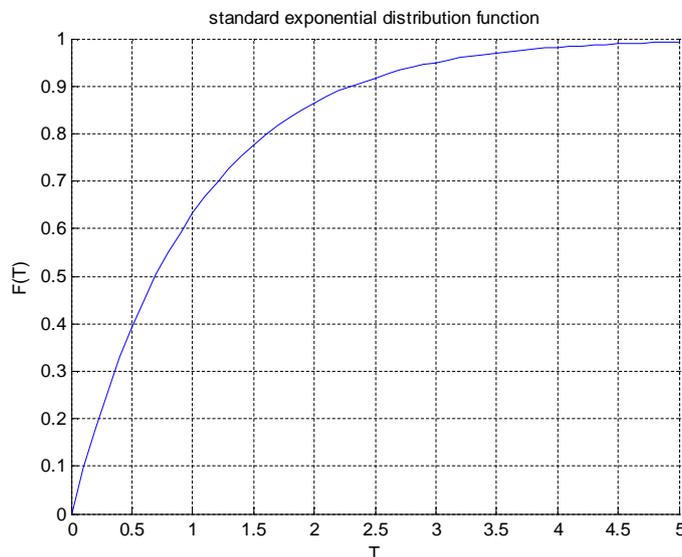


Figure 2.34 Standard exponential distribution function:  $F(T) = 1 - \exp(-T)$ .

The probability that *no* failure occurs in an interval  $(0,T]$ , given the fact that no failure occurs at  $t=0$ , is equal to:

$$P\left[\min_{t \in (0,T]} \{Z(t)\} \geq 0 \mid Z(0) \geq 0\right] = e^{-vT} \quad (2.189)$$

The probability that no failure occurs in an interval  $[0,T]$  is then equal to:

$$P\left[\min_{t \in [0,T]} \{Z(t)\} \geq 0\right] = [1 - P_F(0)] e^{-vT} \quad (2.190)$$

In which  $P_F(0)$  is the initial probability of failure (see below for more information on this probability), i.e the probability that  $Z < 0$  at  $t=0$ . The probability,  $P_F$ , that failure occurs in an interval  $[0,T]$  is equal to:

$$P_F(T) = P\left[\min_{t \in [0,T]} \{Z(t)\} < 0\right] = 1 - [1 - P_F(0)] e^{-vT} \quad (2.191)$$

If the outcrossing rate,  $\nu$ , and the initial failure probability,  $P_F(0)$ , are small, the probability of failure can be approximated by:

$$P_F(T) \approx P_F(0) + \nu T \quad (2.192)$$

This is an upper bound of the failure probability. In essence, this approximation “double counts” the probability of events in which two or more failures occur in the interval  $[0, T]$ . If  $\nu$  and  $P_F(0)$  are small, the probability of two or more failures occurring in the interval  $[0, T]$  is negligible and therefore equation (2.192) is a good approximation in that case.

In the equations above, failure rate  $\nu$  was assumed to be constant. If this is not the case, equation (2.191) changes into the following, more general, equation:

$$P_F(T) = 1 - [1 - P_F(0)] \exp\left(-\int_0^T \nu(t) dt\right) \quad (2.193)$$

The equations above can also be used if  $Z$  is a function of space. In that case,  $t$  and  $T$  need to be replaced by  $x$  and  $X$ , where  $x$  represents distance, e.g. the longitudinal distance along a dike section.

In Hydra-Ring, the outcrossing method is applied in time as well as in space. First, the probabilities of failure of the smallest “components” are computed with the probabilistic techniques for single components as described in section 2.3. The smallest component considered in Hydra-Ring is a cross section of a flood defense (space) during a tidal period (time). So, the initial result of the probabilistic procedure is the probability that failure occurs at a certain cross section within the time-span of a tidal period. This result will be used as  $P_F(0)$  in the equations above, i.e. the initial failure probability. Subsequently the outcrossing approach is applied for upscaling the probability of failure from a cross-section to a dike section and from a tidal period to a year.

The failure rate  $\nu(t)$  or  $\nu(x)$  needs to be derived from spatial and temporal autocorrelations of the strength and load variables. This will be described in more detail in sections 2.5.2 and 2.5.5. In general, functions  $\nu(t)$  and  $\nu(x)$  are too complex to solve equation (2.193) analytically, which means approximating techniques are required. Hydra-Ring uses different outcrossing approaches for space and time because of mutual differences in autocorrelation structures.

Note that the component for which  $P_F(0)$  is computed in Hydra-Ring has a “width” equal to 0 in space (a cross section) whereas the “width” in time is not equal to zero (a tidal period). This has to do with the fact that the input statistics of random load variables like sea water level, river discharge or wind speed represent probabilities of the maximum value in a tidal period (see also section 3.3.3). These values are therefore suitable to represent the whole tidal period. Correlation in space is mainly determined by strength variables and statistics of strength variables do generally not refer to a maximum value over a considerable length. On the other hand, a dike breach also has a certain width, i.e. does not occur only on a theoretical cross section with zero width. Therefore, in practice often a certain width is awarded to a cross section and the initial failure probability,  $P_F(0)$  is awarded to this width. This means a (slight) reduction in the length of the remainder of the dike section and hence a (slight) reduction in the computed failure probability.

## 2.5 Combining failure probabilities of components – choices and implementations in Hydra-Ring

### 2.5.1 General overview

Hydra-Ring computes the failure probability of a flood defence *system*, which is composed of a (large) number of components. Section 2.3 described techniques for computing failure probabilities for single components and section 2.4 described generic techniques for combining probabilities of single components to derive the failure probability of a system. The current section describes the methods for system analysis that are implemented in Hydra-Ring. The system analysis for failure probabilities in Hydra-Ring involves:

- Starting point: failure probability for a single mechanism and cross-section;
- Combining different closure scenarios of tidal barriers;
- Integrating in space from a cross-section to a flood defence segment;
- Combining different wind directions;
- Integrating in time from small time periods (<1 day) to a year;
- Combining different failure mechanisms;
- Combining different flood defence segments;

Note that the combining methods are also applied on random variables like wind direction and closure scenarios. In that sense, these random variables are treated fundamentally different from other random variables like wind speed and river discharge. The latter group is dealt with in the probability computation for a single component (see section 2.3). The reason why the wind direction is treated differently is because it is a “cyclic variable”. In section 2.2.3 it was already mentioned that cyclic variables can not be represented in a meaningful way by standard normal u-variables. This is because potential outcomes of these variables are not “ordered” from small to large. The domain for these variables is 0-360 degrees, but 360 degrees is not “larger” than 0 degree. For this reason, the wind direction cannot be included in the probabilistic computation for a single component. Instead, potential outcomes of the wind direction are clustered in sectors, e.g. 16 sectors of 22.5 degrees. For each sector, a probabilistic computation of a single component is executed. The resulting probabilities for the 16 sectors are combined with a special version of the Hohenbichler method. Section 2.5.3 describes the implemented method in Hydra-Ring for combining wind directions.

For the same reason, the random variable that represents the “functioning of the storm surge barrier” (which can fail to function with a certain probability) is also treated similar to the wind directions. This means that each potential state of the surge barrier is treated as a “single component” and the failure probabilities of the different states are subsequently combined. This procedure is described in section 2.5.4.

### 2.5.2 Temporal upscaling

#### 2.5.2.1 Introduction

Failure probabilities in Hydra-Ring are first computed for relatively small time scales (<1 day) in which the temporal variation of the relevant hydraulic variables (water level, wind etc) is small enough to be assumed constant. This time scale will be referred to as the *basic time scale*. In Hydra-Ring this time scale is typically set equal to a tidal period (12 hours and 25 minutes) but this can be varied by a user.

Output failure probabilities are generally expressed per year, which means the failure probabilities of the basic time scale need to be scaled up into an annual failure probability. In

Hydra-Ring, the strength/resistance of the flood defence system is assumed to be time-invariant. This means temporal variation is only present in the hydraulic load variables. In the hydraulic load model, two groups of variables are distinguished:

- 1 Fast evolving random variables;
- 2 Slowly evolving random variables.

For the first group, sampled values for successive basic time steps are assumed to be independent, whereas for the second group there is significant dependence. The first group consists of variables like wind speed and sea water level, the second group consists of variables like river discharges and lake water levels.

For the slowly evolving variables a characteristic time scale needs to be defined. The time scale of a load variable is the time at which autocorrelation of the load variable reaches zero. For example, for a discharge hydrograph, the time scale may be on the order of several weeks. Each discharge hydrograph is assumed to be independent from the hydrograph that preceded it (no correlation). The number of time scales in a computation will depend on the number of unique time scales of the load variables. However, in Hydra-Ring there is currently only the option to take into account one time scale for slowly-evolving variables. This means this one time scale has to be applied on all load variables for which autocorrelation in time needs to be taken into account. Multiple time scales may be used for the fast evolving variables, i.e. the variables for which autocorrelation in time can be ignored. The time scale for the slowly evolving variable will always be the largest time scale.

Within the probabilistic model, care must be taken how the smaller scale variables and more slowly-evolving variables are combined. For this purpose, three methods have been implemented in Hydra-Ring: FBC, NTI, and APT. FBC stands for Ferry Borges & Castanheta model, NTI stands for numerical time integration, and APT stands for arbitrary point in time. How the slow load variable is handled in the probabilistic analysis has an impact on the computation structure, which will be described in Appendix B. The sections below describe the mathematical background of the three methods.

### 2.5.2.2 Method 1: Numerical time integration (NTI)

NTI stands for numerical time integration and refers to a method in which the failure probabilities are combined over the smaller time period  $\tau$  within the time scale  $T_m$  of the slowly evolving load variable. Since the slowly evolving variables are typically river discharge or lake level, the temporal evolution within the time scale  $T_m$  can be thought of as a hydrograph with a rising limb and falling limb. In Hydra-Ring these are schematized by a trapezium, as shown in Figure 2.35. The trapezium has a total duration  $T_m$  and is divided into  $T$  components of length  $\tau$ , with  $T = T_m/\tau$ . The duration of the peak value is defined as  $T_p$ .

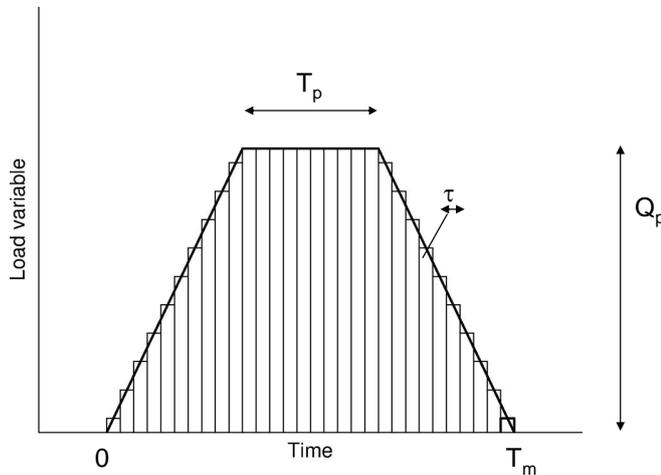


Figure 2.35 Example of a trapezium model used for modeling slowly-evolving load variables; the finer discretization represents a smaller time scale of additional load variable(s)

In NTI, the failure probabilities are determined separately for each period of length  $\tau$ , and subsequently the probabilities are combined to derive the probability of failure for the entire period of length  $T_m$ . Each period of length  $\tau$  can be considered as a separate component of a series system, i.e. the system fails if at least one of the components fails. The probability of failure for the entire period of length  $T_m$  can thus be described as:

$$P[Z_{T_m} < 0] = P\left[\bigcup_{j=1}^T (Z_j < 0)\right] \quad (2.194)$$

Where  $Z_j$  is the Z-function of the  $j^{\text{th}}$  period of length  $\tau$ . To determine this failure probability, consider the trapezium of Figure 2.35, which is characterized by:

- a base duration  $T_m$ ;
- a peak duration  $T_p$ ; and
- a peak value  $Q_p$ .

The first two,  $T_m$  and  $T_p$  are constants that are defined by the person who implements the load model. The third,  $Q_p$ , is a random variable with a user-defined distribution function. In order to make a clear distinction between random variable  $Q_p$  and constants  $T_m$  and  $T_p$ , the hydrograph is described as the product of  $Q_p$  and a dimensionless hydrograph:

$$Q(t) = q(t)Q_p \quad (2.195)$$

In which

- $Q(t)$  = hydrograph of the slowly evolving variable;
- $q(t)$  = dimensionless hydrograph of the slowly evolving variable;
- $Q_p$  = the peak of the hydrograph of the slowly evolving variable.

Function  $q(t)$  is dimensionless and has a peak value of 1. In case of a trapezium,  $q(t)$ , is fully determined by the choice of the user-defined parameters  $T_m$  and  $T_p$  and is therefore deterministic. Each period of time scale  $\tau$  is small enough to assume  $Q$  and  $q$  to be constant

within the interval. For each component  $j, j=1..T$ , the value of function  $q(t)$  is taken equal to the centre value in the interval:

$$q_j = q\left(\frac{j-0.5}{T}\right) ; j=1..T \quad (2.196)$$

The probability of failure within time scale  $T_m$  can be written as follows (according to the theory of total probability):

$$P[Z_{T_m} < 0] = \int P[Z_{T_m} < 0 | Q_p] dQ_p = \int P\left[\bigcup_{j=1}^T (Z_j < 0 | Q_p)\right] dQ_p \quad (2.197)$$

The temporal evolution of the fast evolving load variables is such that the resulting values of these variables in subsequent time steps can be assumed independent. This means that in the simple case where the strength parameters can be considered as known constants, the conditional probabilities of equation (2.197) are mutually independent. This simplifies the probability of failure for the period  $T_m$  to:

$$P[Z_{T_m} < 0 | Q_p] = 1 - \prod_{j=1}^T \{1 - P(Z_j < 0 | Q_p)\} \quad (2.198)$$

In general, however, strength parameters cannot be considered as known constants, which means the uncertainties in these parameters need to be taken into account. The autocorrelation of these strength parameters is approximately equal to one, even for large time scales. Consequently, equation (2.197) is not valid anymore and a different approach is required. For this reason, Hydra-Ring applies the Hohenbichler method of section 2.4.2 to solve equation (2.194). First the failure probabilities of the individual components (the smaller time scales) are quantified. In this step, the peak value of the trapezium,  $Q_p$ , is treated as a random variable. So Hydra-Ring does not determine  $P(Z_j < 0 | Q_p)$ , but directly  $P(Z_j < 0)$ . The advantage of this approach is that  $Q_p$  can be treated as one of the parameters in a FORM computation, instead of having it evaluated through numerical integration. This can save significant computation time. So, for each component  $j$  the probability that failure occurs is determined, taking all potential values of the peak,  $Q_p$ , into account.  $P(Z_j < 0)$  is therefore determined as follows (applying the theorem of total probability on  $Q_p$ ):

$$P[Z_j < 0] = \int P[Z_j < 0 | Q_p] dQ_p = \int P[Z_j < 0 | Q = q_j Q_p] dQ_p \quad (2.199)$$

Note that:

- $q_j$  is a known constant for each  $j$  (see equation (2.196))
- $P(Z_j < 0)$  is different for different intervals, because  $q_j$  is different for different intervals. The failure probabilities for components around the peak value are higher than failure probabilities for components near the beginning or end of the standardized hydrograph.

The conditional probabilities of equation (2.199) are computed for each component separately with the probabilistic computation techniques of section 2.3. The probability of failure for the entire period  $T_m$  is then determined by combining the components with the Hohenbichler

method of section 2.4.2. The Hohenbichler method requires correlations between the Z-functions of the components. According to equation (2.119) these correlations are based on [a] autocorrelations of the random variables and [b]  $\alpha$ -values of the random variables. The  $\alpha$ -values are determined when deriving  $P[Z_j < 0]$  for  $j=1..T$ . The autocorrelations are taken as follows:

- Resistance variables: autocorrelation=1;
- Slowly evolving load variables: autocorrelation=1;
- Fast evolving load variables: autocorrelation=0.

Notes:

- Because the combining techniques of section 2.4 use the assumption of linear Z-functions there will be differences in the results compared to the “exact” formulation as described in equation (2.197).
- Naturally, the assumptions on autocorrelations are simplifications. The main assumption is the assumption that the autocorrelations of the fast evolving variables is equal to 0 between two successive periods of time scale  $\tau$ . This is an assumption that is also done in other probabilistic models in the Netherlands like Hydra-B or Hydra-Zoet. In Hydra-Ring, this assumption on the autocorrelation can be easily changed if a positive autocorrelation is believed to be more realistic.
- In this chapter it is assumed that all resistance variables are constant in time. In some failure mechanisms this is not the case for all stochastic variables. An example of this is the seepage ditch water level, which play a role in the piping mechanism.

Subsequently, the failure probabilities for the hydrographs need to be upscaled to a year:

$$P[Z_{year} < 0] = P\left[\bigcup_{j=1}^{N_T} (Z_{Tm,j} < 0)\right] \quad (2.200)$$

In which  $Z_{year}$  is the Z-function for the period of one year and  $Z_{Tm,j}$  is the Z-function of the  $j^{\text{th}}$  hydrograph within a year and  $N_T$  is the number of hydrographs in a year. In this computation the autocorrelation of all load variables is taken equal to 0. However, the autocorrelation of strength parameters is still higher than zero, which means the Z-functions of the hydrographs are not independent. Therefore, equation (2.200) is solved with the Hohenbichler method (section 2.4.2). If the autocorrelation between subsequent hydrographs blocks is large, the combining method for equal components (section 2.4.3) can also be used.

### 2.5.2.3 Method 2: FBC

The FBC model, named for J. Ferry Borges and M. Castanheta (Ferry Borges & Castanheta 1971), assumes a rectangular shape for the evolution of the slowly-evolving load variable, referred to in this context as a block. Figure 2.36 illustrates the concept.

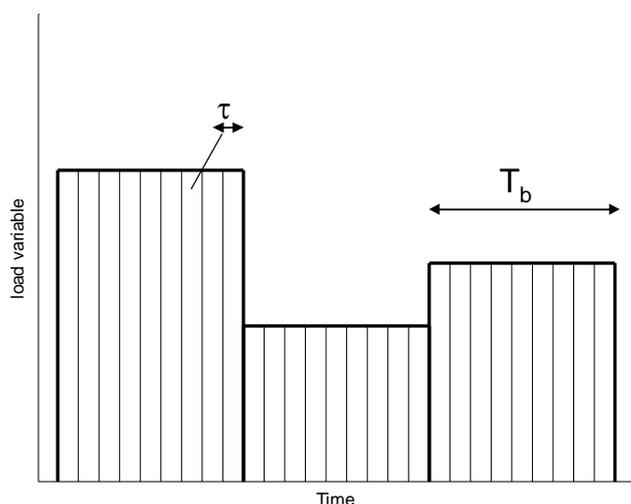


Figure 2.36 Illustration of the FBC model used for modeling slowly-evolving load variables; the finer discretization represents a smaller time scale

Within a block, the represented (slowly varying) load variable is assumed to be fully correlated and between blocks there is constant correlation (in Hydra-Ring often assumed to be zero). For simplicity, the reader can assume a situation in which there are only two relevant time scales:  $\tau$  (fast evolving variables) and  $T_b$  (block duration of slowly evolving variables). The block duration  $T_b$  is typically somewhere between  $T_p$  and  $T_m$  as defined in the previous section.

The block duration does not represent the duration of the evolution of the variable (e.g. a complete discharge hydrograph), but rather represents the duration that the peak value of the block is exceeded on average per exceedance. The duration of a block is derived from a duration curve,  $N(q)$ :

$N(q)$  = average duration of a single exceedance of level  $q$ .

The duration curve is an input function for Hydra-Ring. The method to derive the duration is an iterative procedure. First an initial estimate is done. For this initial estimate, the FORM routine is executed and an initial estimate of the failure probability is derived. Furthermore, the value of  $Q$  in the design point is known:  $Q=q_d$ . The value of  $N(q_d)$  is then used as the estimate of the duration in the second iteration step. The iteration procedure continues until the difference in durations of subsequent iteration steps is sufficiently small. Generally, this procedure converges fast, i.e. only a few iteration steps are required.

The description of supported duration curves and an explanation how such a curve is derived can be found in the statistical chapter, section 3.3. In the FBC method, the probability of failure is first determined for the time period  $\tau$  (i.e. the time scale of the fast evolving variables) using the standard probabilistic techniques as described in section 2.3. Subsequently this probability is scaled-up to one year in two steps:

- 1 Upscaling from time period  $\tau$  (time scale of the fast evolving variables) to the FBC block duration;
- 2 Upscaling from the FBC block duration to a year.

In step 1 the following probability is computed:

$$P[Z_{T_b} < 0] = P\left[\bigcup_{j=1}^{N_b} (Z_j < 0)\right] \quad (2.201)$$

Where  $N_b$  is the number of smaller time steps in a block duration ( $N_b = T_b/\tau$ ) and  $Z_j$ ,  $j=1..N_b$  is the Z-function of the  $N_b$  time steps. Equation (2.201) is solved with the upscaling method of section 2.4.3, i.e. the combining method for equal components. This method requires correlations between the Z-functions of the individual components, i.e. the time steps of length  $\tau$ . According to equation (2.139) these correlations are based on [a] autocorrelations of the random variables and [b]  $\alpha$ -values of the random variables. The  $\alpha$ -values are determined when deriving the probability of failure of the individual components,  $P[Z_j < 0]$  for  $j=1..N_b$ . The autocorrelations are taken as follows (similar to the NTI method):

- Resistance variables: autocorrelation=1;
- Slowly-evolving load variables: autocorrelation=1;
- Fast evolving load variables: autocorrelation=0.

In step 2, the following probability is computed:

$$P[Z_{year} < 0] = P\left[\bigcup_{j=1}^{N_y} (Z_{T_b,j} < 0)\right] \quad (2.202)$$

In which  $Z_{year}$  is the Z-function for the period of one year and  $Z_{T_b,j}$  is the Z-function of the  $j^{\text{th}}$  block within a year. In this computation the autocorrelation of all load variables is taken equal to 0. However, the autocorrelation of strength parameters is still higher than zero, which means subsequent Z-functions are not independent. Therefore, the upscaling in step 2 is done with the combining method for equal components (section 2.4.3).

The FBC method is computationally faster than NTI. Naturally, the representation of time-variation of variables in “blocks” is an approximation, but so is a representation with trapeziums or any other standardized shape.

#### 2.5.2.4 Method 3: Arbitrary point in time (APT)

APT stands for arbitrary point in time, and refers to the stochastic handling of the timing within the standardized hydrograph of the slowly evolving load variable. Just as with NTI, a trapezium-shaped hydrograph is assumed for the slowly evolving load variable (see Figure 2.35) and this hydrograph is divided into  $T$  smaller time periods of length  $\tau$ .

Function  $q(t)$  describes the dimensionless hydrograph (see section 2.5.2.2 on NTI) which is defined on the interval  $[0, T_m]$ , where  $T_m$  is the time scale of the slowly evolving random variable. Consider a random variable  $c$ , that is equal to the value of function  $q$  at an arbitrary moment,  $t_c$ , in the interval  $[0, T_m]$ :

$$c = q(t_c) \quad (2.203)$$

Arbitrary means that  $t_c$  is taken from a uniform distribution function over  $[0, T_m]$ :

$$t_c \sim U(0, T_m) \quad (2.204)$$

Then:

$$P[C < c] = P[q(t_c) < c] \quad (2.205)$$

Figure 2.37 shows the cumulative distribution function,  $F_C(c)$ , of  $c$  in case the standardized hydrograph,  $q(t)$ , is a trapezium. To explain this figure, define  $T_p$  as the duration that the trapezium is at its peak value. Then, the probability that at an arbitrary moment,  $t_c$ , in the interval  $[0, T_m]$ , the function  $q(t_c)$  is equal to its peak value ( $q=1$ ) is:

$$P_1 = P[C = 1] = \frac{T_p}{T_m}, \quad (2.206)$$

So the probability that  $q(t_c)$  is less than 1 is:

$$P[C < 1] = 1 - P_1 = 1 - \frac{T_p}{T_m} = \frac{T_m - T_p}{T_m}, \quad (2.207)$$

Since the rising and falling limb of the trapezium are linear, the probability  $1 - P_1$  is uniformly spread over the interval  $[0, 1)$ , in formula:

$$f_c(C | C < 1) = \frac{T_m - T_p}{T_m}, \quad (2.208)$$

This results in the probability distribution function as displayed in Figure 2.37 and described as:

$$F_C(c) = P[C < c] = \begin{cases} 0 & ; c \leq 0 \\ c \frac{T_m - T_p}{T_m} & ; 0 < c < 1 \\ 1 & ; c \geq 1 \end{cases} \quad (2.209)$$

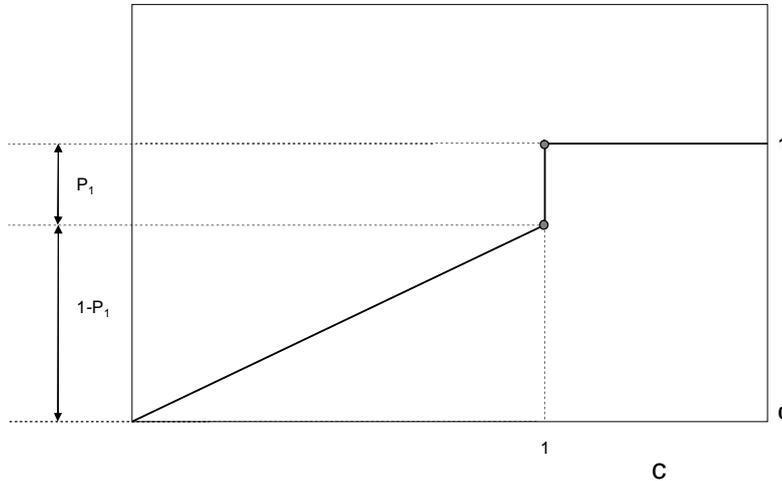


Figure 2.37 – Schematic of the cumulative distribution function of  $c$

To sample variable  $c$ , first a random number,  $v$ , from the uniform distribution is sampled, after which the value of  $c$  is computed as:

$$c = \min \left( \left( \frac{T_m}{T_m - T_p} v \right), 1 \right), \quad (2.210)$$

According to equation (2.195), the value of the slow-evolving variable  $Q$  at an arbitrary moment,  $t_c$ , is the product of the peak discharge  $Q_p$  and the value of the standardized hydrograph,  $c=q(t_c)$ . In other words:  $Q$  can now be described as the product of two random variables,  $Q_p$  and  $c$ . According to the theorem of total probability, the probability that failure occurs at an arbitrary moment,  $t_c$ , can be written as follows:

$$P[Z_{t_c} < 0] = \int \int P(Z_{t_c} < 0 | c, Q_p) f_c(c) f_{Q_p}(Q_p) dc dQ_p \quad (2.211)$$

The conditional probability of failure  $P(Z_{t_c} < 0 | c, Q_p)$  can be computed with the probabilistic computation techniques of section 2.3. Equation (2.211) could be solved with numerical integration over variable  $c$ , which in essence is the NTI method. However, in the APT-procedure, the random variable  $c$  is directly included in the evaluation of  $P(Z_{t_c} < 0)$ . This means that if the FORM-method is applied to solve  $P(Z_{t_c} < 0)$ , variable  $c$  is an additional variable in the FORM-procedure. Since FORM is generally much faster than numerical integration, this saves valuable computation time in comparison with the NTI procedure.

The resulting value of  $P[Z_{t_c} < 0]$  is the probability that failure occurs in the smallest time scale of the model, i.e. the time scale of the fast evolving variables. This probability is scaled-up to one year in two steps:

- 1 Upscaling from time period  $\tau$  (time scale of the fast evolving variables) to the time period  $T_m$  time scale of the slowly-evolving variables).
- 2 Upscaling from  $T_m$  to a year.

In step 1 the upscaling method of section 2.4.3 is used, i.e. the combining method for equal components. This is because each component at time scale  $\tau$  is treated as an arbitrary moment in time, which means each component is essentially the same. This is the essential difference with the NTI method, where each component is associated with a specific moment in the hydrograph, which means each component has a different probability of failure. The upscaling method requires correlations between the Z-functions of the individual components, i.e. the time steps of length  $\tau$ . For this purpose, the following assumptions are made on the autocorrelations of the individual random variables (similar to NTI and FBC):

- Resistance variables: autocorrelation=1;
- Slowly-evolving load variables: autocorrelation=1;
- APT-variable  $c$ : autocorrelation=0;
- Fast evolving load variables: autocorrelation=0.

Upscaling to a period of one year (step 2) is done in the same way as the NTI-method.

### *Required adaptations to APT*

In areas where the slowly-evolving variables are dominant, the APT method is known to have some practical problems that may lead to significant errors. Such areas are for instance river stretches with hardly or no tidal influence. In that case the river discharge, which is the slowly evolving variable, is the dominant factor. The problems are explained below and also the solutions are described.

#### **Problem 1**

In case of a dominant slowly evolving variable,  $c$  becomes a dominant random variable and its value may be close to 1, which is the upper limit of variable  $c$ . If the dimensionless hydrograph,  $q(t)$ , is a trapezium this means  $c$  is in the area where the distribution function  $F_C(c)$  makes a “jump” (see Figure 2.37). This is a discontinuity and this may cause the probabilistic method FORM (see section 2.3.6) not to converge, because the partial derivative of variable  $c$  in equation (2.49) varies strongly around the discontinuity.

This problem can be solved by splitting the trapezium in two components: the part where the trapezium is at its peak and the combined rising and falling limb. The probability of failure for the whole trapezium then becomes:

$$P(Z_{trapezoid} < 0) = P(Z_{peak} < 0 \cup Z_{limb} < 0) \quad (2.212)$$

The first part,  $Z_{peak}$ , can be modeled by an FBC block, as described in the previous section. In the second part,  $Z_{limb}$ , variable  $c$  has a uniform distribution. This mean  $F_C(c)$  has no discontinuities and the APT method can be applied. Subsequently, the two Z-functions are combined with the Hohenbichler method.

#### **Problem 2**

Consider the APT model for  $Z_{limb}$ . Variable  $c$  now has a uniform distribution:

$$\begin{aligned} F_C(c) &= c \quad ; 0 \leq c \leq 1 \\ f_C(c) &= 1 \quad ; 0 \leq c \leq 1 \end{aligned} \quad (2.213)$$

Suppose the slowly evolving variable  $Q$  is the only random variable involved:

$$Z = R - AQ \quad (2.214)$$

In this Z-function, R and A are constants. In the APT method, Q is represented by two random variables: c and  $Q_p$ . Application of the FORM procedure to derive the failure probability for the small time period  $\tau$  will result in  $\alpha$ -values for c and  $Q_p$  and a reliability index,  $\beta$ . The Z-function can then be written as:

$$Z = \beta + \alpha_1 u_1 + \alpha_2 u_2 \quad (2.215)$$

Where  $u_1$  represents c and  $u_2$  represents  $Q_p$ . The next step is to upscale the Z-function to the period of time scale  $T_m$ . For this purpose, the method of section 2.4.3 is used, i.e. the combining method for equal components. In the upscaling process,  $u_1$  is a standard normally distributed variable. This means  $u_1$  does not have an upper limit, even though it represents variable c which does have an upper limit (i.e. 1). The larger values of  $u_1$  will therefore unrightfully contribute to the failure probability in the upscaling process. This will lead to an overestimation of the failure probability, especially if the value of c in the design point is close to 1. Application of the APT method on a number of academic tests revealed that this overestimation indeed occurs.

The APT method therefore needs to be corrected in such a way that this overestimation is ruled out. In Vrouwenvelder et al [2011] a practical solution was found for this. Application of the adapted method on a number of academic test problems showed that the problems were resolved, i.e. failure probabilities were not overestimated anymore.

Ergens In 2013 gaat nog verder onderzoek gedaan worden naar de oplosmethode. Zodra dat beschikbaar is, worden de relevante tekstdelen overgenomen in het onderhavige document en waar nodig aangepast ten behoeve van het beoogde lezerspubliek.

### 2.5.3 Combining over wind directions

#### 2.5.3.1 Straightforward approach; not used in Hydra-Ring

Wind directions can take on any value between 0 and 360 degrees. For practical purposes, the interval [0,360] is divided into a number of sectors  $\phi_1 \dots \phi_{N_s}$ , e.g. 12 sectors of 30 degrees or 16 sectors of 22.5 degrees. Statistical characteristics for all wind directions within a single sector are then assumed to be the same, whereas statistics between different sectors will vary. The combined sectors cover all potential outcomes of the wind direction (0-360 degrees), which means the probability of failure of a flood defence can be written as:

$$P(Z < 0) = P\left(\bigcup_{i=1}^{N_s} Z_i < 0 \cap \phi \in \phi_i\right) \quad (2.216)$$

Where  $Z_i$  is the Z-function for sector i,  $i=1..N_s$  and  $\phi$  is the wind direction. Wind sectors are generally made disjunct, which means equation (2.216) can be written as:

$$P(Z < 0) = \sum_{i=1}^{N_s} P(Z_i < 0 \cap \phi \in \phi_i) \quad (2.217)$$

Taking into account that  $P(A \cap B) = P(A|B)P(B)$ , equation (2.217) can be further rewritten as:

$$P(Z < 0) = \sum_{i=1}^{N_s} P(Z_i < 0 | \phi \in \phi_i) \cdot P(\phi \in \phi_i) \quad (2.218)$$

According to this equation, the probability of failure can be derived *given* the wind direction, and subsequently the different contributions of the wind directions are accumulated. The advantage of this approach is that wind directions are treated separately, taking into account differences in statistics of variables like wind speed and sea water level that depend on the wind direction. Equation (2.218) in essence is the approach where the wind direction is taken into account through numerical integration.

### 2.5.3.2 Method as programmed in Hydra-Ring

Even though the method as described in the previous section is straightforward and easy to comprehend, it is not applied in Hydra-Ring. The reason for this is that the method does not provide an  $\alpha$ -value for the wind direction and this  $\alpha$ -value is required for subsequent combining of failure mechanisms and dike segments (see section 2.5.6). Therefore an alternative method is applied that is largely based on the method of Hohenbichler (section 2.4.2).

Consider equation (2.218). The probability per wind direction that needs to be derived is:

$$P(Z_i < 0 \cap \phi \in \phi_i) = P(Z_i < 0 | \phi \in \phi_i) \cdot P(\phi \in \phi_i) \quad (2.219)$$

Now define function  $Z_i'$  as follows:

$$Z_i' = u_i - \Phi^{-1}(P(\phi \in \phi_i)) \quad (2.220)$$

Where  $u_i$  is a standard normally distributed variable and  $\Phi$  is the standard normal distribution function. It is easy to verify that:

$$P(Z_i' < 0) = P(\phi \in \phi_i) \quad (2.221)$$

So  $Z_i'$  is a Z-function that represents the probability of occurrence of wind direction  $\phi_i$ . This means equation (2.219) can be rewritten as:

$$P(Z_i < 0 \cap \phi \in \phi_i) = P(Z_i < 0 \cap Z_i' < 0) = P(Z_i < 0 | Z_i' < 0) \cdot P(Z_i' < 0) \quad (2.222)$$

In order to combine these two probabilities, first the correlation coefficient,  $\rho$ , between  $Z_i$  and  $Z_i'$  needs to be established. The following formula is used to derive  $\rho$ :

$$\rho(Z_i, Z_i') = \sum_{k=1}^n (\alpha_k)^2 \rho_k \quad (2.223)$$

Where  $\rho_k$  is the autocorrelation of variable  $k$  and  $\alpha_k$  is the  $\alpha$ -value of variable  $k$  as derived in the determination of  $P[Z < 0]$ . The following is assumed for the autocorrelation:

- Resistance parameters:  $\rho=1$ ;
- Slowly evolving variables:  $\rho=1$ ;
- Fast evolving variables:  $\rho=0$ .

In other words: only for the load variables that fluctuate as fast as the wind direction the autocorrelation is assumed to be equal to 0. This assumption is the same as was done in section 2.5.2 for the temporal upscaling methods NTI, FBC and APT. Naturally, the assumptions on autocorrelations are simplifications. The main assumption is that the autocorrelations of the fast evolving variables is equal to 0 between two successive periods of time scale  $\tau$ . This is an assumption that is also done in other probabilistic models in the Netherlands like Hydra-B, Hydra-Zoet. In Hydra-Ring, this assumption on the autocorrelation can be easily changed if a positive autocorrelation is believed to be more realistic.

Function  $Z_i$  can be rewritten as follows

$$Z_i = \alpha_{c,i} u_{c,i} + \alpha_{u,i} u_{u,i} \quad ; i = 1..N_s \quad (2.224)$$

In this description,  $u_{c,i}$  represents the  $u$ -values of the variables that are fully correlated in time, whereas  $u_{u,i}$  represents the  $u$ -values of the variables that are fully uncorrelated in time. Note that:

$$\alpha_{c,i} = \sum_k^n (\alpha_{ik})^2 \rho_k = \rho \quad ; \alpha_{u,i} = \sqrt{1 - (\alpha_{c,i})^2} = \sqrt{1 - \rho^2} \quad (2.225)$$

Naturally, the sum of the squares of  $\alpha_{c,i}$  and  $\alpha_{u,i}$  is equal to 1. The method that is applied in Hydra-Ring takes into account the fact that after combining  $Z_i$  with  $Z'_i$ , the  $u$ -value of the the correlated part,  $u_c$  does not change. This can be explained with the simple example of a  $Z$ -function with only one variable,  $U$ , that is fully correlated to the wind direction:

$$Z = \beta + \alpha u = \beta - u \quad (2.226)$$

The reliability index of this function is equal to  $\beta=u$ , which is also the value of  $u$  in the design point. Now after combining with  $Z'$ , the failure probability,  $P[Z<0]$ , remains the same because  $Z$  is fully correlated to the wind direction. This means the reliability index,  $\beta$ , remains the same and hence the value of  $u$ . This shows that the value in the design point does not change for variables that are fully correlated to the wind direction.

If  $U$  is fully uncorrelated to the wind direction, the combination of  $Z_i$  and  $Z'_i$  can be done by simple multiplication:

$$P[Z_i < 0 \cap Z' < 0] = P(Z_i < 0) P(Z' < 0) \quad (2.227)$$

This means a decrease in failure probability and hence an increase in  $\beta$ . The new  $\beta$  can be computed as follows:

$$\beta_{new} = \Phi^{-1} [P(Z_i < 0) P(Z' < 0)] \quad (2.228)$$

The increase in  $\beta$  is then of course equal to:

$$\Delta\beta = \beta_{new} - \beta \tag{2.229}$$

The examples above show cases where U (and hence Z) is either fully correlated or fully uncorrelated to the wind direction. In reality, Z can also be partially correlated to the wind direction, see the example of equation (2.224). In that cases, the U-value of the correlated part in the design point remains unchanged, whereas the change in  $\beta$  for the uncorrelated part is computed according to equation (2.229). In order to compute the change in beta for the complete Z-function, the value of  $\Delta\beta$  needs to be multiplied by the  $\alpha$ -value of the uncorrelated part (since only the uncorrelated part causes a change in the reliability index):

$$\Delta\beta = \Delta\beta\alpha_{u,i} = \Delta\beta\sqrt{1-\rho^2} \tag{2.230}$$

This value of  $\Delta\beta$  is added to the  $\beta$  of function  $Z_i$ . This is the way in which the influence of the probability of the wind direction is taken into account. This method is schematized in Figure 2.38.

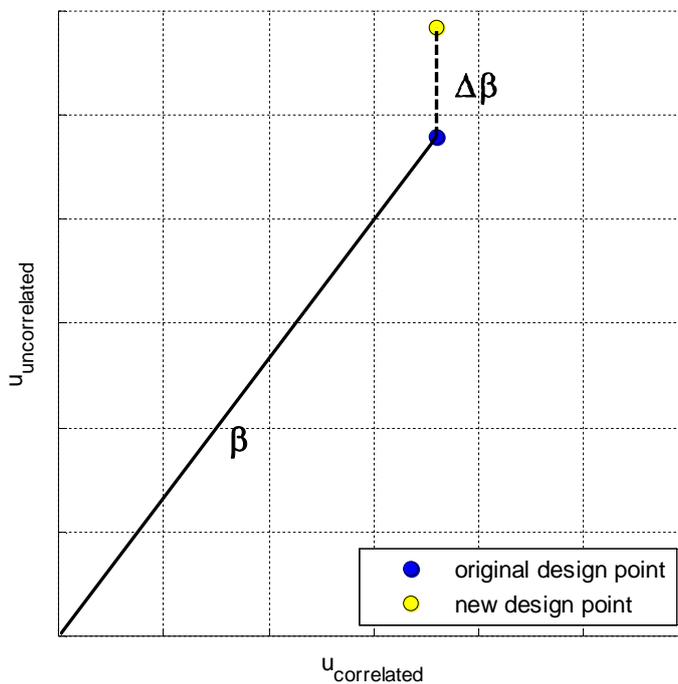


Figure 2.38 Schematised view of computing the new design point and new beta as a result of including the probability of occurrence of the wind direction.

The result of the method for combining  $Z_i$  and  $Z_i'$  is a set of  $\beta$  and  $\alpha$ -values for all wind directions:

$$Z_i^e = \beta_i^e + \alpha_{i1}^e u_{i1} + \dots + \alpha_{in}^e u_{in} \quad ; i = 1..N_s \quad (2.231)$$

The superscript “e” refers to the fact that the Z,  $\beta$  and  $\alpha$  are equivalent values/functions, derived by combining two components. The probability of failure for all wind directions combined is then equal to:

$$P(Z < 0) = P\left(\bigcup_{i=1}^{N_s} Z_i^e < 0\right) \quad (2.232)$$

Note that the probability of occurrence of each wind direction,  $P(\phi=\phi_i)$ , is not explicitly mentioned in equation (2.232) because it is already incorporated in  $Z_i^e$ . Equation (2.232) is solved with the Hohenbichler method. To apply this method, the correlations between the Z-functions are required. The correlation between two functions  $Z_i^e$  and  $Z_j^e$  is derived as follows:

$$\rho(Z_i^e, Z_j^e) = \sum_{k=1}^n \alpha_{ik}^e \alpha_{jk}^e \rho_{ijk} \quad (2.233)$$

Where  $\alpha_{ik}^e$  is the equivalent  $\alpha$ -value of the  $k^{\text{th}}$  variable in the  $i^{\text{th}}$  wind sector and  $\rho_{ijk}$  is the (auto-)correlation between the  $k^{\text{th}}$  variable in the wind sector  $i$  and the  $k^{\text{th}}$  variable in the wind sector  $j$ . For the individual variables, autocorrelation coefficient  $\rho_{ijk}$  is chosen the same as before, i.e. equal to 1 for resistance variables and slowly evolving load variables and equal to 0 for fast evolving load variables.

To illustrate the whole procedure, consider the extreme case where the Z-function only depends on fast evolving variables. In that case the  $\rho$ -values of equations (2.223) and (2.233) are all equal to 0. Functions  $Z_i$  and  $Z_i'$  are modelled as a *series* system in equation (2.222), so a value of  $\rho=0$  means  $P(Z_i < 0)$  and  $P(Z_i' < 0)$  are multiplied:

$$P(Z_i < 0 \cap Z_i' < 0) = P(Z_i < 0) \cdot P(Z_i' < 0) \quad ; \rho(Z_i, Z_i') = 0 \quad (2.234)$$

Note that  $P(Z_i' < 0)$  represents the probability of occurrence of a wind direction,  $P(\phi=\phi_i)$ . Subsequently the wind directions are combined. Wind directions are modeled as a *parallel* system in equation (2.233), so a value of  $\rho=0$  means the failure probabilities of the individual wind directions are accumulated:

$$P(Z < 0) = \sum_{i=1}^{N_s} P(Z_i < 0) \quad ; \rho(Z_i, Z_j) = 0, i \neq j \quad (2.235)$$

So if the Z-function only depends on fast evolving variables, the procedure comes down to the numerical integration procedure as described in equation (2.218), which is the exact solution. So for the case where the Z-function only depends on fast evolving variables, the method as described in the current section provides the correct answer. Note that for each wind direction the failure probability is multiplied by  $P(\phi=\phi_i)$ , which means a reduction of the failure probability (i.e. an increase in  $\beta$ ). Subsequently, the probabilities of the wind directions are accumulated, which leads to an increase in the failure probability (i.e. a decrease in  $\beta$ ).

Now consider the other extreme case where the Z-function only depends on slowly evolving variables:

$$Z = R - a_1 X_1 - \dots - a_n X_n \quad (2.236)$$

Where R is the resistance and  $X_1 \dots X_n$  are slowly evolving variables. The Z-function is the same for each wind direction, because the slowly evolving random variables like river discharge are independent of the wind direction. Furthermore, the  $\rho$ -values of equations (2.223) and (2.233) are all equal to 1. A value of  $\rho=1$  means the combination of  $P(Z_i < 0 \cap Z_i' < 0)$  in equation (2.222) will be equal to  $P(Z_i < 0)$ :

$$P(Z_i < 0 \cap Z_i' < 0) = P(Z_i < 0) \quad ; \rho(Z_i, Z_i') = 1 \quad (2.237)$$

In other words: the probability of failure *given* the wind direction is the same as the probability of failure *AND* the occurrence of the wind direction. Subsequently, the wind directions are combined. A value of  $\rho=1$  means the failure probability of the combined wind directions is equal to the failure probability of the wind direction with the largest probability of failure:

$$P(Z < 0) = \max_i P(Z_i < 0) \quad ; \rho(Z_i, Z_j) = 1, i \neq j \quad (2.238)$$

Function  $Z_i$  is the same for each wind direction, see equation (2.236). This means that after executing the procedure, the original Z-function of equation (2.236) is derived. This shows the procedure provides the correct answer in case the Z-function only depends on slowly evolving random variables.

The examples above show that for the two “extreme” cases, the procedure will result in the correct failure probability. If the Z-function only depends on fast evolving variables, the procedure comes down to the numerical integration approach as described in (2.218). If the Z-function only depends on slowly evolving variables, the procedure results in a single Z-function that is the same as the original Z-function (equation (2.236)). If the Z-function depends on both slowly and fast evolving variables, these two groups are essentially split up and the above described procedures are executed.

### 2.5.3.3 Deriving the alpha value for the wind direction

As stated earlier, the main advantage of the approach as described in section 2.5.3.2 over the method as described in section 2.5.3.1 is that it enables the derivation of  $\alpha$ -values of the wind direction. First of all, a design point value for the wind direction needs to be selected. This is the wind direction ( $\varphi$ ), with the highest probability of failure, i.e. the lowest reliability index. Define  $\beta_\varphi$  as the weighted reliability index *given the wind direction* that was derived for the selected wind direction  $\varphi$ . Furthermore, define  $\beta_{com}$  as the reliability index that is derived after combining all the wind directions. The  $\alpha$ -value of the selected wind direction,  $\alpha_\varphi$ , is derived from the following equation:

$$\beta_\varphi = \beta_{com} \sqrt{1 - \alpha_\varphi^2} \Rightarrow \alpha_\varphi = \sqrt{1 - \left( \frac{\beta_\varphi}{\beta_{com}} \right)^2}, \quad (2.239)$$

This can be explained as follows. Consider a standard description of a normalized linear Z-function:

$$Z = b + a_1 u_1 + \dots + a_n u_n \quad (2.240)$$

As we know from section 2.2.5, this Z-function has a reliability index equal to  $\beta$ . The design point value of each variable  $u_i$  is equal to:  $u_{d,i} = -\alpha_i \beta$ . Now consider an alternative Z-function,  $Z'$ , in which variable  $u_1$  is taken equal to its value in the design point:

$$\begin{aligned} Z' &= b + a_1 (u_{d,1}) + a_2 u_2 + \dots + a_n u_n \\ &= b + a_1 (-a_1 \beta) + a_2 u_2 + \dots + a_n u_n \\ &= b - a_1^2 \beta + a_2 u_2 + \dots + a_n u_n \end{aligned} \quad (2.241)$$

The  $\alpha$ -values in equation (2.241) are taken the same as in equation (2.240). Equation (2.241) is not a normalized Z-function. This is because  $u_1$  is no random variable anymore, so the function  $Z'$  has  $\alpha_2, \dots, \alpha_n$  left as the  $\alpha$ -values. Since  $Z$  is a normalized function, we know that:

$$\sqrt{\sum_{i=1}^n a_i^2} = 1 \quad \text{and} \quad \sqrt{\sum_{i=2}^n a_i^2} = \sqrt{1 - a_1^2} \quad (2.242)$$

This means that equation (2.241) can be normalized by division by  $\sqrt{1 - a_1^2}$ . This results in a new Z-function  $Z''$ :

$$Z'' = \frac{b - a_1^2 \beta}{\sqrt{1 - a_1^2}} + a_2^i u_2 + \dots + a_n^i u_n \quad (2.243)$$

In which:

$$a_i^i = \frac{a_i}{\sqrt{1 - a_1^2}} ; i = 2..n \quad (2.244)$$

$Z''$  is a normalized Z-function with reliability index,  $\beta''$  equal to:

$$b'' = \frac{b - a_1^2 \beta}{\sqrt{1 - a_1^2}} = b \sqrt{1 - a_1^2} \quad (2.245)$$

So, if we set the first u-variable of the Z-function of equation (2.240) equal to its design point value, reliability index  $\beta$  reduces to reliability index  $\beta''$  as described in equation (2.245). The similarity between equations (2.245) and (2.239) is evident. Reliability index  $\beta_\phi$  in equation (2.239) is the reliability index that is obtained if the wind direction is set equal to its design point value, whereas  $\beta_{com}$  is the reliability index if all potential outcomes of the wind direction are taken into account. So if we replace  $\beta''$  and  $\beta$  in (2.245) by  $\beta_\phi$  and  $\beta_{com}$  we obtain the relation as described in (2.239). This shows that equation (2.239) provides the  $\alpha$ -value of the wind direction.

the  $\alpha$ -values for the other random variables involved are taken equal to the  $\alpha$ -values that were derived for wind direction  $\varphi$ , i.e. the wind direction with the highest failure probability. Note, however, that these values need to be multiplied by the a factor  $(\beta_\varphi/\beta_{com})^2$  to take into account the fact that the  $\alpha$ -value of the wind direction is included. This multiplication guarantees that the sum of the squares of the  $\alpha$ -values (including the  $\alpha$ -value of the wind direction) is equal to 1.

## 2.5.4 Closure scenarios for flood barriers

### 2.5.4.1 Basic method description

Flood barriers close off a river or estuary system when high sea water levels are predicted, with the objective to reduce high loads in tidal river systems. The ways in which these barriers operate depend on their design and the conditions under which they should be activated. An open or closed barrier can substantially affect the water level at flood defences (e.g. dikes) protected by it. This automatically means that the probabilities of failure of the flood defences behind the barrier are influenced by the operation of the barrier.

Closure scenarios are the different scenarios for the barrier that need to be taken into account when deriving the probability of failure of a flood defence that is protected by the barrier. These scenarios include, among others, the state of the barrier (open, closed, partially closed, etc.) and the operation rules of the barrier. For example, consider a barrier that is designed to close when the water level,  $h_b$ , at a specified location is greater than some critical water level,  $h_b^*$ , and assume (for the moment) that the barrier never fails to close upon request. The contribution of this scenario to the probability of failure at a flood defence that is protected by the barrier is equal to:

$$P[\text{failure flood defence with closed barrier}] = P(\text{failure flood defence} \cap (h_b > h_b^*)) \quad (2.246)$$

To explain the method used in Hydra-Ring to solve equation (2.246), Figure 2.39 presents a schematization of this scenario. This figure helps conceptualize the influence of a barrier on the failure probability of a flood defence at some location along the river or estuary. In this figure, the water level at the flood defence is assumed to be determined by the combination of:

- Water level at sea ( $h_{sea}$ )
- Upstream river discharge ( $Q$ )
- Closure situation of the barrier (open or closed).

Other influences like wind speed have been left out of Figure 2.39 for the sake of simplicity. In each panel of Figure 2.39, three lines are displayed:

- 1 a limit state function,  $Z_c(Q, h_{sea})=0$  in case the barrier is closed
- 2 a limit state function,  $Z_o(Q, h_{sea})=0$  in case the barrier is open
- 3 the closure criterion:  $h_b(Q, h_{sea})=h_b^*$ .

The first two lines are limit state functions that define for which combinations of  $Q$  and  $h_{sea}$  failure occurs *at the considered dike location for the considered failure mechanism*, in the situation when the barrier is closed or open respectively. The failure domain is to the upper right of these lines, which shows failure is more likely to occur in case the barrier is open.

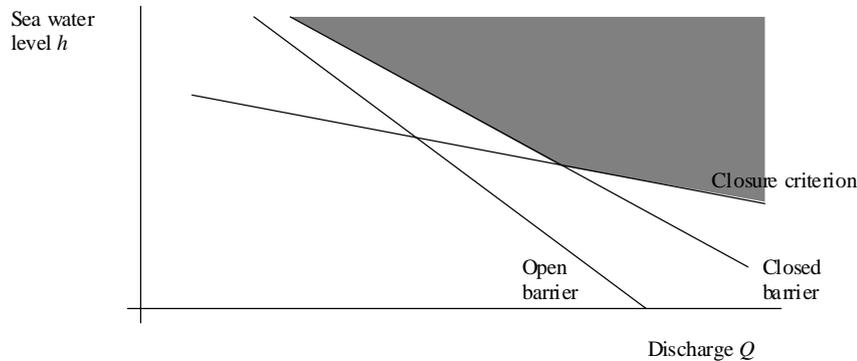


Figure 2.39 Schematic representation of the failure domain (of the flood defence) for the scenario of the closed barrier.

The 'closure criterion' (third line) does not represent a critical level for failure (Z-function) like the other two lines. However, for computational reasons it can be described as a Z-function:

$$Z_b = h_b^* - h_b \quad (2.247)$$

In which  $h_b$  is a function of random variables such as  $Q$  and  $h_{\text{sea}}$ . In some cases, the closure of the barrier is slightly complicated by the fact that predictions of water levels are used to determine whether the barrier should be closed. The predicted water level contains some error  $\varepsilon$ , with a user defined distribution. To take this prediction error into account, the Z-function for the barrier is described as:

$$Z_b = h_b^* - h_b(\varepsilon) \quad (2.248)$$

In which  $h_b(\varepsilon)$  is the water level as a function of the prediction error  $\varepsilon$ . Note that the three Z-functions,  $Z_o$ ,  $Z_c$  and  $Z_b$ , are presented in Figure 2.39 as straight lines. In reality they will be non-linear of variables like discharge ( $Q$ ) and sea water level ( $h$ ). In equation (2.248), function  $Z_b$  is a linear function of  $h_b$  (the local water level at the barrier) but  $h_b$  is a non-linear function of  $Q$  and  $h$ . Note further that the closure criterion line is the same for each dike location, while the other two lines are location-specific.

In Figure 2.39, the grey area represents the failure domain that is associated with the scenario of equation (2.246). The failure probability for this scenario can be described as follows:

$$P[\text{failure flood defence with closed barrier}] = P(Z_c < 0 \cap Z_b < 0) \quad (2.249)$$

In a similar manner, the failure probability for the scenario of the open barrier can be described:

$$P[\text{failure flood defence with open barrier}] = P(Z_o < 0 \cap Z_b \geq 0) \quad (2.250)$$

The failure domain for this scenario is displayed in Figure 2.40.

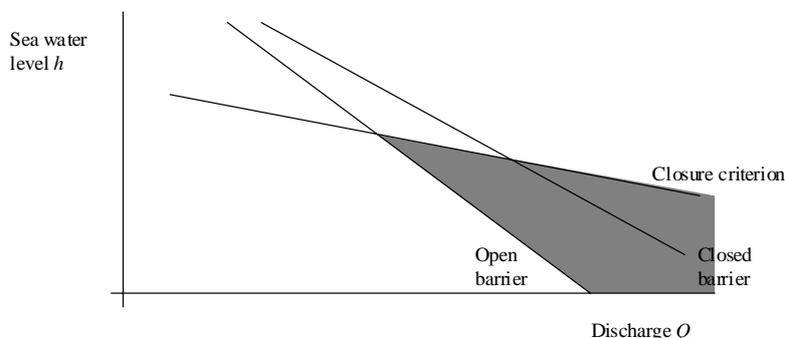


Figure 2.40 Schematic representation of the failure domain for the scenario of the open barrier.

Failure at a dike location will either occur with open or closed barrier. The failure probability for such a location is therefore a combination of equations (2.249) and (2.250):

$$P(Z < 0) = P[(Z_c < 0 \cap Z_b < 0) \cup (Z_o < 0 \cap Z_b \geq 0)] \tag{2.251}$$

So in order to determine failure probabilities for locations that are protected by a barrier, three Z-functions ( $Z_o$ ,  $Z_c$  and  $Z_b$ ) are combined. The Z-functions are described by databases of hydraulic model results in which the water level at each model output location is a function of multiple input variables. In case of the tidal area in the Netherlands, these variables are river discharge, sea water level, wind speed and wind direction. For this area, two databases are available, representing the cases of open and closed barrier respectively. Function  $Z_c$  is based on the database for the closed barrier, functions  $Z_o$  and  $Z_b$  are based on the database for the open barrier. The latter is explained by the fact that the moment at which the barrier should be closed is determined in the situation in which the barrier is still open. For function  $Z_b$ , only the hydraulic model results for the location that determines the closure criterion are relevant.

### 2.5.4.2 Potential failure of the barrier

In the previous section it was assumed that the barrier closes on each request. However, in reality the barrier may fail to do so, e.g. due to mechanical or human errors. This can be taken into account in Hydra-Ring. For a barrier with two closing situations (fully open or fully closed), there are four possible scenarios, described as follows:

Table 2.6 Description of the four possible closure scenarios for a barrier with two closing situations (open or closed)

Scenario	closure criterion situation	closure situation	Correctly open/closed
S <sub>1</sub>	$h_b \geq h_b^*$	Closed	yes
S <sub>2</sub>	$h_b \geq h_b^*$	Open	no
S <sub>3</sub>	$h_b < h_b^*$	Closed	no
S <sub>4</sub>	$h_b < h_b^*$	Open	yes

The probability of failure for a flood defence that is protected by the barrier can be described as:

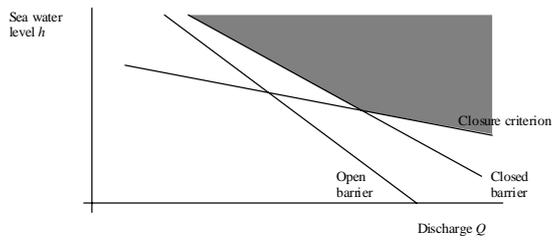
$$P(Z < 0) = \bigcup_{i=1}^4 P(Z < 0 \cap S_i) \tag{2.252}$$

Figure 2.41 shows the failure domain for the four scenarios (grey areas). Each of the four schematizations is explained as follows:

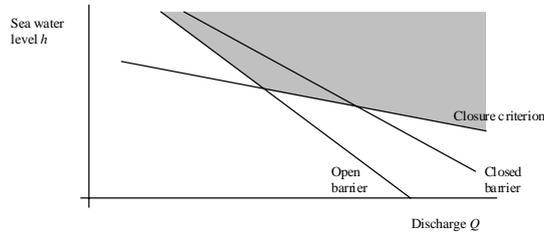
- |   |   |  |
|---|---|--|
| 1 | In scenario 1, the barrier is correctly closed:   | failure domain = $Z_c < 0 \cap Z_b < 0$ .    |
| 2 | In scenario 2, the barrier is incorrectly open:   | failure domain = $Z_o < 0 \cap Z_b < 0$ .    |
| 3 | In scenario 3, the barrier is incorrectly closed: | failure domain = $Z_c < 0 \cap Z_b \geq 0$ . |
| 4 | In scenario 4, the barrier is correctly open:     | failure domain = $Z_o < 0 \cap Z_b \geq 0$ . |

Substitution of these functions in equation (2.252) gives:

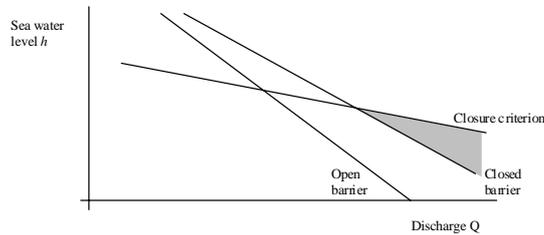
$$P(Z < 0) = P \left[ \begin{array}{l} (Z_c < 0 \cap Z_b < 0) \cup (Z_o < 0 \cap Z_b < 0) \cup \\ (Z_c < 0 \cap Z_b \geq 0) \cup (Z_o < 0 \cap Z_b \geq 0) \end{array} \right] \quad (2.253)$$



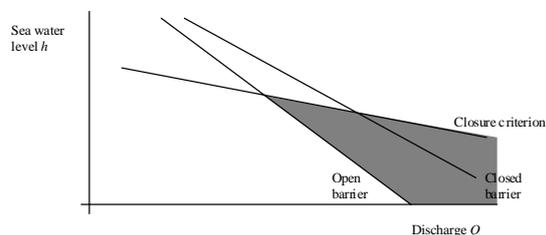
a) Failure domain, Scenario 1



b) Failure domain, Scenario 2



c) Failure domain, Scenario 3



d) Failure domain, Scenario 4

Figure 2.41 Schematic representation of the failure domain for the four closure scenarios

Consider the general failure probability formula for a single scenario (out of four scenarios). From basic probability theory, the following is true by definition:

$$P(A \cap B \cap C) = P((A \cap B) \cap C) = P(A \cap B) \cdot P(C | A \cap B) \quad (2.254)$$

The definitions of A, B and C differ for the four scenarios, see Table 2.7

Table 2.7 Description of events A, B, and C for the four closure scenarios described in Table 2.6

scenario	A	B	C
1	$Z_b < 0$	$Z_c < 0$	Barrier closed
2	$Z_b < 0$	$Z_o < 0$	Barrier open
3	$Z_b \geq 0$	$Z_c < 0$	Barrier closed
4	$Z_b \geq 0$	$Z_o < 0$	Barrier open

The probability of C is fully determined by the status of A, i.e. it is not influenced by the status of B. This means equation (2.254) simplifies to:

$$P(A \cap B \cap C) = P(A \cap B) \cdot P(C | A) \quad (2.255)$$

$P(A \cap B)$  is determined by combining the two related Z-functions, using e.g. the techniques as described in section 2.4.  $P(C|A)$  is an input table in Hydra-Ring, containing information about the probability of failure of the barrier per request for closure or opening (so no annual failure probability!). These probabilities can be derived e.g. through a fault tree analysis. An example of such an input table is presented in section 4.5.1.6.

### 2.5.4.3 Potential extensions of the method

The method as described in the previous section can also be applied on more complex systems. For instance the closure criterion may be related to the water level at more than just one location. This is for instance the case for the barrier in the Rhine/Meuse delta in the Netherlands. This barrier will be closed if the (predicted) water level at either Rotterdam or Dordrecht exceeds a critical level. This means a Z-function needs to be derived for each of these two locations and subsequently these two Z-functions need to be combined. In more general terms this can be formulated as follows:

$$Z_b = \bigcup_{i=1}^{N_L} Z_{bi} \quad (2.256)$$

In which  $Z_b$  is the combined Z-function for the closure criterion and  $Z_{bi}$  describes the closure criterion of location  $i$  (out of  $N_L$  locations).

Another potential extension is the option that a barrier may also 'partially' fail to close upon request, e.g. it is only half closed. This would lead to an increase of potential states of the barrier, additional to 'open' or 'closed'. For each state a separate hydraulic database is required that describes the local water level as a function of the random variables (discharge, sea water level etc.). Consequently, for each barrier state, a Z-function can be defined. These Z-functions need to be combined in a similar manner as described in the previous section to derive the total failure probability.

A third potential extension is the fact that a river delta may contain more than just one barrier. This would also lead to an increase in potential barrier states similar to the previous example. Basically, each combination of states of the various barriers in the system can be considered as one state of the entire barrier system. For each combination, a separate hydraulic database is required that describes the local water level as a function of the random variables, in order to formulate a Z-function for each combination. With multiple barriers and multiple potential states for each barrier this may lead to a large amount of combinations of barrier states and hence a large amount of required hydrodynamic model simulations.

#### 2.5.4.4 Trouble-shooting

As stated before in previous sections, techniques using linearizations of Z-functions rely on the fact that the error introduced by the linearisation is small. Unfortunately, this turned out not to be the case for a practical case in the Netherlands. More specifically: the linearisation of function  $Z_b$  that represents the closure criterion (equation (2.248)) introduced an error in the computation of the failure probability that is unacceptably large. This was (mainly) due to the inclusion of the prediction error,  $\varepsilon$ , in function  $Z_b$ . To solve this problem, this variable is dealt with through numerical integration. In this approach the following integral, based on the theorem of total probability, is solved:

$$P(Z_b < 0) = \int P(Z_b < 0 | \varepsilon) f(\varepsilon) d\varepsilon \quad (2.257)$$

Note that for  $P(Z_b \geq 0)$  a similar approach can be used. In the numerical integration approach, a numerical grid is defined to represent the domain of outcomes of  $\varepsilon$ . Assume for simplicity that the grid is equidistant with step size  $\Delta\varepsilon$ . Then, the integral of equation (2.257) is approximated by:

$$P(Z_b < 0) = \sum_{j=1}^N P(Z_b < 0 | \varepsilon_j) f_\varepsilon(\varepsilon_j) \Delta\varepsilon \quad (2.258)$$

Where  $N$  is the number of grid cells. Now, define:

$$Z_j = (Z_b | \varepsilon_j) \quad (2.259)$$

Then:

$$P(Z_b < 0) = \sum_{j=1}^N P(Z_j < 0) f_\varepsilon(\varepsilon_j) \Delta\varepsilon \quad (2.260)$$

Substitution of equation (2.260) in equation (2.253) provides the failure probability for the 4 scenarios of Table 2.7 combined:

$$P(Z < 0) = P \left[ \begin{array}{l} \sum_{j=1}^N P(Z_j < 0 \cap Z_c < 0) f_\varepsilon(\varepsilon_j) \Delta\varepsilon \cup \\ \sum_{j=1}^N P(Z_j < 0 \cap Z_o < 0) f_\varepsilon(\varepsilon_j) \Delta\varepsilon \cup \\ \sum_{j=1}^N P(Z_j \geq 0 \cap Z_c < 0) f_\varepsilon(\varepsilon_j) \Delta\varepsilon \cup \\ \sum_{j=1}^N P(Z_j \geq 0 \cap Z_o < 0) f_\varepsilon(\varepsilon_j) \Delta\varepsilon \end{array} \right] \quad (2.261)$$

The approach in Hydra-Ring is to derive a linearised version of each function  $Z_j$ ,  $j=1..N$ . For this purpose a failure computation for a single component (see section 2.3) is executed  $N$  times, each time for a different constant value of  $\varepsilon_j$ . The result of this procedure consists of  $N$  sets of  $\beta$  and  $\alpha$ -values, that are used subsequently to solve equation (2.261) with combining techniques such as the Hohenbichler method of section 2.4.2. Note that the set of  $\beta$  and  $\alpha$ -values for  $Z_j < 0$  are the same as the set of  $\beta$  and  $\alpha$ -values for  $Z_j \geq 0$ , multiplied by  $-1$ .

#### 2.5.4.5 Deriving the alpha value for the prediction error

Since prediction error  $\varepsilon$  is dealt with through numerical integration, no  $\alpha$ -value is provided for this variable at first. The  $\alpha$ -value of  $\varepsilon$  therefore needs to be derived in similar fashion as for the wind direction, see section 2.5.3.3. This means first of all that a design point value for  $\varepsilon$  needs to be selected. This is the value of the prediction error,  $\varepsilon^*$ , for which the failure probability  $P(Z_j < 0 \cap Z_c < 0)$  is highest. Define  $\beta_{\varepsilon^*}$  as the associated reliability index and define  $\beta_{com}$  as the reliability index that is derived after combining all values of  $\varepsilon$  through numerical integration. The  $\alpha$ -value of the prediction error,  $\alpha_\varepsilon$ , is derived with the following equation:

$$\alpha_\varepsilon = \sqrt{1 - \left( \frac{\beta_{\varepsilon^*}}{\beta_{com}} \right)^2} \quad (2.262)$$

The reader is referred to section 2.5.3.3 for an explanation of this equation. Similar to section 2.5.3.3, the  $\alpha$ -values for the other random variables involved are taken equal to the  $\alpha$ -values that were derived for  $\varepsilon^*$ , i.e. the prediction error with the highest failure probability. Note, however, that these values need to be multiplied by the a factor  $(\beta_{\varepsilon^*}/\beta_{com})^2$  to take into account the fact that the  $\alpha$ -value of the prediction error is included. This multiplication guarantees that the sum of the squares of the  $\alpha$ -values (including the  $\alpha$ -value of the prediction errors) is equal to 1.

An alternative method to derive the  $\alpha$ -value of  $\varepsilon$  is to base it directly on the prediction error,  $\varepsilon^*$ , for which the failure probability  $P(Z_j < 0 \cap Z_c < 0)$  is highest. The corresponding u-value of  $\varepsilon^*$ ,  $u^*$ , is

$$\alpha_{\varepsilon} = -\frac{\beta_{com}}{u^*} \quad (2.263)$$

This method is applicable for the prediction error because the  $u$ -value of  $\varepsilon$  has a real meaning. For wind direction, this method can not be applied because wind direction is a 'cyclic' variable, which means the  $u$ -value has no meaning.

#### 2.5.4.6 Pre-processing

Since  $Z_b$  describes the behaviour of the tidal barrier, the linearisation of this function is the same for each flood defence segment that is protected by the barrier. The procedure therefore needs to be repeated each time a failure probability of a flood defence segment is computed that is protected by the barrier. To reduce computation time, the linearization process as described in section 2.5.4.4 is executed once as preprocessing and the resulting sets of  $\beta$  and  $\alpha$ -values that describe the linearised  $Z$ -functions are stored in an input file for Hydra-Ring. For this purpose a special pre-processing procedure is implemented. In this pre-processing procedure, the required values for  $\alpha$  and  $\beta$  are determined for  $Z_b$ . This is done by executing one of the computation techniques of section 2.3. In this procedure, the following probability that  $Z_b < 0$  is computed. The resulting values of  $\alpha$  and  $\beta$  provide the required description of the linearization of  $Z_b$ . Note that in principle the same procedure needs to be executed for  $Z_b \geq 0$ . However, the resulting  $\alpha$ -values and  $\beta$  for this case will be -1 times the  $\alpha$ -values and  $\beta$  for  $Z_b < 0$ . Therefore, the procedure will not be executed twice to save computation time.

The procedure needs to be executed for each combination of:

- wind direction;
- grid value of  $\varepsilon$ , i.e. the water level prediction error (only in case of function  $Z_{bo}$ );
- temporal upscaling method: FBC, APT and NTI;

With regard to the temporal upscaling methods, it is relevant to note that failure computations with FBC in Hydra-Ring will be done with block durations (see section 2.5.2.3) that can be different for different dike sections or mechanisms. To facilitate this, the preprocessing procedure should also be executed for different block durations. Similarly, for application in the NTI method, the preprocessing should be applied for each phase of the standardized hydrograph (see section 2.5.2.2). However, these additional computations are left out of the procedure as they will not influence the computed failure probability significantly. For FBC, this means a representative block duration is selected whereas for NTI the computation is executed for the peak of the standardized hydrograph.

### 2.5.5 Spatial upscaling - from cross section to flood defence segment

#### 2.5.5.1 Computing the failure probability

The spatial upscaling technique as described in the current section is done over homogeneous reaches of flood defense. Homogenous in this case means the statistical characteristics remain constant. It is therefore relevant that the flood defence system is divided into segments for which the assumption of homogeneity is valid. So, if a dike segment is inhomogeneous, it needs to be split up into smaller, homogenous, segments. Note: upscaling over different segments to derive the probability of failure of the entire flood defence system is described in section 2.5.6.

Spatial upscaling is subject to a concept known as the length effect. The length effect essentially has to do with the increase in failure probability when going from a cross-section to a longitudinal segment and from a single segment to a flood defense system (interconnected segments). That is, the length effect refers to the effect that an increase in length has on the probability of failure. Note that this effect is also present when upscaling over time; the failure probability will increase as the considered time period increases.

The mathematical description of the length effect is the ratio of the failure probability of the larger length to that of the shorter. For the upscaling from cross-section to longitudinal segment (assuming statistical homogeneity!) this would be as follows:

$$\text{Length effect} = \frac{P_{f,segment}}{P_{f,crossSection}} \tag{2.264}$$

where  $P_{f,segment}$  refers to the failure probability of the longitudinal segment and  $P_{f,crossSection}$  refers to the failure probability of the cross section within that longitudinal segment. To derive the ratio of equation (2.264), a notion of the spatial correlation within the segment is required, for each random variable,  $X$ , involved. In Hydra-Ring this correlation is described with the following model:

$$\rho(\Delta y) = \rho_x + (1 - \rho_x) \exp\left[-\left(\frac{\Delta y^2}{d_x^2}\right)\right] \tag{2.265}$$

Where  $\rho$  is the correlation between two locations within the segment,  $\Delta y$  is the distance between these two locations,  $\rho_x$  is the residual correlation length of variable  $X$  and  $d_x$  is the spatial correlation length of variable  $X$ . Parameter  $d_x$  determines how quickly the correlation of variable  $X$  decreases over distance and  $\rho_x$  is the minimum correlation of variable  $X$  between two locations of the same (homogeneous) segment. The parameters  $d_x$  and  $\rho_x$  need to be determined for each variable  $X$ , based on a combination of measurements and expert judgement.

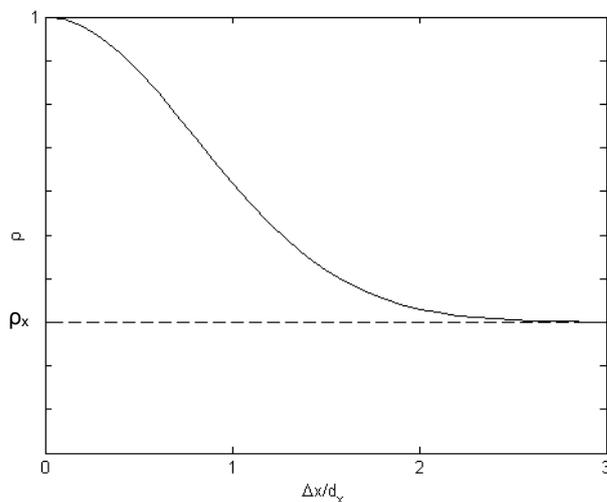


Figure 2.42 Autocorrelation function, correlation within a dike section

The correlation model of equation (2.265) and Figure 2.42 in principle is applied for each strength variable (load variables can generally be assumed to have correlation 1 within a single segment). This results in a similar model for the Z-function, i.e. in values  $d_z$  and  $\rho_z$ :

$$\rho(\Delta y) \approx \rho_z + (1 - \rho_z) \exp \left[ - \left( \frac{\Delta y^2}{d_z^2} \right) \right] \quad (2.266)$$

The parameters  $d_z$  and  $\rho_z$  can be derived as follows:

$$\rho_z = \sum_{i=1}^n \alpha_i^2 \rho_i \quad (2.267)$$

$$\frac{1}{d_z^2} = \frac{1}{1 - \rho_z} \sum_{i=1}^n \alpha_i^2 (1 - \rho_i) \frac{1}{d_i^2} \quad (2.268)$$

In which:

- $d_i$  = correlation length of random variable  $i$
- $\rho_i$  = residual correlation length of random variable  $i$
- $\alpha_i$  = influence coefficient of random variable  $i$

Note that:

- 1 Coefficients  $\alpha_1, \dots, \alpha_n$  are determined in the probabilistic computation for a “representative” cross-section within the flood defence segment. For this purpose the probabilistic techniques for a single component are used (see section 2.3).
- 2 Function (2.266) is an approximation of the correlation function for Z.

To derive the probability of failure of a dike segment, the segment is divided into components of equal length  $\Delta L$ . The number of components is equal to:

$$n_e = \frac{L}{\Delta L} \quad (2.269)$$

Where  $L$  is the length of the dike segment. The probability of failure for the entire dike segment is then equal to:

$$P_{F,segment} \approx (1 + n_e) P_{F,cross-section} = \left( 1 + \frac{L}{\Delta L} \right) P_{F,cross-section} \quad (2.270)$$

This means the continuous process, in which failure can occur at any location along the dike is now replaced by a discrete process in which the dike segment is composed of a finite number of components, each of which has a failure probability that is equal to the probability of failure of a cross-section. This simplification/approximation is only valid for a well selected value of  $\Delta L$ . If we assume that the spatial variation of Z is a Gaussian ergodic process (i.e.  $\rho_z=0$ ), the length  $\Delta L$  should be taken equal to:

$$\Delta L = d_z \sqrt{\pi} / \beta \quad ; \text{if } \rho_z = 0 \quad (2.271)$$

Where  $\beta$  is the reliability index as derived in the probabilistic computation for a cross-section (see section 2.3). The value of  $\Delta L$  is a result of the outcrossing approach (see section 2.4.5.3) in which the spatial variation of  $Z$  is assumed to be a Gaussian ergodic process. The derivation of  $\Delta L$ , as described in equation (2.271), is described in Jongejan [2012].

With the assumption of a Gaussian ergodic process, the failure probability of a dike segment of length  $L$  is approximately equal to (combine equations (2.270) and (2.271)):

$$P_{F,segment} \approx \left( 1 + \frac{L\beta}{d_z \sqrt{\pi}} \right) \Phi(-\beta) \quad ; \text{if } \rho_z = 0 \quad (2.272)$$

If  $\rho_z > 0$ , the assumption of a Gaussian ergodic process does not hold and an alternative solution is required. In that case,  $\rho_z > 0$  represents the part of the correlation function that does not contribute to the length effect, because it is the correlation that persists over the entire dike segment. In that case the  $Z$ -function is split in an ergodic part (with  $\rho$  approaching zero over long distances) and a non ergodic part (with  $\rho$  constant):

$$Z = \beta - v\sqrt{\rho} - u\sqrt{1-\rho} \quad (2.273)$$

Where  $v$  is the non-ergodic constant and  $u$  is the ergodic stochastic process with:

$$\rho(\Delta y) = \exp \left[ - \left( \frac{\Delta y^2}{d_z^2} \right) \right], \quad (2.274)$$

Where  $\rho$  is the correlation between two locations within the segment and  $\Delta y$  is the distance between two locations. Using the theorem of total probability, the failure probability of the flood defence segment can be described as follows:

$$P[Z < 0] = \int P[Z < 0 | v] f_v(v) dv, \quad (2.275)$$

Where  $f_v(v)$  is the standard normal density function. The conditional failure probability,  $P[Z < 0 | v]$ , in equation (2.275) can be written as (see Jongejan, 2012):

$$P[Z < 0 | v] = 1 - (1 - P(Z_{cross} < 0)) e^{-N_f} \quad (2.276)$$

$$N_f = \frac{L}{2\pi} e^{-\frac{\beta^{*2}}{2}} \frac{\sqrt{2}}{d_z}$$

$$\beta^* = \frac{\beta_{cross} - v\sqrt{\rho_z}}{\sqrt{1-\rho_z}}$$

Where  $Z_{cross}$  and  $\beta_{cross}$  are the  $Z$ -function and reliability index of the cross section and  $\Phi$  is the standard normal distribution function. The combination of equations (2.275) and (2.276)

provide the probability of failure for a flood defence segment. More details on the derivation of equations (2.275) and (2.276) can be found in Jongejan [2012].

Note: In formula 2.276 the width of the mechanism is not taken into account. In the current version of Hydra-Ring the width of the mechanism is taken equal to  $\Delta L$ . In the formula for  $N$ , the length  $L$  is replaced by  $L - \Delta L$ . The idea behind this “correction” is that for stretches smaller than  $\Delta L$  it is not possible to have an increase in failure probability as a result of the length effect.

In earlier versions of PC-Ring, the predecessor of Hydra-Ring, the following approximation for equations (2.275) and (2.276) was implemented to save computation time:

$$P[Z < 0] = \left( 1 + \frac{L\beta\sqrt{1-\rho_z}}{d_z\sqrt{\pi}} \right) \Phi(-\beta), \quad (2.277)$$

This approximation is only valid for small values of  $\rho_z$ . With the current day computation power, equation (2.276) can be evaluated in a split second, so it is recommended not to use the approximation as described with equation (2.277).

Note: Formula 2.277 is only valid for value of  $\rho_z > 0$ . For values of  $\rho_z \leq 0$  in Hydra-Ring the Hohenbichler method together with the outcrossing approach is used.

#### 2.5.5.2 Computing equivalent alpha-values

As stated in the previous section, the flood defence segment can be thought of to consist of identical components of  $N$  identical components of length  $\Delta L$ . Upscaling to a dike section in essence is therefore the same as upscaling over  $N$  identical components. The last step in such an upscaling process, is the derivation of new equivalent  $\alpha$ -values for the individual random variables, see section 2.4.3.2. The first step in this method is to determine the  $\alpha$ -value of the correlated part of the  $Z$ -function of equation (2.273), i.e. variable  $v$ . This is done in the standard way by perturbing the mean value of  $v$  with a small value  $\varepsilon$  and quantifying the effect on the computed  $\beta$ -value of the dike section of a small perturbation ( $\varepsilon$ ) in the mean value of  $v$ . The  $\alpha$ -value of  $v$  is thus equal to

$$\alpha_v = \frac{\partial \beta_{\text{section}}}{\partial \bar{v}} \quad (2.278)$$

Equation (2.177) states that the equivalent value,  $\alpha_k^e$ , of variable  $k$  can then be derived as follows:

$$\alpha_k^e = \sqrt{1-\alpha_v^2} \frac{\alpha_k \sqrt{1-\rho_k}}{\sqrt{1-\rho}} + \alpha_v \frac{\alpha_k \sqrt{\rho_k}}{\sqrt{\rho}} \quad (2.279)$$

In which  $\alpha_k$  is the  $\alpha$ -value of variable  $k$  before upscaling and  $\rho_k$  is the correlation of variable  $k$  between two components. Since components in this case have length  $\Delta L$ , this correlation is equal to (see equation (2.265):

$$\rho_k = \rho_{kr} + (1 - \rho_{kr}) \cdot \exp\left[-\left(\frac{\Delta L}{d_k}\right)^2\right] \quad (2.280)$$

Where,  $\rho_{kr}$  is the residual correlation length of variable  $k$  and  $d_x$  is the spatial correlation length of variable  $k$ . For load variables,  $\rho_k$  can be assumed to be equal to 1, which simplifies equation (2.279) to:

$$\alpha_k^e = \alpha_v \frac{\alpha_k}{\sqrt{\rho}}; \text{ load variables} \quad (2.281)$$

## 2.5.6 Combination order over mechanisms and flood defense segments

In Hydra-Ring, the failure probability is computed per flood defence segment (e.g. dike segment) and per failure mechanism. The method of Hohenbichler (see section 2.4.2) is used for this purpose. This section describes the order of combining failure probabilities over failure mechanisms and segments.

Consider an example with two failure mechanisms and  $m$  dike segments, in which the failure probability for each segment and mechanism has already been computed. The  $Z$ -functions per segment and mechanism can be given in terms of their  $\alpha$  and  $\beta$  values.

Table 2.8  $Z$  functions per segment and failure mechanism written in terms of their  $\alpha$  and  $\beta$  values.

Segment	Failure mechanism 1	Failure mechanism 2
1	$Z_{11} = \beta_{11} - \alpha_{111}u_{11} - \dots - \alpha_{11n}u_{1n}$	$Z_{21} = \beta_{21} - \alpha_{211}u_{11} - \dots - \alpha_{21n}u_{1n}$
2	$Z_{12} = \beta_{12} - \alpha_{121}u_{21} - \dots - \alpha_{12n}u_{2n}$	$Z_{22} = \beta_{22} - \alpha_{221}u_{21} - \dots - \alpha_{22n}u_{2n}$
...		
$m$	$Z_{1m} = \beta_{1m} - \alpha_{1m1}u_{m1} - \dots - \alpha_{1mn}u_{mn}$	$Z_{2m} = \beta_{2m} - \alpha_{2m1}u_{m1} - \dots - \alpha_{2mn}u_{mn}$

In Table 2.8,  $Z_{ij}$  represents the  $Z$ -function for failure mechanism  $i$  and segment  $j$ . The variable  $u_{jk}$  represents the  $k^{\text{th}}$  random variable in the  $j^{\text{th}}$  segment and  $\alpha_{ijk}$  is the  $\alpha$ -variable for the  $k^{\text{th}}$  random variable in the  $j^{\text{th}}$  segment and the  $i^{\text{th}}$  failure mechanism.

There are two ways in which the combination can be carried out:

- 1) First combine the failure probability over the segments (per failure mechanism), and then combine over the failure mechanisms.
- 2) First combine the failure probability over the failure mechanisms (per segment), and then combine over the segments.

It turns out that application of the first option leads to a practical problem and therefore option (2) is programmed in Hydra-Ring. To explain this issue, both options will be discussed and the difficulty with option (1) will be explained.

### 2.5.6.1 Option 1: First combining over segments, then over mechanisms

The combination order described by option (1) first combines the failure probability over the segments per failure mechanism. This step results in the following:

Table 2.9 Z functions per failure mechanism of the combined segment

Segment	Failure mechanism 1	Failure mechanism 2
combined (1-m)	$Z_1^e = P(Z_{11} < 0 \cup Z_{12} < 0 \cup \dots \cup Z_{1m} < 0)$	$Z_2^e = P(Z_{21} < 0 \cup Z_{22} < 0 \cup \dots \cup Z_{2m} < 0)$

The variables  $Z_1^e$  and  $Z_2^e$  in Table 2.9 represent the equivalent Z functions for the two failure mechanisms, combined over the  $m$  segments.  $Z_1^e$  and  $Z_2^e$  can also be written in terms of their respective  $\alpha$  and  $\beta$  values:

$$Z_1^e = \beta_1^e + \alpha_{11}^e u_{11} + \dots + \alpha_{1n}^e u_{1n} \quad (2.282)$$

$$Z_2^e = \beta_2^e + \alpha_{21}^e u_{21} + \dots + \alpha_{2n}^e u_{2n}$$

In equations (2.282),  $\alpha_{ik}^e$  represents the equivalent  $\alpha$ -value for the  $i^{\text{th}}$  failure mechanism and the  $k^{\text{th}}$  random variable. Similarly,  $u_{ik}$  refers to the  $u$ -value of the  $i^{\text{th}}$  failure mechanism and the  $k^{\text{th}}$  random variable.

To combine over the failure mechanisms, using the method of Hohenbichler, the correlation coefficient between the two Z-functions needs to be calculated according to formula (2.119). Within that formula, the correlation coefficients between the random variables in equation (2.282) are required, i.e.  $\rho(u_{1k}, u_{2k})$ ,  $k=1..n$ . The problem is that once the Z-functions have been combined over all the segments, it is no longer known what the correlation is between random variables for the different failure mechanisms. Take for example the wave height. Within a segment, the correlation between the wave height for failure mechanism 1 and failure mechanism 2 is equal to 1, because it concerns the same location. Once the segments have been combined it is no longer known how the wave height in failure mechanism 1 is correlated with the wave height in failure mechanism 2. It is for this reason that the combination takes place first over the failure mechanisms (where the correlation between variables is 1 within a segment), and then over the segments, where the spatial correlation between the random variables is known.

### 2.5.6.2 Option 2: First combining over mechanisms, then over segments

The combination order described by option (2) first combines the failure probability over the failure mechanisms per segment. This step results in the following:

Table 2.10 Z functions per segment of the combined failure mechanism

Segment	Combined failure mechanism
1	$Z_1^e = P(Z_{11} < 0 \cup Z_{21} < 0)$
2	$Z_2^e = P(Z_{12} < 0 \cup Z_{22} < 0)$

...	
$m$	$Z_m^e = P(Z_{1m} < 0 \cup Z_{2m} < 0)$

The variables  $Z_1^e, Z_2^e, \dots, Z_m^e$  in Table 2.9 represent the equivalent Z-functions for the  $m$  segments, combined over the two failure mechanisms. The functions  $Z_1^e, Z_2^e, \dots, Z_m^e$  can also be written in terms of their  $\alpha$  and  $\beta$  values:

$$Z_j^e = \beta_j^e + \alpha_{j1}^e u_{j1} + \dots + \alpha_{jn}^e u_{jn} \quad ; j=1..m \quad (2.283)$$

In equation (2.283),  $\alpha_{jk}^e$  represents the equivalent  $\alpha$ -value associated with the  $j^{\text{th}}$  segment and the  $k^{\text{th}}$  random variable. The Z-functions of equation (2.283) are obtained by combining over the failure mechanisms, using the method of Hohenbichler. For this purpose, the correlation coefficient between the Z-functions of the failure mechanisms are needed. The correlation coefficient between the random variables in one failure mechanism and the random variables in the other failure mechanism is equal to 1 within a segment, i.e.  $\rho(u_{1k}, u_{1k})=1, k=1..n$ .

To combine over the segments, the correlation coefficient between Z-functions again needs to be calculated according to formula (2.119). Within that formula, the correlation coefficient between the random variables in one segment and the random variables in another segment are needed. This correlation coefficient is simply the spatial correlation between the random variables, which can be computed from the residual correlation and the correlation length, see equation (2.265). The residual correlation and the correlation length are given as input. For details on computing the correlation coefficient from the residual correlation and the correlation length, the reader is referred to section 2.5.5.

Note: In option 2 the correlation between stochastic variables is more realistic taken into account. For example: For the mechanisms within one section some stochastic variables are equal and used in more than one mechanism. An example is the dike height. For these variables the correlation should be set to 1.

In the case of option 1 in the combination of the mechanisms it is not sure the dike height is in the same section. Therefore the correlation should be set to 0 for these variables.

### 2.5.7 Order of combining: wind directions

In addition to combining over segments and failure mechanisms, failure probabilities are also combined over wind directions. The computation structures for computing failure probabilities in Hydra-Ring are carried out per failure mechanism, per segment and per wind direction. The wind directions are combined in “the inner loop”. This means that by default the order of combinations occurs first over wind directions, prior to the combination over the dike segments and failure mechanisms. The reason to combine the wind directions before combining over segment is the same as why the mechanisms are combined before the segments are combined (see previous section). However, there is no particular reason why the wind direction is combined before the mechanisms are combined, i.e. the other way around would also have been possible.

## 2.6 Procedure for computing the failure probability of a flood defence system

### 2.6.1 Introduction

Section 2.4 and 2.5 described the various combining techniques for system reliability that are implemented in Hydra-Ring. The current section presents an overview of the whole procedure that is applied in Hydra-Ring to derive the failure probability of a flood defence system. This means the interaction between the various combining techniques is explained. Furthermore, some 'finer details' of the various methods will be explained that have been left out of the description so far.

The procedure is explained for the three different alternatives for upscaling probabilities in time: FBC, NTI and APT (see section 2.5.2). The choices and assumptions in these methods are such that they strongly influence the procedure for computing failure probabilities for flood defence systems. The user can select which method is to be used for upscaling in time. The choice of the method determines the order in which the subroutines are executed.

### 2.6.2 Procedure for FBC

The procedure for computing the failure probability of a flood defence system, using the FBC procedure for upscaling in time, can be summarized as follows:

1. Determine the failure probability of the smallest component, i.e. a combination of one cross section, one sub-mechanism one wind direction, one closure situation and one basic time step.
2. Combine the failure probabilities of the sub-mechanisms to one failure probability per failure mechanism.
3. Repeat steps 1 and 2 iteratively to derive the appropriate block duration for each load variable.
4. Combine failure probabilities of the closure situations with the possible closure scenarios.
5. Upscale the failure probability of a cross section to a dike section.
6. Combine the failure probability for each wind direction with the probability of occurrence of that wind direction.
7. Combine the failure probabilities of all wind directions.
8. Derive the influencing factor of the wind direction.
9. Upscale the failure probabilities to a year.
10. Combine failure probabilities of all failure mechanisms.
11. Combine failure probabilities of all segments.

The remainder of this section elaborates on each of these steps.

#### 2.6.2.1 Step 1: Failure probability per sub-mechanism

In this step, the failure probability of the smallest component is computed, i.e. a combination of one cross section, one sub-mechanism one wind direction, one closure situation and one basic time step. For this purpose the computation techniques for the failure probability for single components are used (see section 2.2). The choice of the method is user-defined.

#### 2.6.2.2 Step 2: Combine failure probabilities of the submechanisms

For failure mechanisms that consist of multiple submechanims, step 1 has to be applied for each submechanism. An example of such a failure mechanism is 'piping and heave', which only occurs if the two submechanisms 'piping' and 'heave' both occur (at the same time at the same cross section). Subsequently, the failure probabilities of the submechanims need to be combined, with one of the combination techniques of section 2.4. For failure mechanisms that

do not consist of multiple submechanisms, the result of step 2 is the same as the result of step 1.

#### 2.6.2.3 *Step 3: Determine the block duration*

The block duration is essential in the FBC procedure (see section 2.5.2.3). This duration needs to be determined for each load variable. For fast evolving random variables like wind speed and sea water level, the block duration is equal to the basic time step (generally one tidal period). For slowly evolving variables like river discharge or lake level, the block duration depends on the threshold value under consideration. Within a block, the represented (slowly varying) load variable is assumed to be fully autocorrelated and between blocks there is constant correlation (in Hydra-Ring often assumed to be zero).

The block duration of variable  $Q$  is derived from a duration curve:  $N(q) =$  the average duration of a single exceedance of threshold  $q$ . By definition,  $N$  is a non-increasing function of  $q$ . This function is input of Hydra-Ring. The block duration for variable  $Q$  is taken to be equal to  $N(q_d)$ , where  $q_d$  is the derived value of  $Q$  in the design point. The problem, however, is that the block duration is required to determine the design point (to be explained below). Therefore, the block duration and the design point need to be derived in an iterative manner. The procedure starts with an initial estimate of the block duration. Subsequently, steps 1 and 2 of the previous subsections are executed, which results in a designpoint. From the design point a new block duration  $N(q_d)$  is determined. This procedure is repeated until the block duration has converged. Generally, this only takes a few steps. The result of the procedure is a combination of a block duration and design point that are mutually consistent.

As stated before, the design point that is determined in steps 1 and 2 depends on the block duration. This has to do with the fact that the statistics of variable  $Q$  are “rescaled” in such a way that they are valid on the time scale of the block duration. The reason for this is that in the FBC approach the block duration of variable  $Q$  is the basic time scale for variable  $Q$ . Each block can be considered as a separate sample of variable  $Q$ . Therefore, the statistics that are used for variable  $Q$  need to be applied on the time scale of the block period. Generally, statistics are available on the annual time scale, which is why the rescaling procedure is required. The details of the rescaling procedure will be explained later on, in section 3.3.3, but it is immediately evident that the considered time scale will have an influence on the statistics that are applied. The probability of exceedance of a certain threshold level  $q$  will be different if we consider e.g. a period of one year than if we consider just one day. The probability will increase (formally: not decrease) if the considered time scale increases. This shows that the selected block duration for variable  $Q$  influences the probabilities of exceedance of threshold values of  $Q$ .

#### 2.6.2.4 *Step 4: Combine the failure probability for closure scenarios*

In areas where flood defences are protected by a storm surge barrier, the state of the barrier influences the water level and consequently the probability of failure of the flood defences. The barrier can be in various states (open, closed, partially closed etc.) and the barrier may fail to/open/close upon request. This results in several scenarios for a flood barrier (correctly closed, incorrectly open etc) as described in section 2.5.4. Steps 1-3 need to be executed separately for each barrier scenario and subsequently the failure probabilities for each scenario need to be combined using e.g. the combination techniques of section 2.4.

Before the scenarios can be combined, first a temporal upscaling step is required. The reason is that for the slowly evolving load variables like river discharge and lake level a block duration has been determined in step 3. This block duration is determined separately for each

barrier scenario and therefore most likely will be different for each scenario. This means the scenarios cannot be combined directly because the components representing the scenarios represent different time scales. In order to overcome this discrepancy, the components are upscaled to the largest block duration over all the considered scenarios. This step is schematically depicted in Figure 2.43.

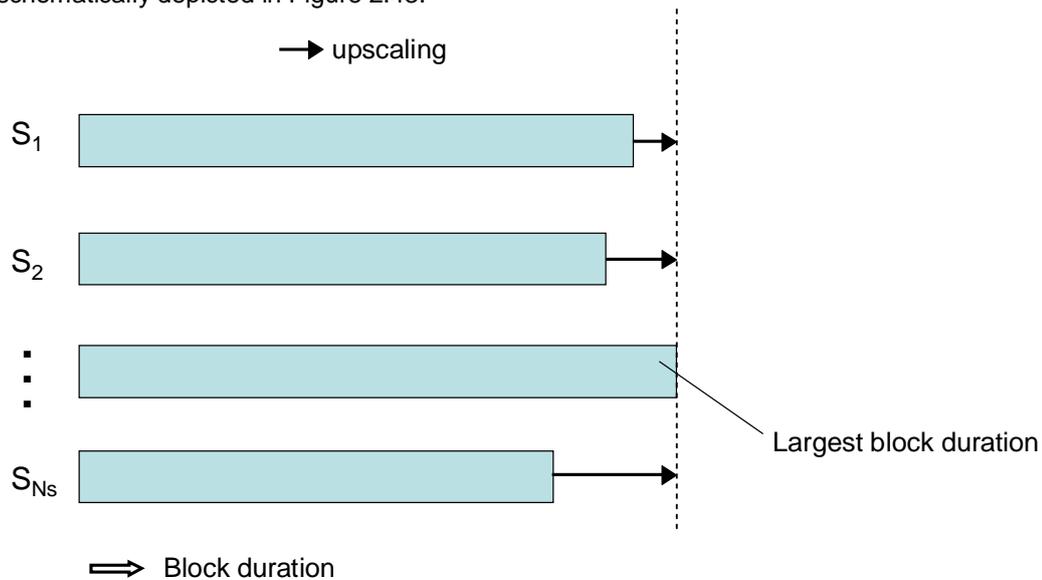


Figure 2.43 Schematic view of upscaling the block duration to the largest block duration of all  $N_s$  barrier scenarios. This step is executed for each random load variable.

This step is executed for each random variable. In other words; for each random variable the scenario with the largest block duration is determined and the components that represent the other scenarios are upscaled in order to guarantee that the block duration for this random variable is the same for all scenarios. The upscaling for 1 scenario is carried out as follows:

Consider variable  $Q$  with a block duration  $T_i$  for scenario  $i$ . This variable needs to be upscaled to the maximum block duration  $T_m$ , with  $T_i \leq T_m$ . This means the component representing scenario  $i$  needs to be upscaled with a factor  $N = T_m/T_i$ . In the FBC model, variable  $Q$  for scenario  $i$  is fully autocorrelated within the block of time scale  $T_i$  whereas there is no autocorrelation over larger time scales. In the upscaling from time scale  $T_i$  to  $T_m$ , the autocorrelation of variable  $Q$  is therefore set equal to 0. Since only variable  $Q$  is considered (the other variables are upscaled in subsequent steps) the autocorrelation of all the other variables is set equal to 1. This means the other variables have no influence on the failure probability in the upscaling process. The autocorrelation of the component that is scaled up in time, is therefore equal to:

$$\rho_t = \alpha_q \cdot 0 + \sqrt{1 - \alpha_q^2} \cdot 1 = \sqrt{1 - \alpha_q^2} \quad (2.284)$$

In which  $\alpha_q$  is the derived  $\alpha$ -value of random variable  $Q$ . So the upscaling procedure for random variable  $q$  and scenario  $i$  involves the upscaling of an component with autocorrelation  $\rho_t$  according to equation (2.284) over a factor  $N = T_m/T_i$ . For this purpose, the method for upscaling over  $N$  equal components is the most efficient (see section 2.4.3), but also the Hohenbichler method can be used or other combining techniques. Note that the fact that  $N$  is likely not to be an integer in this case is no obstacle whatsoever.

The procedure is repeated for all scenarios and all random variables. The result is a set of  $\alpha$ -values and  $\beta$  for each barrier scenario, that represent mutually consistent time scales. Subsequently an overall failure probability for all scenarios combined can be computed with e.g. the combining techniques of section 2.4.

#### 2.6.2.5 Step 5: Upscaling from cross section to a flood defence section.

For this purpose the method as described in section 2.5.5 is applied. The probability of failure for a cross section is upscaled to the probability of failure for an entire flood defence section, taking into account the spatial correlation of all the random variables involved. This step will increase the failure probability, except in the hypothetical case were the spatial correlation is equal to 1.

#### 2.6.2.6 Step 6: Incorporation of the probability of occurrence per wind direction.

Steps 1-5 are all executed separately per wind direction. The result is a probability of failure of a flood defence section (for one failure mechanism and one closure scenario) *given* the wind direction. The next two steps (step 6 and 7) are to incorporate the probability of occurrence of the wind direction and to combine the failure probabilities of all wind directions. Before these steps are executed a temporal upscaling step is required.

The next step is to incorporate the probability of occurrence of the wind direction in order to obtain the probability of failure *and* the wind direction. In formula, the following is computed:

$$P(Z_i < 0 \cap \phi \in \phi_i) \quad (2.285)$$

Where  $Z_i$  is the Z-function for wind direction  $i$ ,  $\phi$  is the wind direction and  $\phi_i$  is wind sector  $i$ . The manner in which this probability is computed is explained in section 2.5.3.2.

#### 2.6.2.7 Step 7 Combine the failure probability for wind directions

In this step the probabilities of failure per the wind direction are combined to compute an omni-direction failure probability:

$$P(Z < 0) = P\left(\bigcup_{i=1}^{N_s} Z_i < 0 \cap \phi \in \phi_i\right) \quad (2.286)$$

Where  $N_s$  is the number of wind direction sectors,  $Z$  is the omnidirectional Z-function and the other variables are as defined in section 2.6.2.6. The manner in which this probability is computed is explained in section 2.5.3.2.

Before the failure probabilities of the wind directions are combined, first a temporal upscaling procedure needs to be done in order to make the block durations of all random variables consistent over the wind directions. This is a similar step as described in step 4, in which the barrier scenarios were combined. For each random variable involved, the largest block duration over the  $N_s$  wind directions is determined and subsequently all components representing the wind directions are upscaled to this largest block duration.

#### 2.6.2.8 Derive the influencing factor of the wind direction

This step is described in section 2.5.3.3.

### 2.6.2.9 Step 9: Upscaling to a year.

The previous steps result in a failure probability for a single dike section and a single failure mechanism. The time scale for which this failure probability is derived is the block duration, which can be different for the various random load variables. Therefore, first an upscaling needs to be done to scale all variables to the same block duration, i.e. the largest block duration of all random load variables. The upscaling is done step by step from the smallest block duration to the second smallest, then to the third smallest etc. up to the largest block duration. This process is schematically depicted in Figure 2.44.

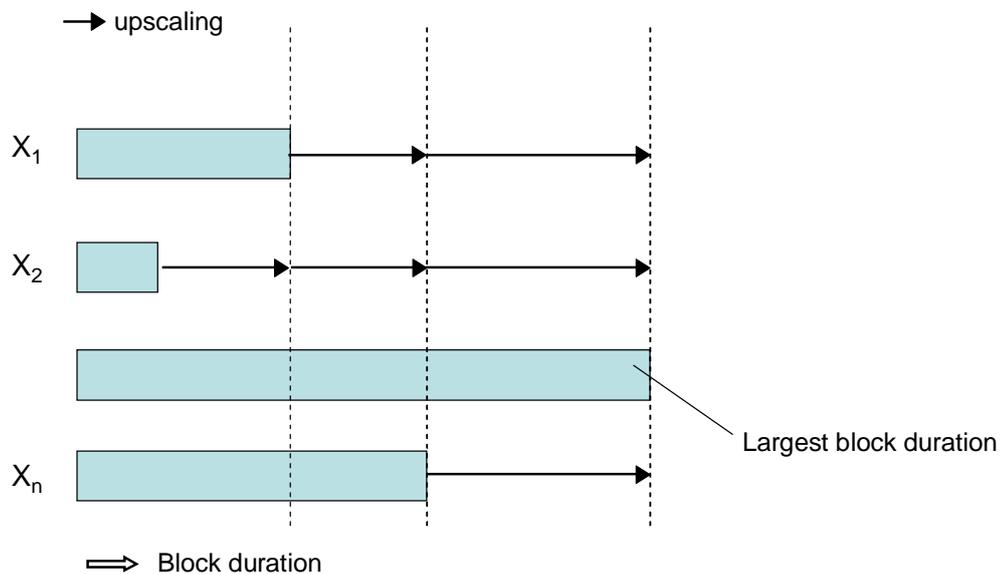


Figure 2.44 Schematic view of upscaling the block duration to the largest block duration of all  $n$  random load variables.

Consider a single upscaling step from duration  $D_1$  to duration  $D_2$ . As described above, the procedure is such that there are no block durations between  $D_1$  and  $D_2$ . This means all load variables with a block duration  $< D_2$  have a block duration  $\leq D_1$ . In the FBC model, a random variable is fully autocorrelated within the period of its block duration, whereas there is no autocorrelation over larger periods. This means in the upscaling from  $D_1$  to duration  $D_2$ , all variables with a block duration  $\leq D_1$  have an autocorrelation equal to 0, whereas all the other variables have an autocorrelation equal to 1. These autocorrelations are used in the following equation to compute the autocorrelation of the component over the interval  $[D_1, D_2]$ :

$$\rho = \sum_{k=1}^n (\alpha_k)^2 \rho_k = \sum_{k \in D_2^*} (\alpha_k)^2 \quad (2.287)$$

In which  $\rho$  is the autocorrelation of the component,  $\rho_k$  is the autocorrelation of the  $k^{\text{th}}$  load variable,  $\alpha_k$  is the  $\alpha$ -value of variable  $k$  and  $D^*$  is the set of variables for which the block duration is  $\geq D_2$ . Each upscaling step can be executed using e.g. the combination techniques of section 2.4. The result of one upscaling step is a (new) set of  $\alpha$ -values for all random variables and a (new)  $\beta$  for the upscaled component.

After all variables have been upscaled to the time scale of the largest block duration, the next step is to upscale the component from the largest block duration to a year. In this step, the autocorrelation for all random variables is equal to 0 because the upscaling is done for time scales that are larger than the largest block duration. This means that the autocorrelation for the entire component is equal to 0. The upscaling is once again done with the combination techniques of section 2.4.

#### 2.6.2.10 Step 10: Combining failure probabilities of mechanisms

The previous steps result in a failure probability for a single dike section and a single mechanism for the period of a year. The next step is to combine the failure probabilities for all considered mechanisms. This is done with the combining methods of section 2.4. The result is the annual failure probability for one dike section.

Note: this procedure is relatively straightforward in PC-Ring, for Hydra-Ring it will be slightly more complex. Not so much from a computational point of view, but more from a programming point of view. This has to do with the fact that in PC-Ring all failure mechanisms are evaluated on the same set of segments. Within Hydra-Ring it is the objective to have a subdivision of the dike ring in subsections that can be different for different failure mechanism. The difficulty is that failure mechanisms cannot be combined directly anymore, because they do not necessarily refer to the same dike section. So the combining needs to be done for the parts of the sections where there is overlap and that brings with it some "administration" as there are many combinations of schematisations possible.

The method in Hydra-Ring is to define "presentation segments" for which the combining results of the mechanisms will be derived (see the design document and Figure 2.45). The results for each mechanism first need to be rescaled to the presentation segments. Subsequently, the mechanisms are combined to derive the failure probability of the presentation segment with the combining techniques of section 2.4. Theoretically, a dike segment can be spread out over multiple presentation sections. In practice, however, presentation segments will be at least the scale of the dike segments, which means one dike segment is spread out over 2 presentation segments at most. The representative cross section of the dike segment therefore needs to be upscaled to:

- 1 The dike section;
- 2 The overlap of the dike section with the first presentation segment;
- 3 The overlap of the dike section with the second presentation segment.

The upscaling technique will be the same as the one that is used in step 5. Note that the  $\alpha$ -values and  $\beta$  of the cross section are required, so these upscaling steps either need to be done in step 5, or the  $\alpha$ -values and  $\beta$  of the cross section need to be stored and used as additional input of step 9.

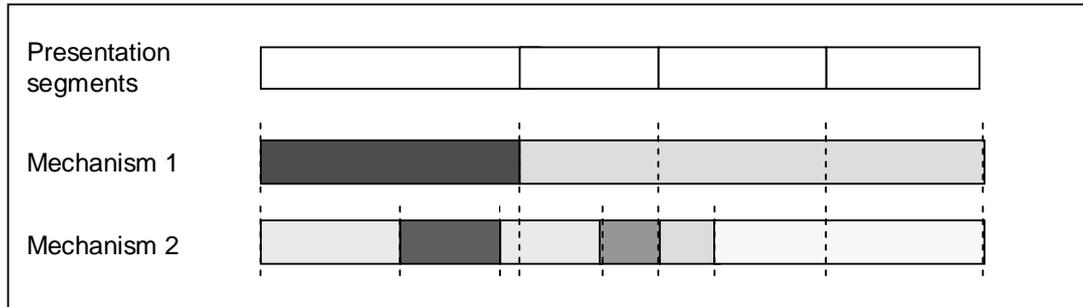


Figure 2.45 Illustration of different divisions for different mechanisms, and an additional division for presentation of combined results.

#### 2.6.2.11 Step 11: Combining failure probabilities of segments

The last step is to combine the failure probabilities of the segments. This is done with the combining techniques of section 2.4.

#### 2.6.3 Procedure for NTI

The procedure for computing the failure probability of a flood defence system, using the NTI procedure for upscaling in time, can be summarized as follows:

1. Determine the failure probability of the smallest component, i.e. a combination of one cross section, one sub-mechanism one wind direction, one closure situation and one basic time step.
2. Combine the failure probabilities of the sub-mechanisms to one failure probability per failure mechanism.
3. Combine failure probabilities of the closure situations with the possible closure scenarios.
4. Upscale the failure probability of a cross section to a dike section.
5. Combine the failure probability for each wind direction with the probability of occurrence of that wind direction.
6. Combine the failure probabilities of all wind directions.
7. Derive the influencing factor of the wind direction.
8. Combine failure probabilities of all basic time steps (generally a tidal period) within the time scale of the slowly evolving random variables.
9. Upscale the failure probabilities to a year.
10. Combine failure probabilities of all failure mechanisms.
11. Combine failure probabilities of all segments.

Relevant differences with the FBC procedure are:

- The combining for wind directions can be executed directly after the loop over all wind directions is finished. The same holds for the closure scenarios. The difference with FBC is that for FBC an upscaling procedure for block durations is required before the combining can be carried out.
- Failure probabilities are computed separately for each basic time step in the schematized hydrograph of the slowly varying random variable(s). The resulting failure probabilities are different for each time step, as each time step represents a different phase in the evolution of the hydrograph. This means steps 1-7 need to be repeated  $T$  times, where  $T$  is the number of basic time steps within the hydrograph. This is the reason why NTI requires more computation time than FBC.

- Upscaling to a year is started from a different time scale. For FBC this time scale is the largest block duration over all random load variables, whereas for NTI this is the duration of the schematised hydrograph of the slowly varying random variable(s), see section 2.5.2.

#### 2.6.4 Procedure for APT

The procedure for computing the failure probability of a flood defence system, using the APT procedure for upscaling in time, can be summarized as follows:

1. Determine the failure probability of the smallest component, i.e. a combination of one cross section, one sub-mechanism one wind direction, one closure situation and one basic time step.
2. Combine the failure probabilities of the sub-mechanisms to one failure probability per failure mechanism.
3. Combine failure probabilities of the closure situation with the possible closure scenarios.
4. Upscale the failure probability of a cross section to a dike section.
5. Combine the failure probability for each wind direction with the probability of occurrence of that wind direction.
6. Combine the failure probabilities of all wind directions.
7. Derive the influencing factor of the wind direction.
8. Upscale the failure probabilities from the basic time step (generally a tidal period) to the time scale of the slowly evolving random variables.
9. Upscale the failure probabilities to a year.
10. Combine failure probabilities of all failure mechanisms.
11. Combine failure probabilities of segments.

Relevant differences with the NTI procedure are:

- For NTI, failure probabilities are computed separately for each basic time step in the schematized hydrograph of the slowly varying random variable(s). The resulting failure probabilities are different for each time step, as each time step represents a different phase in the evolution of the hydrograph. This means steps 1-7 need to be repeated T times, where T is the number of basic time steps within the hydrograph. For APT, steps 1-7 are executed only once and the resulting failure probability can be upscaled directly to the time scale of the hydrograph. The different phases in the hydrograph are represented by a scaling factor c, which is a random variable in the APT approach.

### 3 Hydraulic load models: generic set-up

#### 3.1 Introduction

The main objective of Hydra-Ring is to quantify the probability of flooding of areas that are protected by a system of flood defences such as dikes, dunes or hydraulic structures ('dike ring'). As described in chapter 2, this is done by first computing failure probabilities of flood defences at individual locations (cross sections) and subsequent integration to the failure probability of an entire dike ring, taking spatial correlations into account. Failure occurs when the resistance ( $R$ ) of a flood defence is exceeded by the hydraulic load ( $S$ ).

For each cross section of the flooddefence system, the probability that  $S > R$  is quantified. The load,  $S$ , typically consists of the combination of water levels and waves and in some cases currents. In order to determine failure probabilities, the relevant statistical features of water levels and waves are required. Unfortunately, for most dike sections there are no measurements available of water levels and waves directly in front of the flood defence. Statistics of water level and waves therefore need to be estimated from other sources. For this purpose, a hydraulic load model is required.

This section describes the relevant features of hydraulic load models as implemented in Hydra-Ring. There are multiple hydraulic load models in Hydra-Ring because of the differences in characteristics of the various water systems. However, the set-up of the load models follows as much as possible a generic pattern.

Section 3.2 describes the general concept of hydraulic load models and the role of these models in probabilistic failure computations. The subsequent sections (3.3-3.5) describe the major components of the hydraulic load models. Section 3.3 describes random variables and the associated probability distribution functions that quantify their statistical properties. Section 3.4 describes correlation models that quantify the statistical dependence between random variables. Section 3.5 describes the hydrodynamic models that are applied to derive the hydraulic load at the flood defence.

#### 3.2 Hydraulic load models in probabilistic failure computations

##### 3.2.1 Introduction

In chapter 2 it was described that Hydra-Ring uses the following two-step procedure to quantify the probability of failure of a dike ring:

1. quantification of the probability of all the individual components of the system, and
2. integration of the failure probabilities of the components to derive the failure probability of the entire system (system analysis).

A single component in Hydra-Ring refers to a combination of one cross section, one failure mechanism, one wind direction, one closure scenario and one relatively small (<1 day) time interval during which load conditions are assumed to be constant. The information that is contained in the load models is mainly used to determine the failure probability of individual components (step 1), as will be described in section 3.2.2. But also for the system analysis (step 2) some statistical information on hydraulic loads is required, especially on spatial and temporal correlations.

### 3.2.2 Hydraulic load models in probabilistic failure computations for single components

Section 2.3 describes the probabilistic methods that are implemented in Hydra-Ring for computing failure probabilities for individual components: FORM, numerical integration and three variants of Monte Carlo (Crude Monte Carlo, directional sampling and importance sampling). For practical purposes, these methods work with standard normally distributed variables  $U_1, \dots, U_n$  that represent the 'real world' variables  $X_1, \dots, X_n$ . Variables  $X_1, \dots, X_n$  typically describe load characteristics like sea water level or river discharge and flood defence characteristics like dike height or grain size. In each of the probabilistic methods, the dependence of the limit state function,  $Z$ , on the  $U$ -variables is explored in order to quantify the probability that  $Z < 0$ , i.e. to quantify the probability of failure. Essentially this is done according to the following procedure:

- 1 Translation of 'hydraulic'  $U$ -variables into 'real world'  $X$ -variables;
- 2 Translation of 'hydraulic'  $X$ -variables into the hydraulic load ( $S$ ) at the flood defence;
- 3 Translation of 'resistance'  $U$ -variables in 'real world'  $X$ -variables;
- 4 Quantification of the resistance ( $R$ ) of the flood defence from the 'resistance'  $X$ -variables;
- 5 Comparison of load and resistance to determine  $Z=R-S$

In the probabilistic computations, this procedure is repeated multiple times to provide sufficient insight in the function  $Z(U)$  and to subsequently provide an estimate of the failure probability. Steps 1 and 2 above are essentially the hydraulic load model. In step 1 realisations of the  $U$ -variables are transformed into corresponding realisations of the  $X$ -variables. The generic concept of the transformation is explained in section 2.2.3. If the  $X$ -variables are mutually independent, the transformation between  $U$ -variables and  $X$ -variables is based completely on the probability distribution functions of the individual  $X$ -variables. If the  $X$ -variables are not mutually independent, correlation models are required to describe the transformation between  $U$ -variables and  $X$ -variables. Furthermore, correlations in space and time need to be taken into account.

Each load model roughly consists of the following four components (see Figure 3.1):

- 1 probability distribution functions of random load variables;
- 2 correlations between the random variables and correlations in space and time;
- 3 physical relations that translate possible realizations of the random variables into hydraulic loads at flood defences; and
- 4 additional load parameters.

These four items are discussed in more detail below.

**[1]** The random variables describe the potential events that may result in failure of the flood defences. These random variables include river discharge, wind speed, sea water levels or malfunctioning storm surge barriers. The probability distributions describe:

- probabilities of exceedance of (high) threshold values;
- probabilities of occurrence of "discrete" events, such as the probability that the wind direction is within a given sector or the probability of a malfunctioning barrier;
- probabilities of durations of exceedances of threshold levels

Section 3.3 provides more detail on probability distribution functions

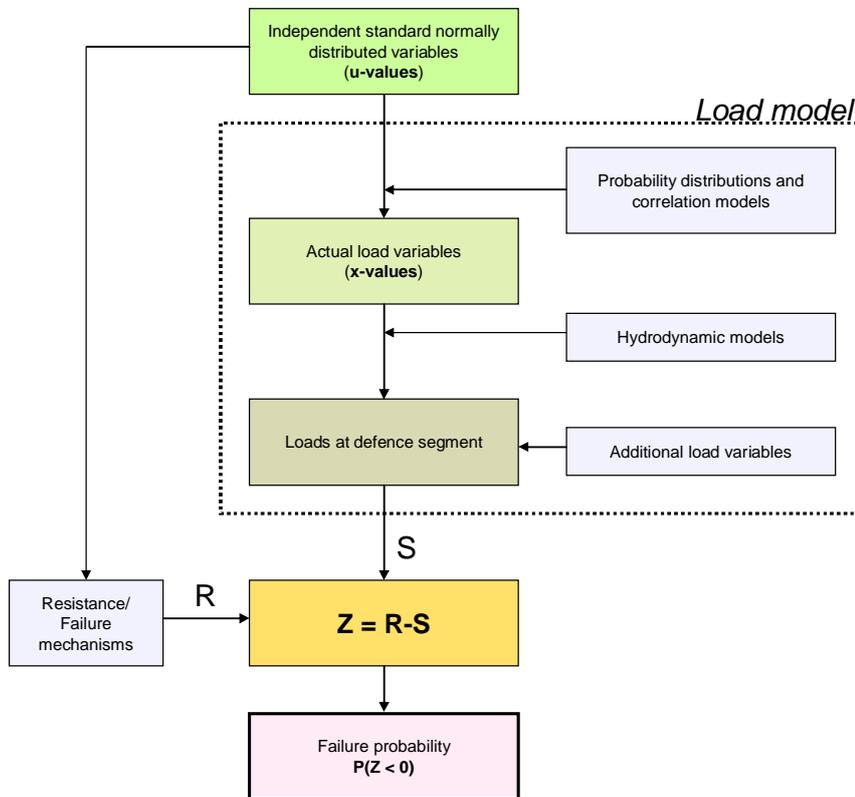


Figure 3.1 Incorporation of the load model in the probabilistic failure computation of a single component.

[2] Statistical dependence among random hydraulic variables generally stems from the fact that they have a common meteorological cause. Usually, this dependence increases the probability of occurrence of floods. Therefore, in many applications in flood risk analysis there is a need for techniques that describe the statistical dependence among random variables. Hydra-Ring describes statistical dependence through application of a library of generic correlation models.

Correlation in time (autocorrelation) is mainly relevant for relatively slowly varying random variables like river discharge and lake level. The basic model time step in Hydra-Ring currently is one tidal period, i.e. a little over 12 hours. The probability of failure of a flood defence is first computed for a single tidal period and subsequently the probability is integrated to a period of one year. For the slowly varying variables like river discharge and lake level, the correlation between two successive tidal periods is close to one, and therefore needs to be taken into account. Also for resistance variables, the correlation in time between two successive tidal periods is close to one.

For “faster” varying variables like wind speed and sea water level, the (auto-)correlation is assumed to be small enough to be neglected. In reality, there is of course some correlation. Nevertheless the error introduced, especially for higher thresholds of wind speed and river discharge, is small. This is due to the fact that statistics of extremes for these variables are generally derived at the annual scale, and then rescaled to the tidal scale, using the same assumption of zero autocorrelation (see section 3.3.3). Subsequent use of the same

assumption in the probabilistic computations will result in reproduction of the correct exceedance probabilities at the annual scale. For lower threshold of the wind speed and sea water level, the error introduced may be more substantial. Potential consequences of this assumption for the non-tidal river area in the Netherlands have been analysed in Geerse en van Veen, [2007]. It was concluded that this may lead to a small overestimation of the failure probability.

Correlation in space for load variables is generally much larger than for variables describing the strength of the flood defence. Within specified regions in the Netherlands, a 100% spatial correlation of random load variables is generally assumed. Inbetween regions, spatial correlation is sometimes taken to be lower than 100%.

Section 3.4 provides more details on correlation models.

**[3]** For all potential events, described by the random variables, the resulting hydraulic load at the flood defences (water levels, waves, currents) needs to be derived. This is generally achieved by executing a large number of hydrodynamic model simulations. Well-selected realisations of the random variables are used as boundary conditions for the hydrodynamic model simulations. The hydrodynamic model simulations are executed prior to the application of Hydra-Ring and the relevant model results are stored as input databases. These databases serve as a look-up table to link the random variables (input) to the hydraulic load at the flood defence (output). In some hydraulic load models, relatively straightforward empirical relations are used as an alternative to input, for example the Bretschneider equation for estimating wave height as a function of the wind and the river geometry.

Section 3.5 provides more details on hydrodynamic models.

**[4]** Some models of failure mechanisms require additional hydraulic input variables that are not directly related to high water events, for example the mean sea water level. These variables are not influenced by the random variables of step 1, and are therefore not part of the loads that are derived in step 3. Others, such as the water level at the inner slope of a dike, are *treated* as being completely independent of the random variables of step 1, even though in reality there may be some dependence. These type of additional loads need to be derived for each dike section separately. This is a pre-processing step for Hydra-Ring and the results are stored in input databases.

The additional load variables will be described in chapter 5, where all the failure mechanisms are described.

### 3.2.3 Generic, modular set-up of load models in Hydra-Ring

Hydra-Ring uses system units that are referred to as *regions*. For each region a separate load model is created in Hydra-Ring. Regions are therefore combinations of locations for which the same load model is used. The choice of the regions is therefore determined by the clustering of locations for which it makes sense to apply the same load model, i.e. locations that are related to the same set of random variables and hydraulic load models. Generally, the regions are associated with different water systems (i.e. rivers, lakes, sea) but sometimes also different types of flood defences (dikes, dunes) require different load models.

Currently, there are 16 regions defined in the Netherlands (see Table 4.1). The associated load models are described in chapter 4. It is likely that Hydra-Ring will be extended with additional load models for other areas, in the Netherlands or abroad. Furthermore, multiple

versions of the load model may be created for a single region, e.g. for purposes of research, development (beta-versions) and impact assesment of climate change or flood mitigating measures. This increases the need for Hydra-Ring to be set-up in a clear, structural and transparent manner. For this reason a generic, modular structure for describing load models is designed and implemented in Hydra-Ring.

In the generic structure, a modular approach is used, making sure the load models make use (as much as possible) of the same set of routines. Routines that more or less serve the same purpose are clustered in libraries. The libraries are designed in such a way that the input/output structure of all the modules in the library is the same. Example libraries in the computational core of Hydra-Ring are:

- Library of probabilistic computation techniques for single components (see section 2.3)
- Library of statistical distribution functions (see section 3.3)
- Library of correlation models (see section 3.4)
- Library of failure mechanisms (see chapter 5)

The specific characteristics of a load model are *all* stored in input databases. Among others these databases contain parameter values of distribution functions and correlation models of the random variables and results of hydrodynamic model simulations. The use of input databases means the model code does not contain any region-specific information. The names of the regions are therefore not part of any "if-statement" in the code. One of the main advantages of this set-up is that new load models can be added to Hydra-Ring without having to change the code. Furthermore, it makes the code easier to maintain as no region specific exceptions need to be verified when the code is changed.

### 3.3 Statistical distribution functions of random variables

#### 3.3.1 Generic description of the application of distribution functions

As stated in the previous section, probabilities of exceedance of random variabels like river discharge, sea water level and wind speed need to be described by the hydraulic load model. The following three types of functions can be used to describe the statistical properties of random variables:

- 1 Cumulative distribution function (CDF);
- 2 Inverse cumulative distribution function (inverse CDF);
- 3 Probability density function (PDF).

The CDF,  $F(x)$ , provides the probability of non-exceedance,  $p$ , of each potential realisation,  $x$ , of random variable  $X$ . The inverse CDF,  $F^{-1}(p)$ , provides the realisation  $x$  that has a probability of non-exceedance  $p$ . The relation between the CDF and the inverse CDF is thus as follows:

$$F(x) = p \Leftrightarrow x = F^{-1}(p) \quad (3.1)$$

The PDF,  $f(x)$ , is the derivative of the CDF:

$$F(x) = \int_{-\infty}^x f(\tau) d\tau \Leftrightarrow f(x) = \frac{dF}{dx}(x) \quad (3.2)$$

# Deltares

1206006-004-ZWS-0001, 2 January 2013, draft

The PDF provides the probability density for any given value of  $x$ . The probability density is the probability *per unit* value. For the normal distribution function, the pdf is the “famous” bell-shaped curve. As an example, Figure 3.2 shows the CDF, inverse CDF and PDF of the standard normal distribution function.

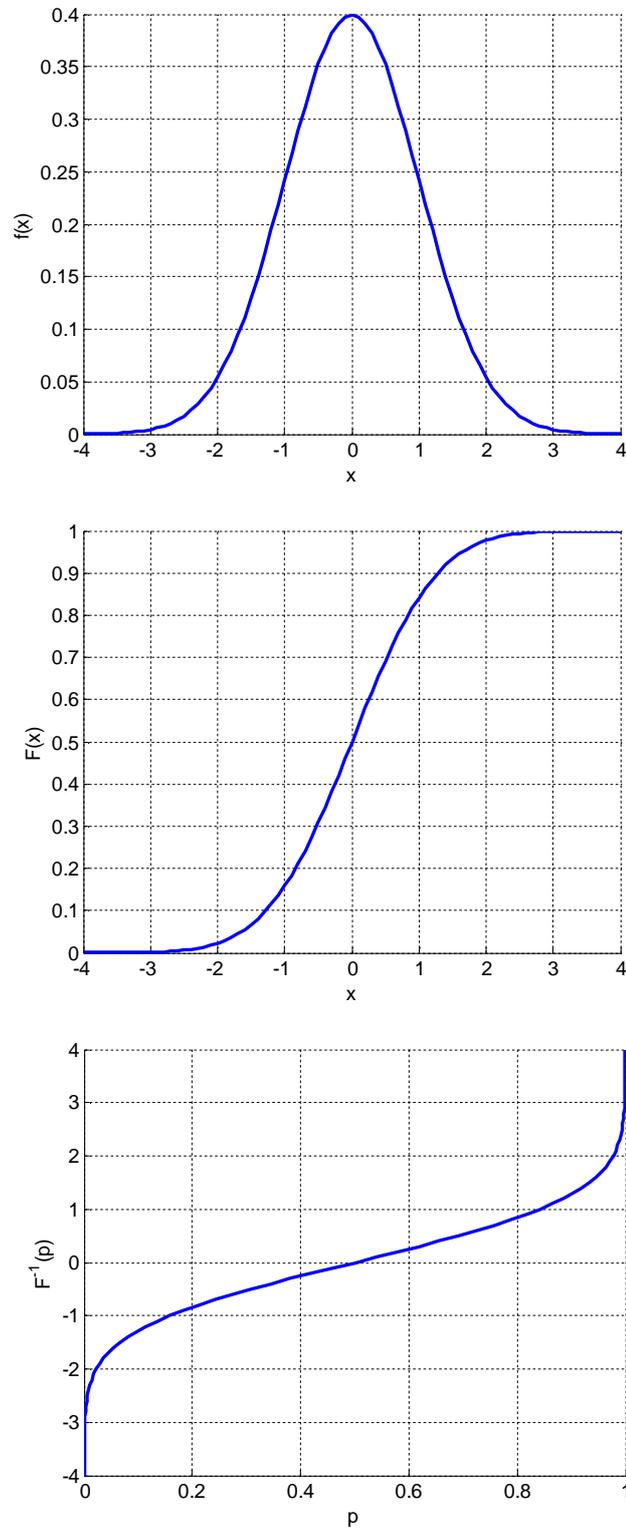


Figure 3.2 PDF (top), CDF (middle) and inverse CDF (bottom) of the standard normal distribution function

The CDF,  $F(x)$ , has the following properties:

- 1  $F(x)$  is non-decreasing;
- 2  $\lim_{x \rightarrow -\infty} F(x) = 0$ ;
- 3  $\lim_{x \rightarrow \infty} F(x) = 1$ .

Property 1 can be easily proven: if  $x_1 < x_2$ , then  $P[X \leq x_1] \leq P[X \leq x_2]$  and thus  $F(x_1) \leq F(x_2)$ . For a formal proof of properties 2 and 3, the reader is referred to Grimmet and Stirzaker [1982, pp 20]. But even without a proof it is intuitively clear that a realization from a probability function will be lower than  $\infty$  and higher than  $-\infty$ .

Since  $F(x)$  is non-decreasing, the inverse CDF,  $F^{-1}(p)$ , is also a non-decreasing function. In probabilistic computations in Hydra-Ring, mainly the inverse CDF,  $F^{-1}(x)$ , of a variable  $X$  is applied, as schematically depicted in Figure 3.3. The library of distribution functions in Hydra-Ring therefore mostly consists of inverse CDF's. The procedure of Figure 3.3 is explained below.

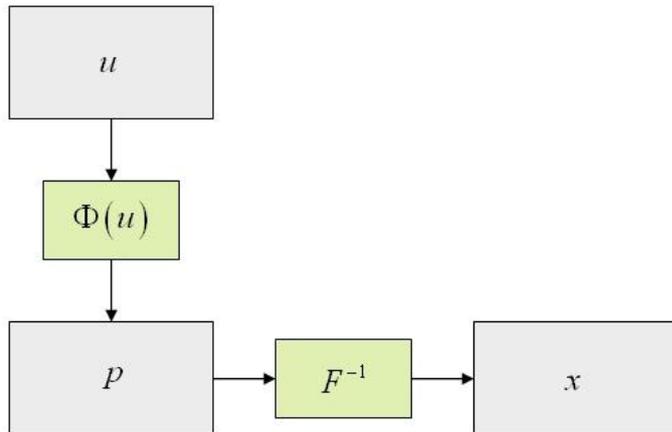


Figure 3.3 Procedure for determining a load variable associated with a randomly selected standard normally distributed variable ( $u$ -value), for the case of uncorrelated variables

As described in section Section 2.3 and 3.2.2, random variables are represented by standardised  $U$ -variables in the probabilistic computations in Hydra-Ring, and the function  $Z(U)$  is explored to derive an estimate of the failure probability. In order to evaluate function  $Z(U)$ , the realisations of the  $U$ -variables are first translated to the corresponding realisations of the  $X$ -variables and subsequently the  $Z$ -value is determined. Assume for the sake of simplicity that the  $X$ -variables are mutually independent (correlations will be dealt with in section 3.4). As explained in section 2.2.3, the transformation from a realization,  $u$ , of variable  $U$ , to realization  $x$ , of variable  $X$ , is done in such a way that the (non-)exceedance probabilities of  $u$  and  $x$  are equal. This transformation, as depicted in Figure 3.3, can be formulated as follows:

$$\Phi(u) = F(x) \Rightarrow x = F^{-1}(\Phi(u)), \tag{3.3}$$

where:

$\Phi$  = standard normal distribution function  
 $F$  = CDF of variable  $X$   
 $F^{-1}$  = inverse CDF of  $X$   
 $x$  = realization of  $X$   
 $u$  = realisation of  $U$

This procedure automatically guarantees that variable  $x$  is a realization from distribution function  $F(x)$  and therefore correctly represents the statistical properties of variable  $X$ . This is demonstrated below.

First it needs to be shown that the value  $p = \Phi(u)$  is a realization from a standard *uniform* distribution function. The standard uniform distribution function is the CDF in which each value in the range  $[0,1]$  has equal probability density. The CDF of this function is as follows (see also Figure 3.4):

$$F(x) = \begin{cases} 0 & ; x \leq 0 \\ x & ; 0 < x < 1 \\ 1 & ; x \geq 1 \end{cases}, \quad (3.4)$$

Consider a realization,  $u^*$ , of the standard normal distribution function with a probability of non-exceedance equal to  $p^* = \Phi(u^*)$ . By definition this means that the probability that a random sample  $u$  from the standard normal distribution function does not exceed  $u^*$  is equal to  $p^*$ . In formula:

$$P[u \leq u^*] = p^*, \quad (3.5)$$

Since  $\Phi$  is a CDF, it is a non-decreasing function. Therefore it follows from equation (3.5) that:

$$P[\Phi(u) \leq \Phi(u^*)] = p^*, \quad (3.6)$$

And since by definition  $p^* = \Phi(u^*)$ , this simplifies to:

$$P[\Phi(u) \leq p^*] = p^*, \quad (3.7)$$

So the probability that  $\Phi(u)$  does not exceed a given value  $p^*$  ( $0 \leq p^* \leq 1$ ) is equal to  $p^*$ . This shows that  $\Phi(u)$  is a realization from a standard uniform distribution function, as described by equation (3.4) and depicted in Figure 3.4.

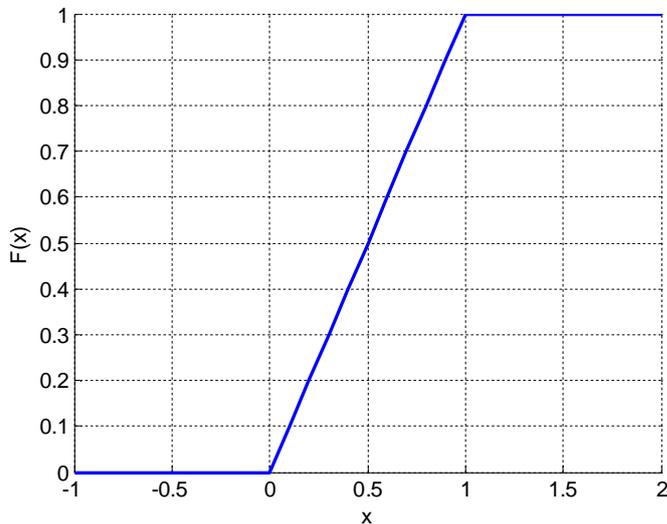


Figure 3.4 Standard uniform distribution function (CDF)

It has been demonstrated that the value  $p$  in Figure 3.3 is a realization from the standard uniform distribution function. The next step is to show that the value  $x = F^{-1}(p)$  in Figure 3.3 is a realization from distribution function  $F(x)$ . For this purpose, consider a value  $x^*$  with probability of non-exceedance  $p^*$ . This means:  $F(x^*) = p^*$  and  $x^* = F^{-1}(p^*)$ . The value  $p$  in Figure 3.3 is taken from a standard uniform distribution function (as proven above), which means:

$$P[p \leq p^*] = p^* \tag{3.8}$$

Since  $F$  is an inverse CDF, it is a non-decreasing function and therefore it follows from equation (3.8) that:

$$P[F^{-1}(p) \leq F^{-1}(p^*)] = p^* \tag{3.9}$$

By definition  $x^* = F^{-1}(p^*)$  and  $x = F^{-1}(p)$ , which means equation (3.9) simplifies to:

$$P[x \leq x^*] = p^* \tag{3.10}$$

Since  $p^* = F(x^*)$ , this means:

$$P[x \leq x^*] = F(x^*) \tag{3.11}$$

This shows that value  $x$  in Figure 3.3 is a realization from distribution function  $F(x)$ .

To summarize: In probabilistic computations in Hydra-Ring, Hydra-Ring works with standard normalized U-variables. The limit state function  $Z(U)$  is explored to derive an estimate of the failure probability. In order to evaluate function  $Z(U)$ , the realisations of the U-variables are first translated to the corresponding realisations of the X-variables to be able to determine

the Z-value. In this transformation, the inverse CDF of variable X is applied, to provide variable X with the correct statistical properties.

### 3.3.2 Standard distribution functions and parameter values

Table 3.1 shows the set of statistical distribution functions that are included in the library of Hydra-Ring. Each of these distributions and the required input parameters will be described in detail in chapter 6. As described in the previous section, the library contains the “inverse cumulative distribution functions” which means the input consists of a probability,  $p$ , of non-exceedance and the output consist of the associated realization,  $x$ , of the variable that is described with this distribution function.

Besides  $p$ , the input of the inverse CDF's also consists of a set of parameter values  $\theta = (\theta_1, \dots, \theta_n)$ . These parameter values quantify the relation between  $p$  and  $x$ . Note that the value of  $n$  can be different for different distribution functions, as shown in the second column of Table 3.1. Each distribution function as mentioned in Table 3.1 has a fixed *number*,  $n$ , of parameters but the *values* of these parameters will be different for different variables. So, for instance, it is possible to describe both river discharge and sea water level with the lognormal distribution, but the values of the two parameters of this distribution function for river discharge will be different from the values that are used for sea water level. In other words: the same module can be used to describe probabilities of different random variables and differences between the variables are characterized by differences in parameter values.

Table 3.1 Probability distributions of load variables, supported in Hydra-Ring.

Probability distributions	Number of parameters
Uniform	2
Normal	2
(Shifted) lognormal	3
(Shifted) exponential	2
Gumbel	2
Weibull	3
Rayleigh	2
Pareto	3
Triangular	3
Multi-linear interpolation	Variable
Modified Gumbel	3
Conditional Weibull	4

### 3.3.3 Temporal scaling of statistical distribution functions

The basic model time step in Hydra-Ring is one tidal period, i.e. a little over 12 hours. The probability of failure of a flood defence is first computed for a single tidal period and subsequently the probability is integrated to a period of one year. In order to compute the failure probability for one tidal period, the statistical properties of the random variables are required. However, statistical distribution functions of random variables are generally only available for the annual time scale. For instance, statistics of extreme sea water levels are usually derived from annual maximum observations or all observations above a user-defined threshold. The resulting distributing function,  $F(x)$ , therefore usually expresses the probability that within any given year the value of  $x$  is exceeded. The probability of an exceedance over a period of a year is generally (much) larger than the probability of exceedance within a tidal period. The available annual statistics of the random variables therefore need to be translated to statistics per tidal period.

First consider a fast evolving random variable like sea water level. The values of such a variables in successive tidal periods are assumed to be independent (see section 2.5.2). Suppose there are  $N$  tidal periods in a year, this means the fast evolving variables can be modeled by taking  $N$  independent samples. In that case the following holds:

$$P[X_{my} < x] = P[X_{mt} < x]^N \tag{3.12}$$

Where

- $X_{my}$  = annual maximum value of variable  $X$ ;
- $X_{mt}$  = maximum value of variable  $X$  in a tidal period;
- $N$  = number of tidal periods in a year.

Equation (3.12) shows how the probability of non-exceedance for a year can be derived from the probability of non-exceedance for a tidal period. This transformation is added as an additional step to the procedure as depicted in Figure 3.3. In that procedure, the probabilistic computation begins with a realization,  $u$ , of a standard normally distributed variable  $U$ . Subsequently, the associated probability of non-exceedance is computed. The value of  $p$  is translated to the realization,  $x$ , of random variable  $X$ . For this translation the inverse CDF,  $F^{-1}$ , is required. However, probability  $p$  refers to a tidal period, whereas  $F_x^{-1}$  refers to a year. Therefore, probability  $p$  is transformed to the annual probability through application of equation (3.12). The resulting annual probability is used as input for function  $F_x^{-1}$ , which provides the required value of  $x$  as output. Figure 3.5 shows the inclusion of this additional step to the procedure of Figure 3.3. In formula, this means the value of  $x$  is derived as follows:

$$x = F_{x,year}^{-1} \left( \Phi(u)^N \right), \tag{3.13}$$

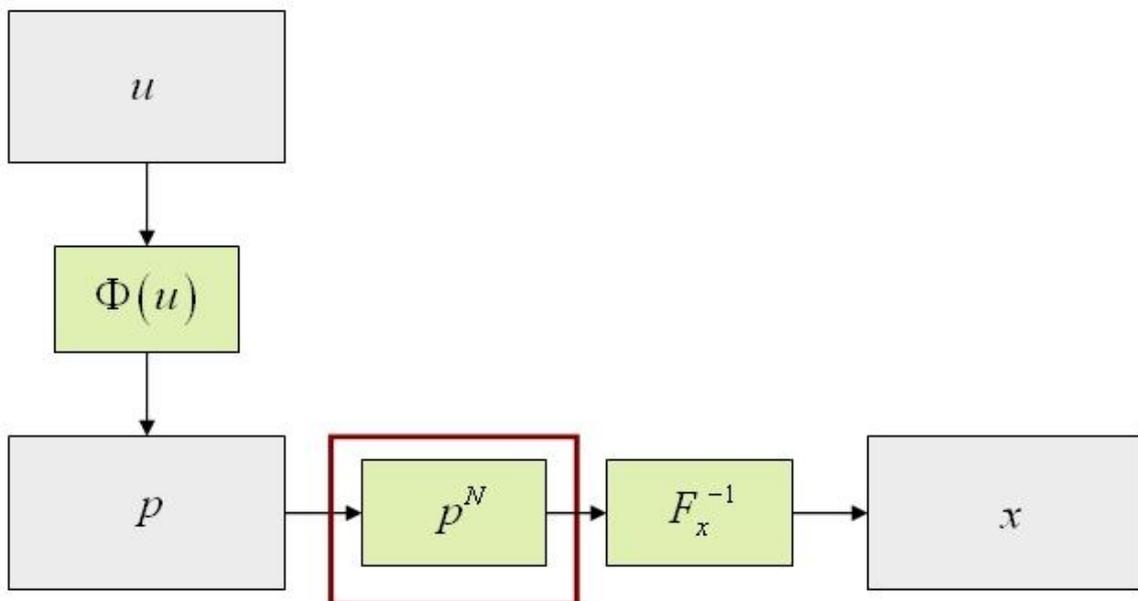


Figure 3.5 Schematization of the procedure going from a  $u$ -value to an  $x$ -value. The step highlighted by the red box is the additional step required when scaling the probability in time.

An important note is that the value of  $N$  will be different for slowly evolving random variables like river discharge and lake water level. The reason is that the realizations of these variables in consecutive tidal periods are highly correlated. This means that equation (3.12) is not valid for these variables, if  $N$  is the number of tidal periods in a year. For slowly evolving variables, the value of  $N$  should therefore be taken equal to the number of independent periods in a year. The value of  $N$  for these variables is therefore related to the value of  $T_b$  or  $T_m$ , depending on whether they are modeled according to the FBC method or the APT/NTI method respectively (see section 2.5.2 for a description of the FBC, APT and NTI methods and the definition of durations  $T_b$  and  $T_m$ ).

Input statistics of Hydra-Ring may represent the annual time scale, whereas others may represent smaller time scales such as a monthly or a tidal period. It is relevant to take these differences in time scales into account. Therefore, the input database of Hydra-Ring demands that the time scale is added to the statistical distribution functions of each random variable.

### 3.3.4 Statistics of durations

The basic model time step in Hydra-Ring currently is one tidal period, i.e. a little over 12 hours. The probability of failure of a flood defence is first computed for a single tidal period and subsequently the probability is integrated to a period of one year. In the upscaling process, the correlation between failure probabilities of subsequent time steps is relevant (see section 2.5.2). This correlation depends on the autocorrelation of the random variables involved. For strength variables, the correlation between all time steps in the year is assumed to be equal to 1. For “faster” varying variables like wind speed and sea water level, the correlation between subsequent time steps is assumed to be small enough to be neglected and is therefore taken equal to 0.

For slowly evolving variables like river discharge and lake water level a hybrid approach is used and for this approach, statistical information on durations is required. The year is considered to be divided into a number of periods. Within a single period, the autocorrelation in time for the slowly evolving variables is taken equal to 1, whereas the autocorrelation between different periods is taken equal to 0 (see section 2.5.2). Hydra-Ring offers three approaches for modelling the temporal evolution and related autocorrelation of slowly evolving random variables (see section 2.5.2):

- 1 NTI
- 2 APT
- 3 FBC

In the first two methods, the time evolution of the slowly evolving variables within a single period is described by a standardised hydrograph (a trapezium, see Figure 3.6) whereas for FBC the value within a single period is assumed to be constant (a block representation, see Figure 3.7). In both cases the peak values are random variables that are considered independent of neighbouring blocks/trapeziums. Besides the peak value, a single trapezium is characterised by the base duration,  $T_m$ , and peak duration  $T_p$ , whereas a block is characterised by the block duration  $T_b$ .

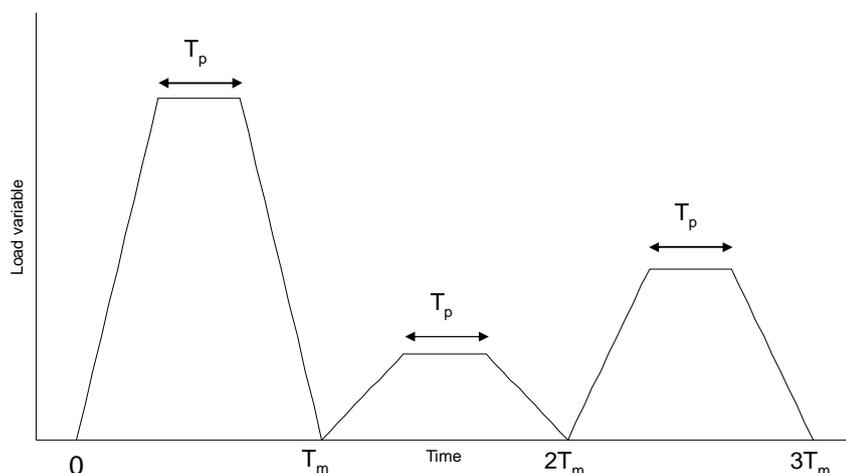


Figure 3.6 Modelling the temporal evolution of discharges and lake levels using trapezia, for probabilistic computation methods APT and NTI

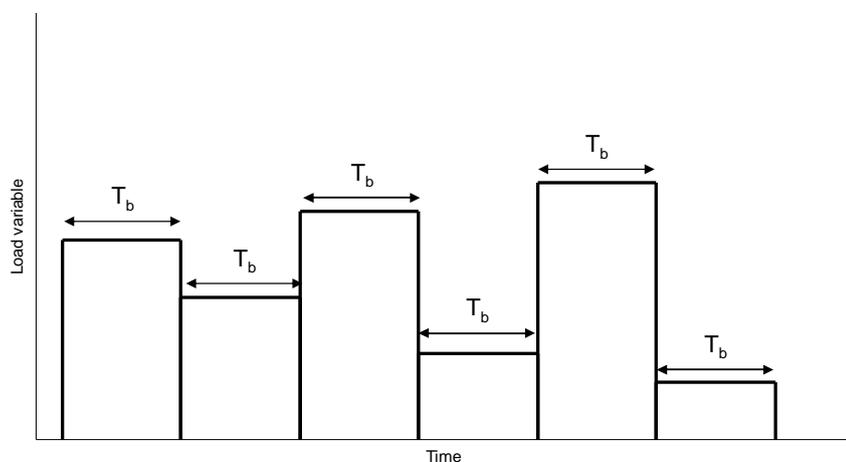


Figure 3.7 Modelling the temporal evolution of discharges and lake levels using “blocks”, for probabilistic computation method FBC.

For example, the value of base duration  $T_m$  is assumed to be constant and taken equal to 30 days for the major rivers in The Netherlands (Rhine and Meuse), 30 days for the IJssel lake and 60 days for the Marker lake. For other rivers or lakes a different value of  $T_m$  may be considered. The choices of the values of  $T_p$  and  $T_b$  are variable and depend on the value of the peak. This is done to guarantee that the resulting statistics are in accordance with available statistics on durations, given as input of Hydra-Ring. The derivation of  $T_p$  and  $T_b$  from these statistics is explained below.

Define, for any given threshold value  $q$  of the slowly evolving variable,  $Q$ , the following statistics:

- $F(q)$  = the average number of exceedances of  $q$  per year;  
 $D(q)$  = the average number of days per year that  $q$  is exceeded;  
 $N(q)$  = the average number of days of a single exceedance event of  $q$ .

The relation between these three functions is as follows:

$$N(q) = \frac{D(q)}{F(q)} \quad (3.14)$$

This is easy to comprehend: the total number of days above level  $q$ , divided by the number of exceedances of level  $q$  provides the average duration of a single exceedance of level  $q$ . Function  $N(q)$  is referred to as the “duration curve” and can be obtained from function  $F(q)$  and  $D(q)$  that are generally more easier to obtain directly from the data.

In the FBC method, the duration curve determines the choice of the the block duration,  $T_b$ . The block duration is set equal to the value  $N(q_d)$ , where  $q_d$  is the selected value of the slowly evolving variable in the design point (see section 2.5.2.3 for an explanation on design points). In order to derive the value of peak duration  $T_p$  of the trapezium in the APT and NTI method, the probability is considered that at an arbitrary moment, the value of  $Q$  exceeds threshold  $q$ :

$$P(Q > q) = \frac{1}{T_m} \int_q^{\infty} f(\tau) L(q, \tau) d\tau \quad (3.15)$$

Where  $L(q, \tau)$  denotes the duration above level  $q$  inside a trapezium with peak value  $\tau$  and  $f$  is the derivative of function  $F(q)$  above. Function  $f$  is therefore directly available from the data, whereas  $L(q, \tau)$  depends on the choice of peak duration  $T_p$ . The peak duration should be chosen such that the probability of equation (3.15) is in accordance with function  $D(q)$ , i.e. the probability should be equal to  $D(q)/365$ . This generally means that the peak duration  $T_p$  is dependent on the peak value of the trapezium. An example of the resulting relation between the peak duration and the peak value is shown in Table 4.2 for the river Rhine at location Lobith. A general method for the derivation of this relation is described in Section 8.2 and 8.3 of Geerse [2011].

For APT and NTI a table like Table 4.2 needs to be available for all slowly evolving random variables. For FBC, the duration curve,  $N(q)$ , needs to be derived. Hydra-Ring supports two options for describing the duration curve. The first is a fitted  $n$ -degree polynomial distribution to the values of  $N(q)$ . The parameters of this polynomial are input into Hydra-Ring. The second option is a table containing two columns. The first column gives  $q$  values and the second column gives  $N(q)$  values. Hydra-Ring uses interpolation (either linear or log-linear) to compute  $N$ -values between the discrete values in the table.

### 3.4 Correlation models

#### 3.4.1 Introduction

In Hydra-Ring, three types of correlations are considered:

- Correlations between different random variables, e.g. the correlation between sea water level and wind speed;
- Correlation in time, e.g. the correlation between river discharge on day  $t$  and river discharge on day  $t+1$

- Correlation in space, e.g. the correlation between the wind speed at 2 different locations.

The three types of correlations are discussed in the sections below

### 3.4.2 Correlations between random variables

#### 3.4.2.1 Generic description

In section 3.3 the procedure for applying statistical distribution functions of random variables in Hydra-Ring was explained. In the explanation, the X-variables were assumed to be mutually independent for the sake of simplicity. In many cases, however, random variables of hydraulic load models are not mutually independent. For instance, wind speed and sea water level are correlated and the same generally can be stated for river discharges of adjacent rivers (such as the Rhine and Meuse in the Netherlands). Correlation between two random variables  $X_1$  and  $X_2$  needs to be taken into account in probabilistic analysis because they influence the probability of failure of the component or system under consideration. In Hydra-Ring, correlation is taken into account with correlation models. Different correlation models are used, in order to be able to encapture specific characteristics of correlation structures. This section describes the generic approach of the correlation models.

The generic approach for modeling correlation between two random variables  $X_1$  and  $X_2$  is to generate realizations  $u_{1,cor}$  and  $u_{2,cor}$  from correlated standard normal variables  $U_{1,cor}$  and  $U_{2,cor}$ .<sup>1</sup> The correlated realizations are subsequently translated into realizations  $x_1$  and  $x_2$  from “real world” variables  $X_1$  and  $X_2$  through application of the procedure of section 3.3 (i.e. through application of inverse CDF’s of variables  $X_1$  and  $X_2$ ). The procedure is depicted in the figure below. The horizontal part of this figure is the exact same procedure as depicted in Figure 3.3 and explained in section 3.3. The correlation model can therefore be considered as pre-processing to the procedure in which the distribution functions are applied.

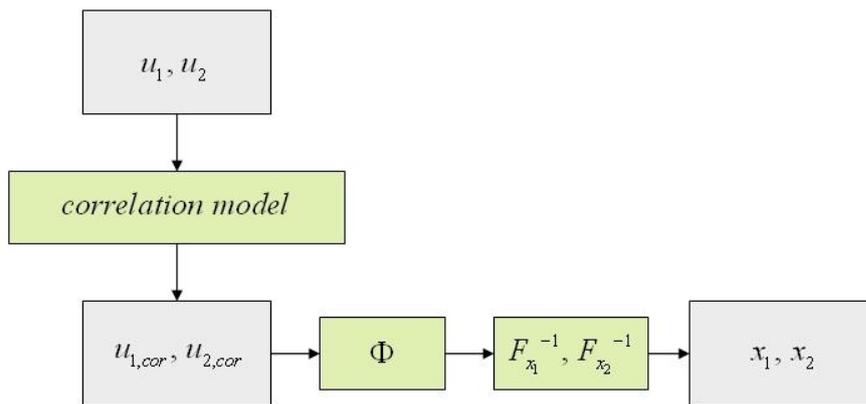


Figure 3.8 Procedure for determining a load variable associated with randomly selected standard normally distributed variables for the case of correlated variables

<sup>1</sup> Note that  $u_1$  and  $u_2$  are strictly speaking only “realisations” if a Monte Carlo procedure is applied, see section 2.3. For methods like FORM and numerical integration,  $u_1$  and  $u_2$  are strategically selected values, not samples from a simulation of a distribution function. However, this fundamental difference in interpretation of  $u_1$  and  $u_2$  has no influence on the applied methods as described in the current section.

Since  $U_{1,\text{cor}}$  and  $U_{2,\text{cor}}$  are correlated variables,  $X_1$  and  $X_2$  are also correlated. Furthermore, since  $U_{1,\text{cor}}$  and  $U_{2,\text{cor}}$  are standard normally distributed and the translation from  $U_{1,\text{cor}}$  and  $U_{2,\text{cor}}$  to  $X_1$  and  $X_2$  is done in the exact same way as described in section 3.3, it is automatically taken care of that  $X_1$  and  $X_2$  are distributed according to their prescribed distribution functions  $F_{X_1}$  and  $F_{X_2}$ . The remainder of this section therefore focuses on the first part of the procedure: the generation of samples  $u_{1,\text{cor}}$  and  $u_{2,\text{cor}}$  of correlated variables  $U_{1,\text{cor}}$  and  $U_{2,\text{cor}}$ .

The generation of  $u_{1,\text{cor}}$  and  $u_{2,\text{cor}}$  starts with the generation of realizations  $u_1$  and  $u_2$  of independent standard normally distributed random variables  $U_1$  and  $U_2$ . Subsequently,  $u_1$  is transformed into a sample  $v$  of variable  $V$  with distribution function  $F_V(v)$ . The transformation is done in similar style as explained in section 3.3, i.e. by making sure the probability of (non-)exceedance of  $u_1$  and  $v$  are equal:

$$\Phi(u_1) = F_V(v) \Rightarrow v = F_V^{-1}(\Phi(u_1)) \quad (3.16)$$

There are no limitations in the choice of distribution function  $F_V(v)$ , but generally standardized distribution functions, such as the standard normal, standard uniform or standard exponential distribution function are used for convenience. Subsequently, a sample  $w$  is introduced that is dependent on  $v$  and  $u_2$ :

$$w = G(v, u_2) \quad (3.17)$$

Were  $G$  is a function of two variables. The fact that  $w$  is a function of  $v$  introduces correlation between  $v$  and  $w$ . Subsequently,  $v$  and  $w$  are translated to samples  $u_{1,\text{cor}}$  and  $u_{2,\text{cor}}$  of standard normally distributed variables  $U_{1,\text{cor}}$  and  $U_{2,\text{cor}}$  by setting equal their respective probabilities of (non-)exceedance. This means  $u_{1,\text{cor}}$  is exactly the same as  $u_1$ , but  $u_{2,\text{cor}}$  will be different from  $u_2$  because of the use of function  $G$ .

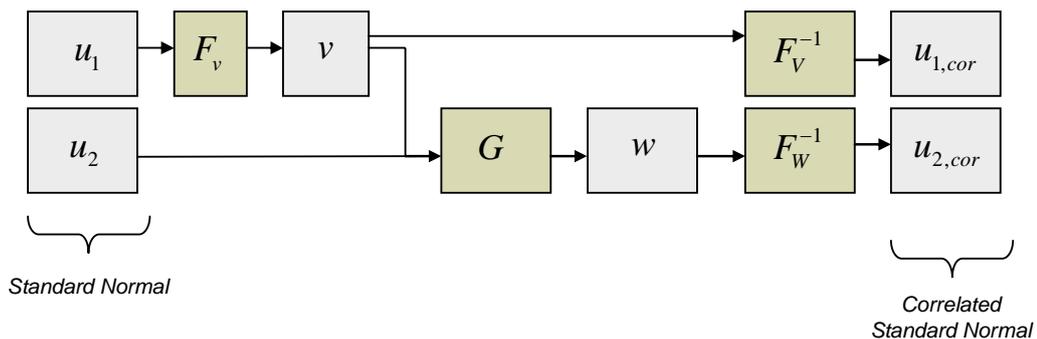


Figure 3.9 Procedure for samples  $u_{1,\text{cor}}$  and  $u_{2,\text{cor}}$  of correlated standard uniform random variables  $U_{1,\text{cor}}$  and  $U_{2,\text{cor}}$ .

Function  $G$  is essentially the correlation model for variables  $X_1$  and  $X_2$ . Virtually any bivariate function  $G$  (note:  $G$  is not a distribution function) can be selected, also because the subsequent transformations guarantee that variables  $X_1$  and  $X_2$  are indeed distributed according to the pre-defined distribution functions  $F_{X_1}$  and  $F_{X_2}$ . Naturally, function  $G$  should be such that it mimics the observed correlation between variables  $X_1$  and  $X_2$  as accurate as possible. For a more detailed background on the derivation and application of correlation models in flood risk analysis, the interested reader is referred to the paper of Diermanse and Geerse [2012].

The correlation models in Hydra-Ring are set up in a modular way, which means additional correlation models can easily be added, as long as they use the structure as described above. The following sections present some example of correlaton models, in order to make the general description more tangible.

### 3.4.2.2 Example 1: the HES-model

The HES model is shorthand for “heteroscedastic model”. The name refers to the fact that this model allows for variation in correlation. The model is very flexible and allows the user to set up a broad range of correlation structures. The model is described in the paper of Diermanse and Geerse [2012]. Note: in some other documentations of Hydra-Ring this correlation model is referred to as the ‘NL-model’.

The model is developed to describe the correlation between two random variables,  $X_1$  and  $X_2$  and ascoiated standard normal random variables  $U_1$  and  $U_2$ . In the model,  $U_1$  is the independent variable and  $U_2$  the dependent variable (see Figure 3.9 for reference). The first step in the model is to translate the independent variable,  $U_1$ , to a standard exponentially distributed variable  $V$ .

$$v = F_V^{-1}(\Phi(u_1)) = -\ln(1 - \Phi(u_1)) \tag{3.18}$$

The transformation is such that  $v$  and  $u_1$  have the same probability of (non-)exceedance. (see section 2.2.3 for more information on this type of transformations). The next step is to compute the realization,  $w$ , of the dependent variable  $W$ . To describe the computation of  $w$ , first a couple of definitions are given. Let  $\lambda(t)$  be a probability density function with a mean of 0 and a standard deviation of 1. This can be, but does not necessarily need to be, a standard normal density function. From  $\lambda(t)$ , distributions of the same type, but with different standard deviations, can be derived through the following transformation:

$$\lambda_\sigma(t) = \frac{1}{\sigma} \lambda(t/\sigma) \tag{3.19}$$

Density function  $\lambda_\sigma(t)$  has a mean of 0 and a standard deviation equal to  $\sigma$ . The cumulative distribution function,  $\Lambda_\sigma$ , is as follows:

$$\Lambda_\sigma(t) = \Lambda(t/\sigma) \tag{3.20}$$

Where  $\Lambda$  is the CDF that is associated with  $\lambda$ . The value of  $w$  is computed as a combination of the realization,  $v$ , of the independent variable (correlated part) and the realization,  $u_2$ , of the standard normal variable,  $U_2$ , (uncorrelated part):

$$w = v + \delta + \Lambda_{\sigma(v)}^{-1}(\Phi(u_2)) \tag{3.21}$$

This is essentially the description of function  $G$  of formula (3.17), see Figure 3.9. A schematic showing the relationship between  $w$ ,  $v$ , and  $u_2$  is shown in Figure 3.10. The dashed line shows the correlated part of the relationship between  $v$  and  $w$ . The distribution around that line, in this specific case illustrated as a normal distribution, is also shown for two selected values of  $v$ . Note that in Figure 3.10 the standard deviation can vary with  $v$ . That is, for increasing values of  $v$ , the standard deviation of  $w$  values around  $v$  can increase or decrease.

In order to establish a constant variance over  $v$ , the function  $\sigma(v)$  is simply set equal to a constant. The relationship  $\sigma(v)$  needs to be determined apriori and included as input to the correlation model. The procedure to determine  $\sigma(v)$  is an iterative one, and is described in Diermanse & Geerse [2012].

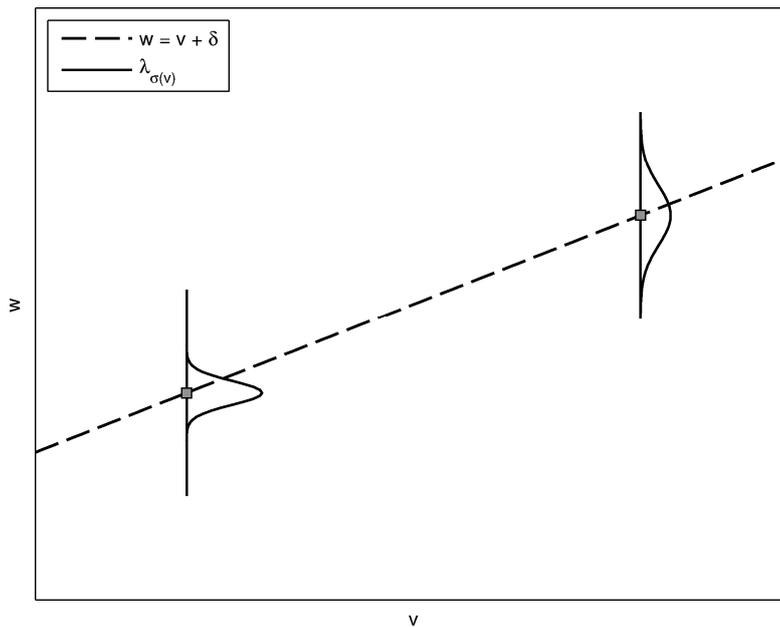


Figure 3.10 Schematic illustrating the relationship between  $w$ ,  $v$ , and  $u_2$  in the NL model.

As stated above, equation (3.21) is essentially the description of function  $G$  of Figure 3.9., Function  $G$  relates the transformed dependent variable  $W$  to the transformed independent variable  $V$ . The next step in the correlation model is to transform the variables  $V$  and  $W$  to standard uniform variables  $U_{1;corr}$  and  $U_{2;corr}$ . For this purpose, the distribution functions  $F_V(v)$  and  $F_W(w)$  are required.

Variable  $V$  is standard uniformly distributed, so the transformation from  $V$  to  $U_{1;corr}$  is done as follows:

$$u_{1;corr} = \Phi^{-1}(F_V(v)) = \Phi^{-1}(1 - e^{-v}) \quad (3.22)$$

This transformation is the inverse of the transformation in equation (3.18), which means  $u_{1;corr}$  is exactly the same as  $u_1$ . For variable  $W$  the transformation to the corresponding variable  $U_{2;corr}$  is less straightforward, since there is generally no analytical description available of distribution function  $F_W(w)$ . This distribution can be derived from the following integral (according to the theorem of total probability):

$$F_w(w) = \int_0^{\infty} f_V(v) P[W < w | v] dv = \int_0^{\infty} \exp(-v) \Lambda_{\sigma(v)}(w - v - \delta) dv \quad (3.23)$$

In a probabilistic computation procedure, this integral has to be solved numerically for each realization  $w$ . In Hydra-Ring, for computational efficiency, the integral is evaluated prior to the probabilistic computation for the whole range of  $w$ -values. The result is a table of  $w$ -values and associated probabilities, which serves as a lookup-table during the probabilistic computation. The transformation from  $W$  to  $U_{2,corr}$  is done as follows:

$$u_{2,corr} = \Phi^{-1}(F_w(w)) \quad (3.24)$$

### 3.4.2.3 Example 2: correlation model PCR

The correlation model referred to as PCR is given this naming convention because it is the correlation model which was developed and used as the standard correlation model in PC-Ring, i.e. the predecessor of Hydra-Ring. The model contains strong similarities with the HES, except that it includes a simplifying approximation for reasons of computation efficiency. The simplification lies in the fact that variable  $w$  assumed to be exponentially distributed, which means the integral of equation (3.23) is left out of the process. The reduction in computation time was valuable in the period that PC-Ring was developed, but with current computation power this is not the case anymore. The main reason to implement this model in Hydra-Ring as well is to be able to compare results of Hydra-Ring with PC-Ring.

Just as with the HES model, the PCR model works with an independent variable  $V$  and a dependent variable  $W$ .  $V$  is once again standard normally distributed and realizations,  $v$ , of this variable are therefore derived according to equation (3.18). Similarly, realizations of variable  $W$  are derived according to equation (3.21). However, in this case the choices of parameter  $\delta$  and function  $\Lambda$  are related (whereas in the HES model these can be chosen independently):

$$w = v - \frac{\sigma^2}{2} + \sigma u_2 \quad (3.25)$$

In equation (3.25),  $\sigma$  is the parameter of the correlation model and gives the measure of correlation. Small values of  $\sigma$  correspond to high correlation, and large values of  $\sigma$  correspond to weak correlation. If we compare equations (3.21) and (3.25), it can be seen that the latter is a special case of the former. In equation (3.25),  $\Lambda$  is a normal distribution function with standard deviation  $\sigma$ , while parameter  $\delta$  is equal to  $-\sigma^2/2$ . This shows  $\Lambda$  and  $\delta$  are related in this correlation model through the choice of parameter  $\sigma$ .

Figure 3.11 shows the assumed probability density function of variable  $W$  in the PCR model, i.e. the standard exponential density function. Furthermore, it shows the probability density function of the HES model,  $f_w(w)$ , in case of a constant standard deviation. As can be seen,  $f_w(w)$  converges to the standard exponential density function in the right tail. In other words: for larger values of  $w$ , variable  $W$  of the HES model with constant correlation is (asymptotically) standard exponentially distributed.

As just mentioned, in the PCR model the assumption is made that variable  $W$  is standard exponentially distributed over the entire domain. This means the transformation from variable  $w$  to the standard normal variable  $u_{2,corr}$  is done as follows:

$$u_{2,corr} = \Phi^{-1}(F_w(w)) = \Phi^{-1}(1 - e^{-w}) \quad (3.26)$$

This assumption is the reason why the PCR-model requires less computation time than the HES-model, since equation (3.26) is much easier to solve than the combination of equations (3.23) and (3.24). However, by making the assumption of a standard exponential distribution function, the PCR-model introduces an error, especially in the left tail, as can be seen in Figure 3.11. This error will propagate in the value of  $u_{2, \text{cor}}$  and eventually also in  $x_2$ , i.e. the realization of the associated “real-world” variable  $X_2$ . The reasoning behind the PCR model is that for failure computations only large values of load variables are relevant and therefore the error introduced in this correlation model is negligible. Unfortunately, this assumption does not hold in each case, because lower values of load variables can be relevant for failure as well. For instance at a location near the sea in a tidal river system, failure (flooding) most likely occurs in an event with extremely high sea water levels in combination with “average” river discharges. If the river discharge is estimated from the PCR model, significant errors may be introduced in estimating the probability of occurrence of such an event. It is therefore recommended to apply the HES model (section 3.4.2.2) instead of the PCR model, especially since the gain in computation time of the PCR model is very marginal with present day computers.

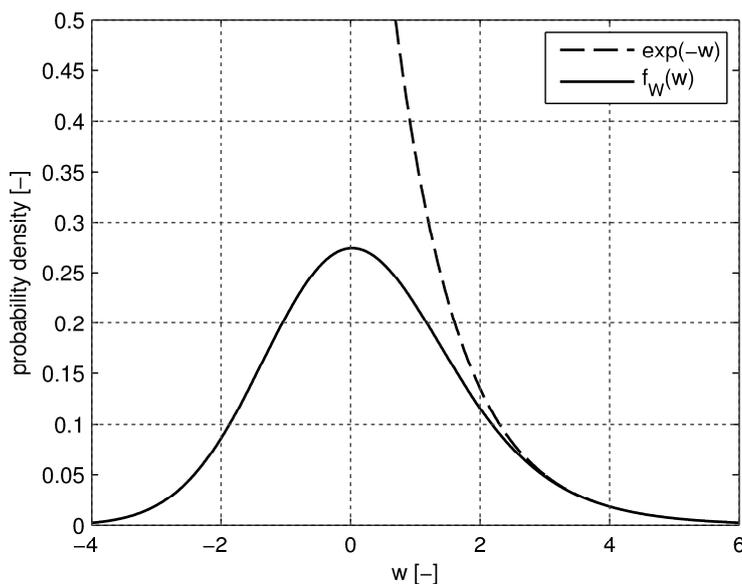


Figure 3.11 Probability density function of variable  $W$  in the PCR-model, compared with the density function of a standard exponential distribution function

#### 3.4.2.4 Other correlation models

Currently, the library of correlation models of Hydra-Ring contains two additional correlation models (additional to the HES model and the PCR model). However, these two models have been designed for one specific application in the Netherlands. They are set-up in a generic way and therefore can be applied to other areas as well, but the specific characteristics of these models make the need for such applications unlikely. These two correlation models are discussed in the sections where the specific load models are described (sections 4.5.1.4 and 4.10.1.2)

#### 3.4.3 Correlation in time and space

For strength parameters the correlation between subsequent time steps is assumed to be equal to 1. For “faster” varying variables like wind speed and sea water level, the correlation

between subsequent time steps is small enough to be neglected and is therefore taken equal to 0. For the slowly varying variables like river discharge and lake level, For slowly evolving variables like river discharge and lake water level a hybrid approach is used as described in section 2.5.2. For these variables, the year is considered to be divided into a number of periods. Within a single period, the autocorrelation in time for the slowly evolving variables is taken equal to 1, whereas the autocorrelation between different periods is taken equal to 0 (see section 2.5.2).

Correlation in space for load variables is generally much larger than for variables describing the strength of the flood defence. Within specified regions in the Netherlands, a 100% spatial correlation of random load variables is generally assumed. In between regions, spatial correlation is sometimes taken to be lower than 100% (see Table 4.41, chapter 4).

## 3.5 Hydrodynamic models

### 3.5.1 Simulating water levels and waves in an entire water system

The previous sections described the statistical properties of the load variables. Combinations of samples of load variables can be considered as synthetic events that may lead to failure of the flood defence system. For each event the resulting hydraulic load at the flood defences (water levels, waves) needs to be derived, in order to provide an estimate of the failure probability. This is generally achieved by executing a large number of hydrodynamic model simulations. In the hydrodynamic simulations, well-chosen realisations of the random variables are used as boundary conditions for the hydrodynamic model. The model simulations are executed prior to the application of Hydra-Ring and the relevant model results are stored as input databases for Hydra-Ring. These databases serve as a look-up table to link the random variables (input) to the hydraulic load at the flood defence (output).

For example, the water level in a tidal river is influenced by several factors, such as the upstream river discharge, the downstream sea water level, the local wind speed and the local wind direction. For such an area, hydrodynamic simulations need to be executed for all potential combinations of river discharge, sea water level, wind speed and wind direction that may lead to failure (flooding) somewhere along the tidal river area.

The events that are simulated should cover the extent of all events that are relevant for flood risk analysis. One of the main challenges in probabilistic flood analysis is to minimize the number of hydrodynamic simulations (to save computation time) and at the same time not to exclude the events and processes that are relevant for estimating flood risks. The choice of simulated events therefore requires sufficient knowledge of the system.

Examples of hydrodynamic models that are used to generate the input of Hydra-Ring are:

- 1 One-dimensional hydrodynamic models for predicting water levels in rivers;
- 2 Two-dimensional hydrodynamic models for predicting water levels in rivers, lakes and along the coast line;
- 3 Wave simulation models for prediction wave characteristics on lakes, rivers and along the coast line.

### 3.5.2 Models for quantification of local (corrections on) hydraulic loads

#### 3.5.2.1 Introduction

The hydrodynamic models as mentioned in section 3.5.1 generally produce water levels and wave conditions at locations at some distance from the toe of the flood defence. In some applications, additional models are required to translate conditions from the model output locations to the toe of the flood defence. Hydra-Ring contains various methods and models for this purpose:

- 1 The Bretschneider model for estimating local wave height as a function of the wind and the river geometry.
- 2 A model to compute wind set-up in rivers
- 3 The foreshore model for estimating the influence of a foreshore on waves;
- 4 The breakwater model estimating the influence of a breakwater on waves;
- 5 Spatial interpolation methods.

These models are relatively straightforward and therefore not very time consuming. They are built into the Hydra-Ring code and are executed multiple times during the probabilistic computations. The following sections elaborate on these models.

#### 3.5.2.2 Wave model Bretschneider

The Bretschneider formulas, which describe wave period and wave height, represent a simplified approach which is only valid under highly schematized conditions. However, given the simplicity of its use, it is often used in practice when a more sophisticated wave propagation model is unavailable.

The Bretschneider formulas rely on the following variables: a representative wind speed ( $U$ ), a representative water depth ( $D$ ), and a representative fetch length ( $F$ ). The formulas return the significant wave height ( $H_s$ ) and significant wave period ( $T_s$ ). In practice, it is typically the peak wave period ( $T_p$ ) that is the required output of load models. To convert from the significant wave period to the peak wave period, the following relationship is often sufficient in case of single-peak wave spectra:

$$T_p = 1.08 \cdot T_s \quad (3.27)$$

The Bretschneider formulas can be written in terms of dimensionless quantities; that is the wind speed  $U$ , water depth  $D$ , fetch length  $F$ , significant wave height  $H_s$  and significant wave period  $T_s$  can be made dimensionless as follows:

$$\tilde{D} = \frac{gD}{U^2} \quad (3.28)$$

$$\tilde{F} = \frac{gF}{U^2} \quad (3.29)$$

$$\tilde{H}_s = \frac{gH_s}{U^2} \quad (3.30)$$

$$\tilde{T}_s = \frac{gT_s}{U} \quad (3.31)$$

Where  $g$  represents the gravitational constant, with units  $\text{m/s}^2$ . The Bretschneider formulas for the dimensionless wave height and wave period are as follows:

$$\tilde{H}_s = A \tanh(k_3 \tilde{D}^{m_3}) \tanh\left(\frac{k_1 \tilde{F}^{m_1}}{\tanh(k_3 \tilde{D}^{m_3})}\right) \quad (3.32)$$

$$\tilde{T}_s = 2\pi \cdot B \tanh(k_4 \tilde{D}^{m_4}) \tanh\left(\frac{k_2 \tilde{F}^{m_2}}{\tanh(k_4 \tilde{D}^{m_4})}\right) \quad (3.33)$$

The empirical constants in equations (3.32) and (3.33) are given below:

A	=	0.283
B	=	1.2
$k_1$	=	0.0125
$k_2$	=	0.077
$k_3$	=	0.530
$k_4$	=	0.833
$m_1$	=	0.42
$m_2$	=	0.25
$m_3$	=	0.750
$m_4$	=	0.375

Note that to get the dimensional quantities  $H_s$  and  $T_s$  from  $\tilde{H}_s$  and  $\tilde{T}_s$ , the relationships in equations (3.30) and (3.31) can be used.

To account for the uncertainty (error) in the empirical model, the formulas for wave height and wave period are multiplied by model factors, which have a mean of one, and a standard deviation represented by the degree of uncertainty expected for the situation under consideration. The model factor for the wave height is denoted  $m_H$ , and the model factor for the wave period is denoted  $m_T$ . Furthermore, the formula for the wave height is multiplied by a wave reduction factor,  $K$ . The final formulas are thus as follows:

$$\tilde{H}_s = m_H \cdot K \cdot A \tanh(k_3 \tilde{D}^{m_3}) \tanh\left(\frac{k_1 \tilde{F}^{m_1}}{\tanh(k_3 \tilde{D}^{m_3})}\right) \quad (3.34)$$

$$\tilde{T}_s = 2\pi \cdot m_T \cdot B \tanh(k_4 \tilde{D}^{m_4}) \tanh\left(\frac{k_2 \tilde{F}^{m_2}}{\tanh(k_4 \tilde{D}^{m_4})}\right) \quad (3.35)$$

And thus (see equations (3.30) and (3.31)) the dimensional quantities  $H_s$  and  $T_s$  are equal to:

$$H_s = \frac{U^2}{g} \tilde{H}_s = \frac{U^2}{g} m_H \cdot K \cdot A \tanh(k_3 \tilde{D}^{m_3}) \tanh\left(\frac{k_1 \tilde{F}^{m_1}}{\tanh(k_3 \tilde{D}^{m_3})}\right) \quad (3.36)$$

$$T_s = \frac{U}{g} \tilde{T}_s = \frac{U}{g} 2\pi \cdot m_T \cdot B \tanh(k_4 \tilde{D}^{m_4}) \tanh\left(\frac{k_2 \tilde{F}^{m_2}}{\tanh(k_4 \tilde{D}^{m_4})}\right) \quad (3.37)$$

In tidal systems, the wave height and wave period can be influenced by swell, which refers to a long wave duration with a low wave height. To account for the influence of swell, the following corrections are made.

$$H_s = \sqrt{H_{s1}^2 + H_{s2}^2} \quad (3.38)$$

$$T_s = \sqrt[4]{\frac{T_{s1}^4 H_{s1}^2 + T_{s2}^4 H_{s2}^2}{H_s^2}} \quad (3.39)$$

where  $H_{s1}$  and  $T_{s1}$  are the significant wave height and wave period computed with the Bretschneider model (equations (3.36) and (3.37)), and  $H_{s2}$  and  $T_{s2}$  are the significant wave height and wave period of swell. These values have to be determined externally and added to Hydra-Ring through input databases. In non-tidal systems, no corrections are applied for swell.

When a shallow foreland is present, this limits the ability of the wave height to increase due to the breaking of waves. To take this physical limitation into account, the wave height computed by Bretschneider is subjected to the following constraint:

$$H_s \leq 0.5h_d \quad (3.40)$$

where  $h_d$  is the water depth at the toe of the flood defence.

The wind speed that is required as input for the Bretschneider formulas is the “open water wind speed”,  $u_{10}$ , at a height of 10 metres above water. In practice, sometimes only statistics of the potential wind speed,  $u$ , are available. In that case a translation from  $u$  to  $u_{10}$  is required to apply the Bretschneider formulas. Table 3.2 shows the relation between  $u$  and  $u_{10}$ . For values of  $u$  in between the tabulated values, linear interpolation is used to obtain  $u_{10}$ .

pot. wind m/s	10 m wind m/s	pot. wind m/s	10 m wind m/s	pot. wind m/s	10 m wind m/s
0	0.00	17	18.53	34	35.59
1	1.12	18	19.56	35	36.56
2	2.25	19	20.59	36	37.53
3	3.37	20	21.62	37	38.50
4	4.49	21	22.64	38	39.47
5	5.61	22	23.66	39	40.43
6	6.74	23	24.68	40	41.39
7	7.86	24	25.69	41	42.34
8	8.97	25	26.69	42	43.30
9	10.06	26	27.69	43	44.25
10	11.14	27	28.69	44	45.20
11	12.21	28	29.69	45	46.14
12	13.28	29	30.68	46	47.08
13	14.34	30	31.67	47	48.03
14	15.39	31	32.65	48	48.96
15	16.44	32	33.64	49	49.90
16	17.49	33	34.62	50	50.83

Table 3.2 Transformation of the potential wind speed  $u$  to the open water wind speed  $u_{10}$ .

### 3.5.2.3 Local wind set up in rivers

Local wind set up in a river leads to a difference in the water levels at the axis of the river and at the toe of the dike. This wind set-up can be estimated with the following formula (see de Waal, 1999 and Geerse, 2011):

$$\Delta h = \frac{\alpha u^2 F \cos \varphi}{d} \text{ voor } d / B > 0.001 \quad (3.41)$$

$$\Delta h = \sqrt{0.7 \cdot 10^{-6} u^2 F \cos \varphi + d^2} - d \text{ voor } d / B < 0.001$$

In which

- F = fetch length (m),
- D = water depth (m)
- B = width of the river (m)
- $\varphi$  = angle between the wind direction and the orientation of the dike
- $\alpha$  = model parameter

### 3.5.2.4 Foreshore model

Water levels and wave conditions can change substantially from the hydrodynamic model output location to the flood defence, particularly when there is a foreshore present. The effect of foreshore on the loads is that the water level increases due to wind effects and the wave height is further dampened. Figure 3.12 illustrates the effect that a foreshore (foreland) can have on the loads at the flood defence.

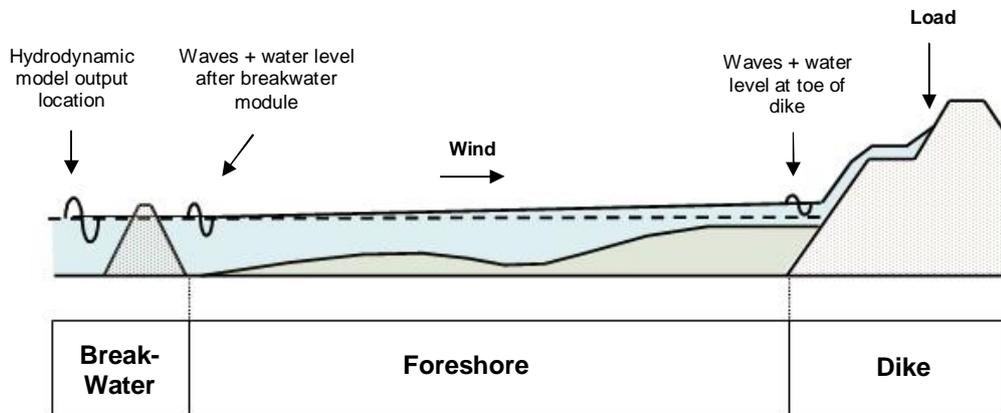


Figure 3.12 Illustration of the effect that the foreland can have on the load conditions at the toe of the dike. The horizontal dashed line indicates the water level without the presence of a foreland (Figure, adapted from De Waal, 1999).

A model to determine the impact of a foreland on the water level and wave conditions has been implemented in Hydra-Ring. Specific information about the models is not currently contained in this documentation but can be found in De Waal [1999].

If a foreshore, with varying bottom levels, is present in front of the dike, the wave conditions at the beginning of the foreshore have to be transformed to wave conditions at the toe of the dike. This transformation is done in the foreshore module with the model ENDEC (acronym for Energy Decay). ENDEC calculates changes in wave height and wave direction, but not of the wave period, due to:

- refraction (bending of waves as a consequence of changes in bottom levels),
- shoaling (changes in wave height due to changes in bottom levels),
- energy loss by breaking of waves,
- energy loss by bottom friction,
- energy gain caused by wave growth due to wind.

ENDEC also calculates the (usually small) changes of the water level as a consequence of wave set-up and wave set-down, whereas changes in the wave period (decrease by breaking or increase by wind) are neglected.

In the original ENDEC model, no changes of the water level caused by wind set-up were taken into account. For some situations this could lead to an underestimation of the water level and waves at the toe of the dike. At a later stage, therefore, wind set-up was implemented in the foreshore module. In the module first the wind set-up is calculated for the specified profile of the bottom. Afterwards, the ENDEC calculation is performed. When using ENDEC, it is important to note that the model is 1-dimensional, and therefore of limited accuracy.

### 3.5.3 Transformation module for a breakwater

A breakwater is a structure which is designed to reduce the wave height; the reduction is a function of the crest height of the structure and the shape of the structure (rubble mound,

caisson, vertical wall). Figure 3.12 illustrates the effect that a breakwater can have on the loads at the flood defence. The effect of a breakwater is that the wave height is dampened; there is no effect on the water level (i.e. the breakwater does not close of the area), wave period and wave direction.

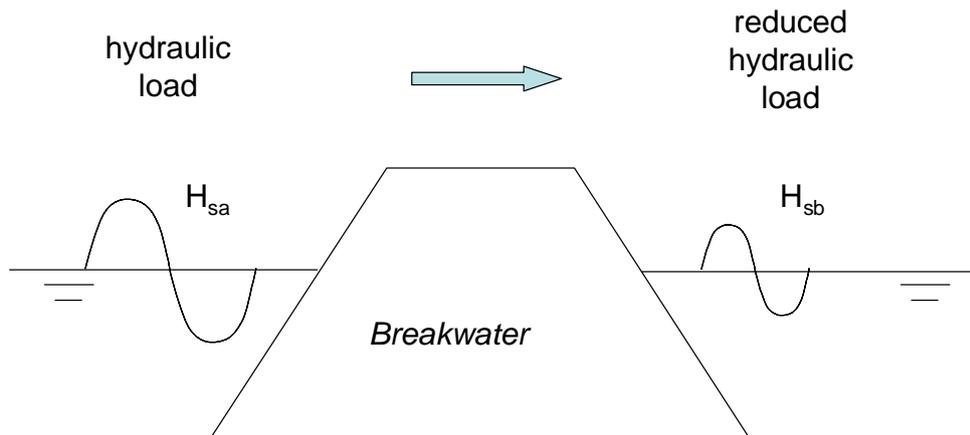


Figure 3.13 Schematic view of wave reduction, caused by a breakwater (figure adapted from De Waal, 1999)

The significant wave height directly behind the dam is assumed to be influenced by wave transmission across the dam only – wave penetration through holes in the dam is neglected. Strictly speaking, the latter contradicts the assumption of equal water levels on both sides of the breakwater in front and behind the dam if between the dam and the dike there is a closed basin and the dam has no holes (a non-porous dam). The actual assumption is: regarding the water level the dam is assumed to have openings, with equal water levels before and after, while regarding the waves the dam is assumed to have no openings. The reduction of the wave height is computed as follows (see de Waal, 1999 and Geerse, 2011):

$$H_{sb} = K_T H_{sa} \tag{3.42}$$

In which

- $H_{sa}$  = incoming significant wave height;
- $H_{sb}$  = reduced significant wave height;
- $K_T$  = reduction factor.

The reduction factor is derived as follows:

$$K_T = \begin{cases} 1 & ; \frac{z_d}{H_{sa}} \leq -\alpha_d - \beta_d \\ \frac{1}{2} \left[ 1 - \sin \left( \frac{\frac{z_d}{H_{sa}} + \beta_d}{\alpha_d} \frac{\pi}{2} \right) \right] & ; -\alpha_d - \beta_d \leq \frac{z_d}{H_{sa}} \leq \alpha_d - \beta_d \\ 0 & ; \frac{z_d}{H_{sa}} \geq \alpha_d - \beta_d \end{cases} \quad (3.43)$$

In which:

$z_d$  = the crest height relative to the water level ( $z_d < 0$  if the crest is under water)  
 $\alpha_d, \beta_d$  = parameters, dependent on the type of breakwater (see Table 3.3)

Figure 3.14 shows the different types of breakwaters and the respective values for  $\alpha_d$  and  $\beta_d$ . Figure 3.15 shows the value of the reduction factor  $K_T$  as a function of the ratio  $z_d/H_{sa}$ . This is the ratio between the crest height (relative to the water level) and the incoming wave height. As can be expected, the reduction factor is close to 1 (i.e. almost no reduction of the wave height) if the crest height is far below the still water level, whereas it is close to 0 (i.e. almost no wave height remaining) if the crest height is far above the still water level.

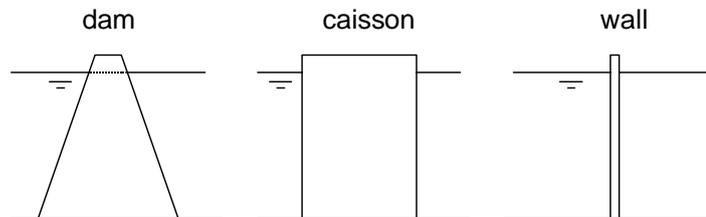


Figure 3.14 Types of breakwaters (figure taken from: Geerse, 2011).

Table 3.3 Parameters  $\alpha_d$  and  $\beta_d$  for different types of breakwaters

Type	$\alpha_d$	$\beta_d$
Dam	2.6	0.15
Caisson	2.2	0.40
wall	1.8	0.10

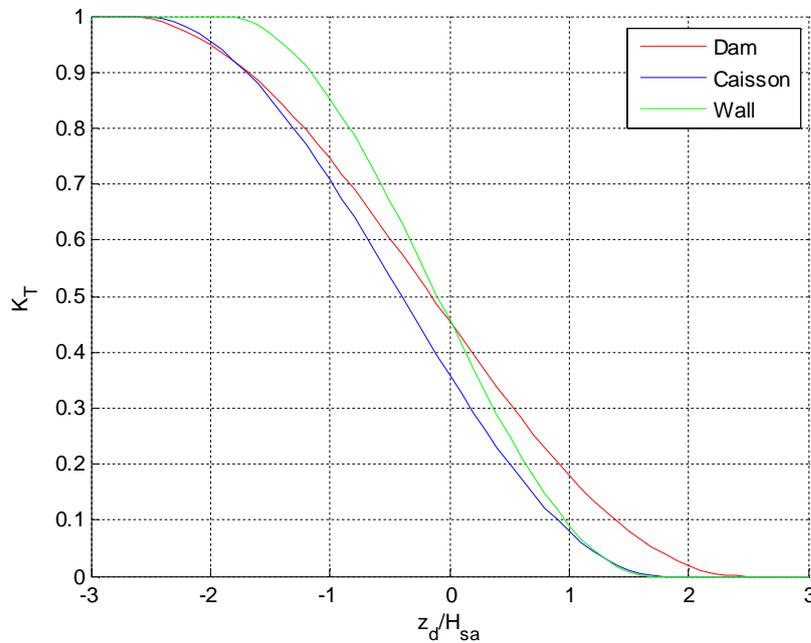


Figure 3.15 Reduction factor  $K_T$  as a function of the relative crest height  $z_d/H_{sa}$

### 3.5.4 Spatial interpolation methods

In hydraulic load models, statistics of load variables are initially available at a limited number of locations. These can either be measurement stations, in which case the statistics can be derived from measurement series directly, or output locations from hydrodynamic models in which case statistics are derived by means of probabilistic computations. In some cases, statistical information is required on locations where no such information is available. In that case, spatial interpolation methods may provide the solution. For water levels, spatial interpolation is generally more acceptable than for waves, since the spatial correlation lengths of waves are generally much smaller than for water levels.

In Hydra-Ring, the spatial interpolation is applied on water levels in river systems. For this purpose, the water levels are interpolated between two adjacent locations for which water level statistics are available. For sea level statistics, a triangular interpolation technique is applied, i.e. 2-dimensional interpolation on water levels of three locations for which water levels statistics are available. This method is applied to be able to model spatial variation of water levels in two dimensions. The triangular interpolation method is explained in more detail in section 4.8.2.1.

## 4 Hydraulic load models for primary water systems in the Netherlands

### 4.1 Introduction

Chapter 3 described the generic setup of hydraulic load models in Hydra-Ring. The current chapter describes the load models that are implemented for the primary water systems in the Netherlands. Table 4.1 shows the regions for the Netherlands that are currently implemented in Hydra-Ring. The map of Figure 4.1 serves as a reference. The load models of these regions will be described in more detail in sections 4.5-4.10. There are 16 load models for the Dutch primary waters in Hydra-Ring, which means 16 different combinations of random variables, correlation models and hydrodynamic models. The last column of Table 4.1 shows that the 16 regions can be clustered into 7 types. Load models of regions that are of the same type have a similar set-up and only differ in the details. In most cases the mutual differences between two load models of the same type are only in the choice of stations for which statistics have been derived that are used to describe the random variables. Sections 4.5-4.10 therefore describes the 7 types of load models.

More background on the derivation of statistics and implementation hydraulic load models for primary water systems in the Netherlands, the interested reader is referred to:

- Coastal systems: Den Heijer et al, [2007] Gautier, en Groeneweg, [2011]
- Inland water systems: Berger, 2007, Geerse, C.P.M [2011]
- Impelmentation in PC-Ring (predecessor of Hydra-Ring) : Diermanse et al [2003], Thonus et al [2003]

Table 4.1 Overview of regions in the TMR2006 database

number	Region	Load model type
1	Rhine branches, non-tidal zone	Upper river, non-tidal
2	Meuse river non-tidal zoner	Upper river, non-tidal
3	Rhine-Meuse delta, tidal area, Rhine dominated	Rhine-Meuse delta; tidal influences
4	Rhine-Meuse delta, tidal area, Meuse dominated	Rhine-Meuse delta; tidal influences
5	IJssel delta	Lake delta
6	Vecht delta	Lake delta
7	IJssel Lake	Lake
8	Marker Lake	Lake
9	Wadden Sea east	Coastal; sea dikes
10	Wadden Sea west	Coastal; sea dikes
11	North Sea coast; north	Coastal; sea dikes
12	North Sea coast; middle	Coastal; sea dikes
13	North Sea coast; south	Coastal; sea dikes
14	Oosterschelde	Coastal; sea dikes; barrier included
15	Westerschelde	Coastal; sea dikes
16	Dunes	Coastal; dunes

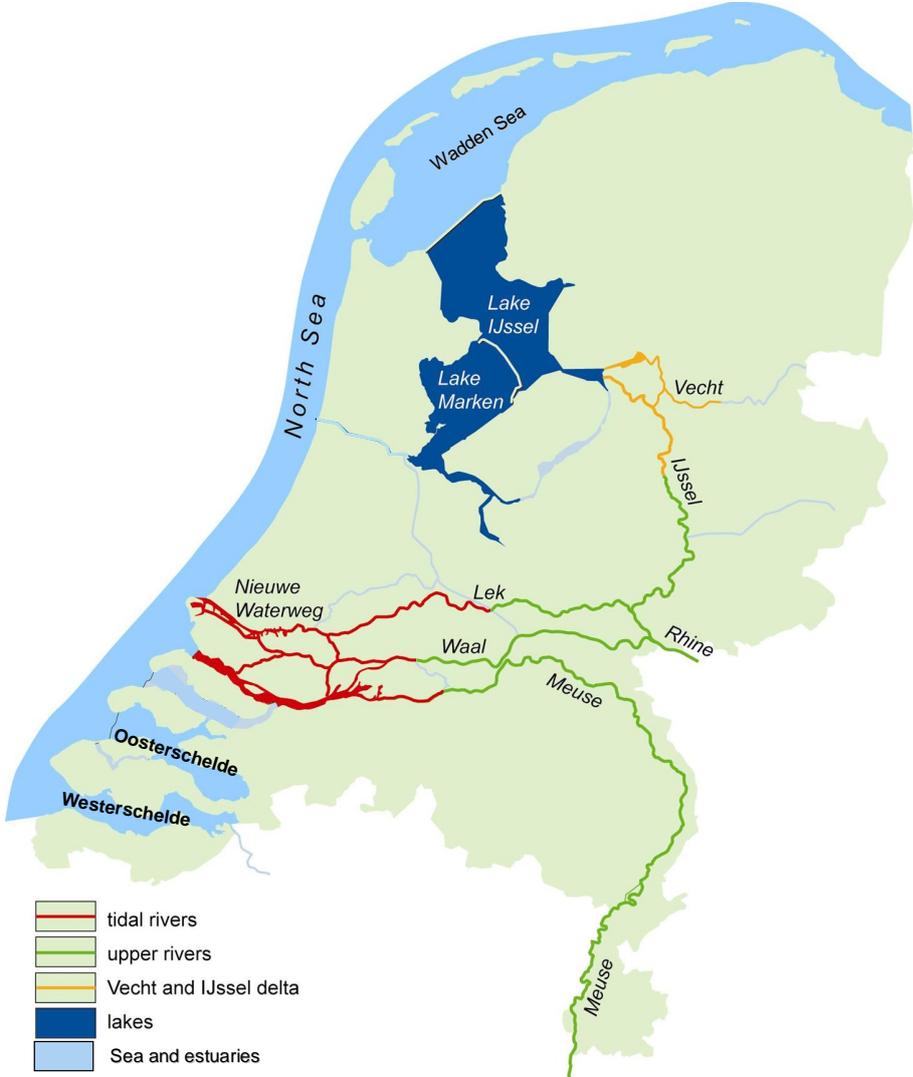


Figure 4.1 Primary water systems of the Netherlands

## 4.2 System characteristics

### 4.2.1 Upper rivers

The upper rivers (green lines in Figure 4.1) refer to the stretches of the river Meuse and the Rhine branches (Waal, Lek, IJssel) in which there is no influence of tide or backwater effects from the lakes. The threats for the dikes in this water system primarily result from high discharges, causing high water levels in the Rhine and Meuse rivers. Additionally, wind waves may cause wave run-up during events of high water. River discharge and wind (speed and direction) are therefore the only random variables of the load models for the upper rivers.

### 4.2.2 Rhine Meuse delta

This water system (red lines in Figure 4.1) refers to the stretches of the river Meuse and the Rhine branches in which water levels are influenced by the tide and storm surge in the North Sea. For this reason, a tidal barrier has been constructed near the river mouth. The barrier closes when high water levels at sea are predicted, to prevent high water levels in the river branches. In the load model, it is taken into account that the barrier can fail to close due to e.g. mechanical failures or human errors. Furthermore, it is taken into account that predicted water levels, required to determine if the barrier needs to be closed, contain uncertainties. Additional random variables in the model are river discharge, sea water level and wind (speed and direction). Since high storm surges in the North Sea are caused by high wind speed, in combination with westerly or northerly wind directions, the correlation between wind and sea water level is a relevant aspect of this load model as well.

### 4.2.3 Lakes

The IJssel Lake and Marker Lake (dark blue areas in Figure 4.1) are separated from each other by the dike, The "Houtrib dike". The inflow of the lakes consist of water from rivers, such as the IJssel and Vecht, but also from polder systems. The IJssel lake is separated from the sea by another dike, the Afsluitdijk. High lake levels in the IJssel Lake are the result of extended periods during which the discharge of the IJssel river exceeds the outflow through the sluices of the Afsluitdijk into the sea. Note that the Lake level refers to the spatially averaged water level; water levels in the lake can vary spatially due to effects of wind set-up. High water levels in the Lakes are therefore caused by a combination of high (spatially averaged) lake levels and local wind set-up. Wind waves causing wave run-up add to the hydraulic load.

Under "normal" conditions water from the IJssel lake is discharged into the sea by gravity, during periods of low tide. In case of high wind set-up at sea, this process may be hampered which causes the lake level to rise. The same holds for the Marker lake, which under normal conditions discharges its water into the North sea (it is connected by the North sea through a canal).

### 4.2.4 Lake delta

The rivers Vecht and IJssel (orange lines in Figure 4.1) are tributaries of the IJssel lake. The downstream stretches of these rivers are therefore referred to as the lake delta or, alternatively, the IJssel and Vecht delta. As stated in the previous section, high water levels in the IJssel Lake are the result of extended periods during which the discharge of the IJssel river exceeds the outflow through the sluices of the Afsluitdijk into the sea. As a result, peak discharges of the IJssel and Vecht rivers on one hand, and peak levels of the IJssel lake on the other hand are correlated. This is relevant for dikes along the IJssel delta, because increased lake levels will lead to increased water levels in the river due to backwater effects.

The correlation between water levels in the lake and river discharges is therefore taken into account in the load model. The load model is further complicated by the fact that there is a barrier in the river mouth of the Vecht that closes when water levels in the lake near the river mouth are too high, due to a combination of high lake levels and additional wind set-up. Similar to the load model of the Rhine-Meuse delta, it is taken into account that this barrier may fail to close due to e.g. mechanical failures or human errors.

#### 4.2.5 Coastal zone; dikes

The coastal zone is depicted by the light blue area in Figure 4.1. High water levels along the coast are a combination of high tides and storm surge. For most locations along the coast the surge is the dominant factor in causing high water levels. High storm surges along the coast are caused by strong westerly or northerly winds. In addition, these winds create waves that further increase the load on the flood defences. Wind and sea water level are the main random variables in these load models. The correlation between wind speed and sea water level is taken into account. Waves are no random variables in coastal load models as they are determined from hydrodynamic load models and as such deterministically related to the random variables water level, wind direction and wind speed. The exceptions to this rule are the regions along the Dutch North sea coast for which the wave period is included as an additional random variable.

#### 4.2.6 Oosterschelde

The Oosterschelde is an estuary in the southwestern part of the Netherlands. It is separated from the sea by a barrier. The barrier closes when high water levels at sea are predicted, to prevent high water levels in the estuary. The effect of the closure of the barrier on water levels in the Oosterschelde complicates the description of hydraulic loads and therefore a separate load model for this area is required. Similar to the load models of the lake delta and the Rhine-Meuse delta, it is taken into account that the barrier can fail to close due to e.g. mechanical failures or human errors. Furthermore, it is also taken into account that predicted water levels, required to determine if the barrier needs to be closed, contain uncertainties. The manner in which the closure of the barrier influences water levels in the Oosterschelde is further influenced by additional variables, such as the phase difference between the storm surge peak and the peak of the tide and the storm surge duration.

#### 4.2.7 Coastal zone; dunes

For dunes a separate load model is created. This is mainly due to the fact that for dikes the load is determined for locations approximately 50-100 m in front of the dike, whereas for dunes this is still the "active zone", i.e. the zone in which large morphodynamical activities take place during storm events. The hydraulic loads for dunes are therefore determined further seaward in comparison with the load models for dikes. The random variables for the load model for dunes are the water level, the wave height and the peak wave period. Correlations between these random variables are also taken into account. The load model for dunes is the only model in which the wind is not taken explicitly into account as a random variable. The influence of wind is considered to be incorporated in the statistics of the water levels and waves.

### 4.3 Random variables

Table 4.1 presents an overview of all random variables in the load models. Note that variables like wind and water level have different statistical properties for different load models and even within a single load model there can be differences. This is explained in the subsequent sections.

Table 4.1 Overview of load variables in the various hydraulic load models.

Random variable	Upper rivers	Sea delta	Lake delta	Lake	Coastal dikes	Ooster-schelde	Coastal dunes
wind speed	x	x	x	x	x	x	
wind direction	x	x	x	x	x	x	
river discharge	x	x	x				
lake level			x	x			
sea level		x			x	x	x
wave height							x
wave period					x		x
functioning of the barrier		x	x			x	
water level predictions for barrier		x				x	
phase tidal peak-surge peak						x	
storm surge duration						x	

#### 4.4 Load model for the upper rivers

##### 4.4.1 Distribution functions of random variables

In the load model of the upper rivers, the following random variables are relevant:

- River discharge
- Wind (speed and direction)

##### 4.4.1.1 River discharge

For river discharge, statistics of three locations are used in the load models of the upper rivers:

- 1 The river Rhine at Lobith;
- 2 The river Meuse at Borgharen; and
- 3 The river Meuse at Lith

Station Lith is located downstream of station Borgharen. For all dike sections upstream of Cuijk, statistics of station Borgharen are used to determine the hydraulic load. For all dike sections downstream of Cuijk, statistics of station Lith are used to determine the hydraulic load. For each station, three types of statistics are used in Hydra-Ring:

- [1] Exceedance frequencies of peak discharges;
- [2] Relationship between the peak value and peak duration of a discharge wave (see section 3.3.4) and
- [3] A duration curve that describes the mean duration of an exceedance of discharge  $Q$  (see section 3.3.4).

Statistics on [1] and [2] are input as tabulated values, see Table 4.2-Table 4.4 (see Geerse, 2008/2012 for a description of how the statistics have been derived). For the duration curve, a third degree polynomial function is used:

$$N(Q) = aQ^3 + bQ^2 + cQ + d \quad (3.1)$$

Where Q is the discharge and N is the average number of days that Q is exceeded during a single exceedance event. The parameters a-d are displayed in Table 4.5

Table 4.2 - Lobith parameters for the exceedance probability and peak durations

LOBITH			
Exceedance Probability of peak discharge		Peak duration	
Q [m <sup>3</sup> /s]	probability [1 / (30 days)]	Q [m <sup>3</sup> /s]	Peak duration [hours]
750	1	750	720
1000	0.97	6000	12
1500	0.8	30000	12
3500	0.3	-----	-----
4500	0.22	-----	-----
5893.3	0.16667	-----	-----
7017	0.083333	-----	-----
10850	0.0066667	-----	-----
16000	0.00013333	-----	-----
18000*	0.00002918	-----	-----

Table 4.3 - Lith parameters for the exceedance probability and peak durations

LITH			
Exceedance Probability of peak discharge		Peak duration	
Q [m <sup>3</sup> /s]	probability [1 / (30 days)]	Q [m <sup>3</sup> /s]	Peak duration [hours]
0	1	0	720
75	0.995	1315.1	12
200	0.88	6000	12
300	0.76	-----	-----
500	0.55	-----	-----
1315.1	0.16667	-----	-----
3652	0.00013333	-----	-----

Table 4.4 - Borgharen parameters for the exceedance probability and peak durations

Borgharen			
Exceedance Probability of peak discharge		Peak duration	
Q [m <sup>3</sup> /s]	probability [1 / (30 days)]	Q [m <sup>3</sup> /s]	Peak duration [hours]
0	1	0	720
76.756	0.995	1315.1	12
204.683	0.88	6000	12
307.025	0.76	-----	-----
511.708	0.55	-----	-----
1350.15	0.16667	-----	-----
4020.68	0.000133	-----	-----

Table 4.5 - Duration curve parameters (FBC model)

Station	Duration curve parameters			
	a	b	c	d
Lobith (Rhine)	-7.150x10 <sup>-13</sup>	5.872x10 <sup>-8</sup>	-1.624x10 <sup>-3</sup>	18.821
Borgharen (Meuse)	-8.387x10 <sup>-11</sup>	1.12x10 <sup>-6</sup>	-5.493x10 <sup>-3</sup>	13.75

Station	Duration curve parameters			
Lith (Meuse)	$-8.387 \times 10^{-11}$	$1.12 \times 10^{-6}$	$-5.493 \times 10^{-3}$	13.75

#### 4.4.1.2 Wind

The statistics of the wind are described in section 4.5.1.3

#### 4.4.1.3 Correlation between Rhine and Meuse discharge

In the non-tidal area, upstream of the confluence of Rhine and Meuse, water levels are determined only by either the Rhine or the Meuse discharge. Nevertheless, the correlation between Rhine and Meuse discharges is still relevant in this area. This is because a number of dike rings are bordered by the Meuse in the South and by the Rhine branches in the North, which means the correlation between the river discharges influences the total failure probability of the dike ring.

The correlation between Rhine and Meuse discharges can be described by different correlation models. The PCR-model of section 3.4.2.3 can be used, in which parameter  $\sigma$  is taken equal to 1 (based on the analysis of Thonus and Diermanse, 2003). The preferred option, however, is to use the HES-model of section 3.4.2.2, since that model provides more reliable results for the entire range of discharges, whereas the PCR-model is only reliable for high (extreme) river discharges. In order to make the HES model compatible with the PCR-model for extreme discharges, the associated distribution function,  $\Lambda$ , for the independent part should be normally distributed with  $\sigma$  taken equal to 1. Parameter  $\delta$  should be taken equal to -0.5.

#### 4.4.2 Hydrodynamic models

For the upper river area, model simulations were carried out with the 2-dimensional hydrodynamic model WAQUA. The simulations were carried out for 9 discharges, covering the whole range from average discharges to extreme discharges. The simulations have been carried out for the rivers Rhine and Meuse separately. The output consists of water levels and results are available approximately every kilometer along the river axis. For dike sections inbetween these location, the local water level is computed by means of spatial interpolation, using the available water levels of the closest upstream and downstream WAQUA model output location. As such, the relation between discharge and water level is established.

For discharges other than the 9 simulated discharges, the associated water level is based on linear interpolation. For example, if the WAQUA simulations show that at location L the water level is equal to 10 m+NAP for a discharge of 16000 m<sup>3</sup>/s and equal to a water level of 11 m+NAP for a discharge of 18000 m<sup>3</sup>/s, then the water level for a discharge of 17000 m<sup>3</sup>/s is assumed to be equal to 10.5 m+NAP and the water level for a discharge of 16500 m<sup>3</sup>/s is assumed to be equal to 10.25 m+NAP.

##### 4.4.2.1 Water level corrections

The model of section 3.5.2.3 is used to compute local wind set-up. The resulting value is added to the water level. Furthermore, an additional correction is added to take into account the effect of various kinds of oscillations in the water level caused by wind (e.g. seiches). These increments are not computed, but read from the input database.

##### 4.4.2.2 Wave loads

In order to compute wave loads in the river system, the Brettschneider model of section 3.5.2.2 is applied.

## 4.5 Load model for the tidal rivers

### 4.5.1 Distribution functions of random variables

In the load model of the tidal river system, the following random variables are relevant:

- River discharge
- Wind (speed and direction)
- Sea water level
- Functioning of the barrier and water level predictions for the barrier

#### 4.5.1.1 River discharge

See section 4.4.1.1

#### 4.5.1.2 Sea water level at Maasmond (Hoek van Holland)

Location Maasmond is the downstream boundary for the (tidal) river area. The sea water level statistics at location Maasmond are described by a conditional Weibull distribution function (see section 6.13). The parameters of this distribution function are shown in Table 4.6. The report of RIKZ [2006] describes these statistics. The parameters are given per wind direction, showing the dependency between sea water level and wind direction. The time scale for which these parameters have been derived is the tidal scale.

Table 4.6 Parameters of the conditional Weibull distribution for sea water levels at location Maasmond/Hoek van Holland. The associated time scale is a tidal period.

Wind Direction	$\omega$	$\lambda$	$\alpha$	$\xi$
N	1.6208	6.4508E-02	7.8380E-01	8.6598E-02
NNE	1.4537	7.9793E-02	5.0180E-01	1.9821E-03
NE	1.4357	7.8015E-02	4.9832E-01	1.8556E-03
ENE	1.4289	7.2644E-02	4.9798E-01	1.8431E-03
E	1.4328	7.8260E-02	4.9808E-01	1.8467E-03
EZE	1.4618	8.2783E-02	4.9740E-01	1.8226E-03
ZE	1.4703	7.1444E-02	4.9732E-01	1.8192E-03
ZZE	1.4526	6.4981E-02	4.9636E-01	1.7855E-03
Z	1.4365	5.9367E-02	4.9563E-01	1.7601E-03
ZZW	1.1679	6.7515E-01	5.0593E-01	2.1403E-03
ZW	1.3383	7.8215E-01	8.4245E-01	5.8747E-02
WZW	1.3006	1.3774E+00	8.3113E-01	7.5904E-02
W	1.4478	8.2036E-01	7.8752E-01	8.1759E-02
WNW	1.4260	9.1612E-01	7.7652E-01	9.0158E-02
NW	1.2602	8.9691E-01	7.3640E-01	9.0069E-02
NNW	2.1220	2.1788E-02	7.5288E-01	8.9100E-02

#### 4.5.1.3 Wind

A modified Gumbel distribution function is applied to describe the wind speed, see section 6.14 for a generic description of this function. The modified Gumbel distribution is fully determined by a quadratic polynomial function  $K_r$  and its parameters  $a$ ,  $b$  and  $c$  (see Table 4.8 and Table 4.9). Furthermore, the modified Gumbel is truncated with a factor  $d=2\%$ , see section 6.15 for truncation of the Gumbel distribution function.

For various conditions, the frequency of exceedance of the wind speed, as derived from the data, is multiplied with a so-called "Volker factor", which is equal to 0.5. This factor is applied to take into account the fact that the peak of the storm surge can coincide with low tide and therefore may not lead to flood risk in the river system. The Volker factor is only applied if each of the following three conditions are met:

- The considered wind direction is a northerly or westerly wind direction;
- The considered failure mechanism is either overtopping, overflow, uplift, piping or stability; and;
- The considered location is in the tidal river area.

Table 4.7 shows the probability of occurrence for each wind direction sector for two stations: Schiphol and Deelen. Table 4.8 and Table 4.9 present the modified Gumbel parameters of these same locations (see e.g. Volker, 1987, . For station Schiphol, the parameters are shown for the cases with Volker factor (left) and without (right) Volkerfactor. For station Deelen, there is no such distinction because the statistics of this station are applied in the upper river area only.

Table 4.7 Probability of occurrence for each wind direction sector

Wind Direction	Schiphol	Deelen
N	4.714E-02	1.751E-02
NNE	4.524E-02	2.531E-02
NE	5.573E-02	4.682E-02
ENE	6.442E-02	6.097E-02
E	5.753E-02	6.581E-02
EZE	4.145E-02	6.264E-02
ZE	4.444E-02	6.230E-02
ZZE	5.822E-02	5.856E-02
Z	7.450E-02	7.326E-02
ZZW	9.068E-02	1.157E-01
ZW	9.598E-02	1.321E-01
WZW	9.088E-02	1.115E-01
W	7.590E-02	6.369E-02
WNW	5.753E-02	4.322E-02
NW	5.083E-02	3.523E-02
NNW	4.954E-02	2.530E-02

Table 4.8 Parameters for the quadratic function  $K_r(u)$  for location Schiphol (Thonus et al, 2003)

	With Volkerfactor	Without Volkerfactor

Wind Direction	<i>a</i> [s <sup>2</sup> /m <sup>2</sup> ]	<i>b</i> [s/m]	<i>c</i> [-]	<i>a</i> [s <sup>2</sup> /m <sup>2</sup> ]	<i>b</i> [s/m]	<i>c</i> [-]
N	7.292E-04	2.536E-01	-6.351E-01	1.989E-02	7.763E-02	-8.213E-01
NNE	1.960E-02	1.654E-01	-1.461E+00	1.960E-02	1.654E-01	-1.461E+00
NE	2.888E-02	0.000E+00	-1.248E+00	2.888E-02	0.000E+00	-1.248E+00
ENE	2.889E-02	0.000E+00	-1.282E+00	2.889E-02	0.000E+00	-1.282E+00
E	2.895E-02	2.724E-02	-1.058E+00	2.895E-02	2.724E-02	-1.058E+00
EZE	2.802E-02	1.013E-01	-1.160E+00	2.801E-02	1.013E-01	-1.160E+00
ZE	2.935E-02	6.514E-02	-1.168E+00	2.935E-02	6.515E-02	-1.168E+00
ZZE	1.969E-02	1.547E-01	-1.464E+00	1.969E-02	1.547E-01	-1.464E+00
Z	2.241E-02	0.000E+00	-8.438E-01	2.241E-02	0.000E+00	-8.438E-01
ZZW	1.099E-02	1.475E-01	-1.673E+00	1.099E-02	1.475E-01	-1.673E+00
ZW	5.896E-04	3.107E-01	-1.786E+00	8.264E-03	1.711E-01	-1.980E+00
WZW	1.143E-03	2.978E-01	-1.591E+00	6.746E-03	1.970E-01	-2.042E+00
W	9.162E-04	2.017E-01	-9.367E-01	5.186E-03	2.235E-01	-2.088E+00
WNW	2.413E-03	2.002E-01	-7.523E-01	5.497E-03	2.244E-01	-1.959E+00
NW	3.632E-03	2.223E-01	-4.522E-01	5.192E-03	2.613E-01	-2.008E+00
NNW	3.486E-03	2.522E-01	-5.565E-01	9.558E-03	2.162E-01	-1.637E+00

Table 4.9 Parameters for the quadratic function  $K_r(u)$  for location Deelen (Thonus et al, 2003)

Wind Direction	<i>a</i> [s <sup>2</sup> /m <sup>2</sup> ]	<i>b</i> [s/m]	<i>c</i> [-]
N	2.069E-02	0.0E+00	-1.000E-01
NNE	1.865E-02	1.804E-01	-1.365E+00
NE	2.826E-02	3.969E-02	-1.057E+00
ENE	2.747E-02	6.599E-02	-1.235E+00
E	3.202E-02	0.000E+00	-6.885E-01
EZE	3.304E-02	3.591E-02	-6.322E-01
ZE	2.762E-02	1.668E-01	-1.297E+00
ZZE	1.735E-02	2.340E-01	-1.636E+00
Z	1.956E-02	1.031E-01	-1.311E+00
ZZW	7.678E-03	2.487E-01	-1.795E+00
ZW	4.413E-03	3.058E-01	-2.235E+00
WZW	3.836E-03	2.877E-01	-2.151E+00
W	2.752E-03	2.779E-01	-1.970E+00
WNW	2.929E-03	2.664E-01	-1.594E+00
NW	2.638E-03	2.923E-01	-1.500E+00
NNW	6.784E-03	2.377E-01	-1.105E+00

#### 4.5.1.4 Correlation between wind and sea water level

High sea water levels in the North Sea are caused by combinations of high tide and high surge. In the North Sea, potential variations in the surge exceed the potential variations in the tidal levels, which means the surge is the dominant cause of high water levels. High surges in

the North Sea are caused by high wind speeds in combination with northerly or westerly directions. Therefore, the mutual correlation between wind (speed and direction) and sea water level needs to be taken into account for these wind directions. A case-specific correlation model is applied to describe the correlation between these two variables, as will be described below. The model is often referred to as the 'Volkermodel'.

Even though the wind causes high sea water levels, the statistical relation is described by taking the sea water level as the independent variable and the wind speed as the dependent variable. This means that for the sea water, the statistical distribution function of section 4.5.1.2 is used. Subsequently the wind speed,  $X_2$ , is computed as a function of sea water level,  $X_1$ , plus an additional random component. The form of the conditional probability distribution of  $X_2$  is as follows:

$$F_{X_2|X_1}(x_2 | x_1, r) = \frac{1}{1-d} \exp \left[ -\exp \left( \frac{-K_r(u) + \rho_r [x_1 - A_r] / B_r}{M_r} \right) \right] \quad (3.2)$$

Where:

- r = index for wind direction sectors
- A, B, M,  $\rho$  = parameters, see Table 4.10.
- K = quadratic function with parameters a, b and c (see Table 4.8 and Table 4.9)
- d = fraction of the upper tail of the distribution that is truncated (d=0.02)

Note that parameters A, B, M and  $\rho$  depend on the wind direction and are only given for westerly and north-westerly wind directions. For easterly wind directions there is no correlation between sea water level and wind speed, so for those directions the wind speed is derived from the univariate distribution function as described in section 4.5.1.3. This can also be realized by taking parameters A, B, M and  $\rho$  according to the values in the last row of Table 4.10.

Table 4.10 Parameters of the Volker correlation function for westerly wind directions.

Wind Direction (degrees)	$A_r$ (m)	$B_r$ (m)	$\rho_r$ [-]	$M_r$ [-]
N	0.873	0.236	0.356	0.67
ZW	1.227	0.122	0.506	1.00
WZW	1.230	0.169	0.605	1.00
W	1.224	0.228	0.477	0.67
WNW	1.195	0.262	0.613	0.67
NW	0.887	0.326	0.768	0.67
NNW	0.904	0.292	0.677	0.67
other	0	1	0	1

The conditional distribution in equation (3.2) gives the non-exceedance probability of the wind speed,  $X_2$ , given the realization of  $X_1$  (water level) and the wind direction r, and is in fact the correlation model. In the original description of the correlation model [Volker 1987], the values  $K_r(x_2)$  were constants, chosen discretely per value of wind speed and wind direction, such

that the marginal statistics (computed by taking the integral over the water level) agree with the independently computed marginal distribution of the wind speed. Since then, a 2<sup>nd</sup> degree polynomial has been fit to the  $K_r(x_2)$  values as a function of  $x_2$ . The relationship is expressed formally as follows:

$$K_r(x_2) = a_r x_2^2 + b_r x_2 + c_r \quad (3.3)$$

To compute the (correlated) value of the wind speed, the inverse of equation (3.3) needs to be calculated. Note that rearrangement of equation leads to:

$$K_r(x_2) = \rho(x_1 - A)/B - M \ln[-\ln[p \cdot (1 - d)]] \quad (3.4)$$

where  $p$  is the non-exceedance probability of the independent part of variable  $X_2$ , which is computed as:

$$p = \Phi(u_2) \quad (3.5)$$

Where  $\Phi$  is the standard normal distribution function and  $u_2$  is the realization of standard normal variable  $U_2$ , that represents the wind speed (see section 2.2.3 on the manner in which Hydra-Ring uses standard normally distributed variables to represent “real-world” variables).

Substituting the expression for  $K_r(x_2)$ , given in equation (3.4), into equation (3.5) and solving for  $x_2$  using the quadratic formula (see section 6.14 for a detailed explanation of this solution) gives the following result::

$$x_2 = \begin{cases} \frac{-b + \sqrt{b^2 - 4ac'}}{2a}, & (b^2 - 4ac') > 0 \\ \frac{-b}{2a}, & (b^2 - 4ac') < 0 \end{cases} \quad (3.6)$$

where  $a$  and  $b$  are parameters of the polynomial given in Table 4.8, and  $c'$  is given as follows:

$$c' = c - \rho \cdot (x_1 - A)/B + M \ln[-\ln[p \cdot (1 - d)]] \quad (3.7)$$

where  $c$  is the parameter of the polynomial given in Table 4.8. The final step is to convert the  $x_1$  and  $x_2$  values to (correlated) standard normally distributed  $u$ -values, to complete the generic procedure of correlation models in Hydra-Ring as described in section 3.4.2.1. This is done as follows:

$$u_{1,cor} = \Phi^{-1}(F_{x_1}(x_1)) \quad (3.8)$$

$$u_{2,cor} = \Phi^{-1}(F_{x_2}(x_2)) \quad (3.9)$$

The distribution functions  $F_{x_1}$  and  $F_{x_2}$  are the marginal distribution of the water level and the marginal distribution function of the wind, respectively (see sections 4.5.1.2 and 4.5.1.3).

The text above describes the current implementation of the Volker model in Hydra-Ring. Recent analysis showed that this model can be implemented in a more straightforward manner, reducing the number of parameters and making it more transparent. Details are described in [Diermanse en Geerse, 2011] en [Geerse, 2011b]. This alternative implementation is a candidate for future adaptations to Hydra-Ring.

#### 4.5.1.5 Correlation between Rhine and Meuse discharge

In the tidal area, downstream of the Rhine and Meuse branches, the water levels are influenced by both Rhine and Meuse discharges. For these locations, a 100% correlation between Rhine and Meuse discharges is assumed, an assumption that may lead to small overestimations of water levels. The correlation is incorporated in the database with hydrodynamic model results and therefore does not need to be taken into account in the statistical distribution functions. The concept is straightforward: each location in the tidal area is either labeled “Rhine-dominant” or “Meuse dominant”. For a Rhine dominant location, the statistics of the river Rhine at Lobith are derived and this discharge is used to find the corresponding water level in the hydrodynamic database. This water level is based on a simulation in which a discharge of the Meuse river was assumed, based on a deterministic relation with the Rhine discharge (hence: 100% correlation). For Meuse dominant locations a different database is constructed in which the Rhine discharge was taken as a deterministic function of the Meuse discharge.

#### 4.5.1.6 State and prediction error for the storm surge barrier

The method for incorporating the influence of storm surge barriers is described in section 2.5.4. The following input is required:

- 1 Description of the closure criterion
- 2 Probability distribution function for the prediction error (if the decision for opening or closure of the barrier is based on water level predictions).
- 3 Probabilities of success/failure for closure and opening of the barrier upon request.

Furthermore, water levels in the river system as a function of the relevant random variables (sea waterlevel, river discharge etc) need to be available for the situation of the open barrier and for the situation of the closed barrier. This aspect will be described in section 4.5.2.

The Maeslant barrier will be closed if the predicted water level at location Rotterdam exceeds 3 m+NAP or if the water level at location Dordrecht exceeds 2.9 m+NAP. The Z-function for the closure criterion can therefore be described as follows:

$$Z_b = 3.0 - (h_R + \varepsilon) \cup 2.9 - (h_D + \varepsilon) \quad (3.10)$$

In which  $h_R$  is the water level at Rotterdam,  $h_D$  is the water level at Dordrecht and  $\varepsilon$  is the water level prediction error that is assumed to be the same for both locations. Prediction error  $\varepsilon$  is assumed to be normally distributed with mean -0.09 m and standard deviation 0.18 m. This means that the predicted water level is on average 0.09 m lower than the actual water level. Table 4.11 shows the probability of success/failure for opening and closure upon request of the Maeslant barrier. As can be seen from this table, the probability that the barrier is closed during an ‘event’ in which the closure criterion is not met (3<sup>rd</sup> scenario) is equal to 0. This scenario can therefore be left out of the computations to save computation time.

Table 4.11 Probability of success/failure for opening and closure upon request of the Maeslant barrier.

Scenario	Description	Value
1	$P[\text{closed} \mid Z_b < 0]$	0.99
2	$P[\text{open} \mid Z_b < 0]$	0.01
3	$P[\text{closed} \mid Z_b \geq 0]$	0
4	$P[\text{open} \mid Z_b \geq 0]$	1

#### 4.5.2 Hydrodynamic models

The 1-dimensional Sobek model was used to compute water levels in the tidal river system for different combinations of river discharge (Rhine and Meuse), sea water level, wind speed, wind direction and barrier state (open or closed). The computed water levels are stored in Microsoft Access databases. Two databases are distinguished: one for locations where the Rhine is the dominant river and one for locations where the Meuse is the dominant river (see section 4.4.1.3). In both cases, simulations were done for 9 combinations of Rhine and Meuse discharges, but the selected combinations were mutually different. Results are available approximately every kilometer along the river axis.

##### 4.5.2.1 Water level corrections

The model of section 3.5.2.3 is used to compute local wind set-up. The resulting value is added to the water level. Furthermore, an additional correction is added to take into account the effect of various kinds of oscillations in the water level caused by wind (e.g. seiches). These increments are not computed, but read from the input database.

##### 4.5.2.2 Wave loads

In order to compute wave loads in the river system, the Brettschneider model of section 3.5.2.2 is applied.

## 4.6 Load model for the lakes

#### 4.6.1 Distribution functions of random variables

In the load model of the lakes (IJssel Lake and Marker Lake), the following random variables are relevant:

- Lake level (spatially averaged)
- Wind (speed and direction)

##### 4.6.1.1 Lake level (spatially averaged)

Similar to river discharges, three types of statistics are used for the lake levels in Hydra-Ring

- [1] Exceedance frequencies;
- [2] Relationship between the peak value and peak duration of a hydrograph (see section 3.3.4) and
- [3] A duration curve that describes the mean duration of an exceedance of lake level  $X$  (see section 3.3.4).

These statistics are summarized in Table 4.12 (source: Geerse, 2006), Table 4.13 (source: Geerse, 2008) and Table 4.14 (source: Diermanse et al, 2003<sup>2</sup>);

<sup>2</sup> In the formulas in section 4.2.2. of Diermanse et al [2003], parameters  $A$  and  $B$  have been incorrectly switched

Table 4.12 – Lake level exceedance probabilities and peak durations for the IJssel Lake

IJssel Lake			
Exceedance Probability		Peak duration	
X [m]	probability [1 / (30 days)]	X [m]	Peak duration [hours]
-0.40	1.000	-0.4	720
0.05	0.166667	0.05	96
0.40	0.016667	1.8	96
1.07	1.667E-05	-----	-----

Table 4.13 – Lake level exceedance probabilities and peak durations for the Marker Lake

Marker Lake			
Exceedance Probability		Peak duration	
X [m]	probability [1 / (60 days)]	X [m]	Peak duration [hours]
-0.40	1.0	-0.40	1440
-0.22	0.3333	-0.22	96
1.00	1.014e-006	1.80	96

Table 4.14 Duration curve parameters for the IJssel and Marker Lake.

IJssel Lake		Marker Lake	
X [m]	N(X)	X [m]	N(X)
-0.4	35.7	-0.4	29.0
0.084	9.5	0.613	0.9664
0.394	1.0	1.8	0.9664
1.8	1.0		

#### 4.6.1.2 Wind speed

For wind speed, the statistics of Schiphol are used, more specifically the option *without* Volker factor (see section 4.5.1.3 and the last three columns of Table 4.8). Because 12 wind directions are considered for the lakes, the statistics need to be provided per 30-degree wind sector. These values are presented in Table 4.15

Table 4.15 Parameters for the quadratic function  $K_r(u)$  for location Schiphol, 12 wind directions. The second column (P) shows the probability of occurrence per wind direction.

Wind Direction	P	Without Volkerfactor		
		a [s <sup>2</sup> /m <sup>2</sup> ]	b [s/m]	c [-]
N	4.53E-02	2.128E-02	2.460E-03	-8.315E-01
30	6.81E-02	2.744E-02	0.000E+00	-9.125E-01
60	8.12E-02	2.699E-02	0.000E+00	-9.982E-01
E	7.66E-02	2.831E-02	0.000E+00	-7.812E-01
120	6.95E-02	4.056E-02	-2.095E-01	1.481E-02
150	7.99E-02	3.344E-02	-9.576E-02	-6.279E-01
S	1.09E-01	2.122E-02	-1.124E-02	-7.794E-01
210	1.36E-01	1.326E-02	1.950E-02	-6.830E-01
240	1.35E-01	7.887E-03	1.169E-01	-1.003E+00

		<i>Without Volkerfactor</i>		
<b>Wind Direction</b>	<b>P</b>	<b>a</b> [s <sup>2</sup> /m <sup>2</sup> ]	<b>b</b> [s/m]	<b>c</b> [-]
W	1.02E-01	7.570E-03	9.265E-02	-9.224E-01
300	5.93E-02	8.231E-03	8.070E-02	-7.676E-01
330	3.70E-02	1.164E-02	9.876E-02	-1.507E+00

#### 4.6.2 Hydrodynamic models

The 2-dimensional hydrodynamic model WAQUA was used to derive water levels along the lake for several combinations of wind direction, wind speed and the (spatially averaged) lake level. For other combinations, the water level is estimated from 3-dimensional interpolation between wind direction, wind speed and lake level.

A correction is added to the water level to take into account the effect of various kinds of oscillations in the water level caused by wind (e.g. seiches). These increments are not computed, but read from the input database.

The wave simulation model HISWA was used to derive wave characteristics (height, period and direction) for the same combinations of wind direction, wind speed and lake level. Again, for other combination 3-dimensional interpolation is used.

### 4.7 Load model for the lake delta

#### 4.7.1 Distribution functions of random variables

In the load model of the lake delta, the following random variables are relevant:

- Lake level (spatially averaged)
- Wind (speed, duration and direction)
- River discharge
- Functioning of the barrier

##### 4.7.1.1 Lake level (spatially averaged)

For the lake delta, only the lake level of the IJssel Lake is relevant. Statistics of the IJssel lake level are described in section 4.6.1.1.

##### 4.7.1.2 Wind

For wind speed, the statistics of Schiphol are used, more specifically the option *without Volker factor* (see section 4.5.1.3 and the last three columns of Table 4.8). The potential variation in wind duration is also taken into account. The storm graph (wind as a function of time) is modelled by assuming a trapezium shape with a given base duration of 48 hours, and peak duration of 2 hours.

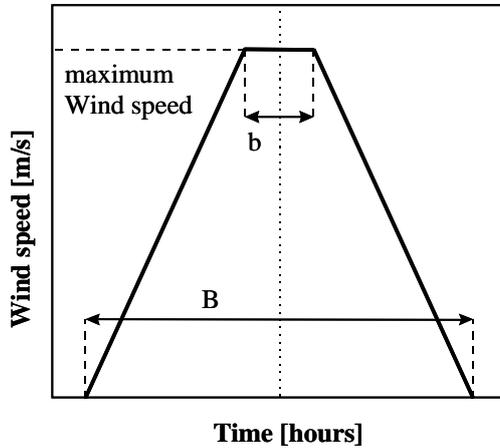


Figure 4.2 Schematisation of the hydrograph of the wind speed.

#### 4.7.1.3 River discharge

For river discharge, statistics of two locations are used in the load models of the lake delta:

- 1 The river IJssel at Olst; and
- 2 The river Vecht at Dalfsen;

For each station, three types of statistics are used in Hydra-Ring:

- [1] Exceedance frequencies of peak discharges;
- [2] Relationship between the peak value and peak duration of a discharge wave (see section 3.3.4) and
- [3] A duration curve that describes the mean duration of an exceedance of discharge  $Q$  (see section 3.3.4).

Statistics on [1] and [2] are input as tabulated values, see Table 4.16 - Table 4.17 (source: geerse, 2006) and Table 4.18 (source: Diermanse et al, 2003). For the duration curve, a third degree polynomial function is used:

$$N(Q) = aQ^3 + bQ^2 + cQ + d \quad (3.11)$$

Where  $Q$  is the discharge and  $N$  is the average number of days that  $Q$  is exceeded during a single exceedance event. The parameters a-d are displayed in Table 4.19.

Table 4.16 Parameters for the exceedance probability for Olst and Dalfsen (Hydra-Zoet version).

Exceedance Probability of peak discharge			
Olst		Dalfsen	
X [m <sup>3</sup> /s]	probability [1 / (30 days)]	X [m <sup>3</sup> /s]	probability [1 / (30 days)]
200	1.0	0	1.0
800	0.16666667	180	0.16666667
2720	1.3333E-04	550	1.3333E-04

Table 4.17 Parameters for the exceedance probability (PC-Ring version).

Dalfsen			
Exceedance Probability of peak discharge		Peak duration	
X [m <sup>3</sup> /s]	probability [1 / (year)]	X [m <sup>3</sup> /s]	probability [1 / (year)]
		0	
		180	
		550	

Table 4.18 Peak durations (hours) for Olst and Dalfsen.

Olst		Dalfsen	
X [m <sup>3</sup> /s]	Peak duration [hours]	X [m <sup>3</sup> /s]	Peak duration [hours]
200	720	0	720
800	24	180	48
4000	24	1000	48

Table 4.19 Duration curve parameters (FBC model)

Station	Duration curve parameters			
	a	b	c	d
X ≤ 1196	-----	-----	-1.33x10 <sup>-2</sup>	2.45x10 <sup>1</sup>
X ≥ 1196	-----	-----	-2.16x10 <sup>-3</sup>	1.12x10 <sup>1</sup>

#### 4.7.1.4 Correlation between Lake and river discharge

High water levels in the IJssel Lake are the result of extended periods during which the discharge of the IJssel river exceeds the outflow through the sluices of the Afsluitdijk into the sea. As a result, peak discharges of the IJssel river and peak levels of the IJssel lake are correlated. This is relevant for dikes along the IJssel delta, because increased lake levels will lead to increased water levels in the river due to backwater effects.

The correlation is described by different correlation models. The PCR-model of section 3.4.2.3 can be used, in which parameter  $\sigma$  is taken equal to 0.85 (based on the analysis of Thonus and Diermanse, 2003). The preferred option is to use the HES-model of section 3.4.2.2, since that model provides more reliable results for the entire range of discharges, whereas the PCR-model is only reliable for high (extreme) river discharges. In order to make the HES model compatible with the PCR-model for extreme discharges, the associated distribution function,  $\Lambda$ , for the independent part should be normally distributed with  $\sigma$  taken equal to 0.85. Parameter  $\delta$  should be taken equal to  $-(0.85)^2/2 = -0.361$ , see section 3.4.2.3. Note that in Geerse [2011], a value for  $\sigma=1.2$  is proposed.

#### 4.7.1.5 Correlation between IJssel and Vecht discharges

Some dike rings are bordered by both rivers (IJssel and Vecht), which means the correlation between the river discharges influences the total failure probability of the dike ring. Therefore, the correlation between IJssel and Vecht discharges is relevant. This correlation can be described by either the PCR-model or the HES model. **Nagaan welke  $\sigma$ -waarde in dit correlatiemodel gebruikt wordt. In VNK1 wordt gesproken van 100% correlatie**

#### 4.7.1.6 Functioning of the storm surge barrier

The Ramspol storm surge barrier has the purpose to protect the Lake delta during periods of high water level on the IJssel Lake. The barrier will be closed if the local water level, measured in the vicinity of the barrier, exceeds 0.50 m+NAP and at the same time the direction of the flow is from west to east, i.e. water flowing from Lake IJssel into the Vecht

delta. The barrier is re-opened as soon as the water level at the eastern side of the barrier exceeds the water level at the western side.

As opposed to the storm surge barrier at Hoek van Holland (see section 2.5.4) water level predictions and potential errors in water level predictions are no issue for the Ramspol barrier. Similar to the the storm surge barrier at Hoek van Holland, the possibility that the barrier fails to close upon request is taken into account. The failure probability is estimated at 1/1000 per closure request. (Note that in Hydra-Zoet, this number is equal to 3.5/1000).

#### 4.7.2 Hydrodynamic models

Model simulations have been carried out with the 2-dimensional hydrodynamic model WAQUA, for different combinations of river discharge, IJssel Lake level, wind direction, wind speed, wind duration and states of the storm surge barrier (open or close).

The model of section 3.5.2.3 is used to compute local wind set-up. The resulting value is added to the water level. Furthermore, an additional correction is added to take into account the effect of various kinds of oscillations in the water level caused by wind (e.g. seiches). These increments are not computed, but read from the input database.

In order to compute wave loads in the river system, the Brettschneider model of section 3.5.2.2 is applied.

### 4.8 Load model for the coastal dikes

#### 4.8.1 Distribution functions of random variables

In the load model of the coastal dikes, the following random variables are relevant:

- Wind (speed and direction)
- Sea water level
- Wave period

Each variable is considered a “slowly evolving variable” (see section 2.5.2). This means the (peak) values of these variables in subsequent tidal periods are assumed to be independent.

##### 4.8.1.1 Wind

Wind speed statistics are given per wind direction. They are described by the conditional Weibull function (see section 6.13). Parameters of 4 stations are presented in Table 4.21. Table 4.20 shows the representative stations for each coastal region.

Table 4.20 Selected wind stations per region

Region	station
Wadden Sea east	Terschelling West
Wadden Sea west	Terschelling West
North Sea coast; north	Texel, de Kooy
North Sea coast; middle	Hoek van Holland
North Sea coast; south	Hoek van Holland
Oosterschelde	Vlissingen
Westerschelde	Vlissingen

Table 4.21 Weibull parameters for wind stations

sector	$\omega$	$\lambda$	$\xi$	$\sigma$
Hoek van Holland				
0	14	0.015799	1.99	7.29
30	13	0.018632	2.11	6.84
60	14	0.012542	2.29	7.84
90	13	0.017007	2.31	7.71
120	13	0.016517	2.29	7.03
150	14	0.012449	2.4	8
180	16	0.011441	2.45	9.37
210	19	0.008507	2.4	11.1
240	20	0.008496	2.17	10.88
270	20	0.01101	1.98	10.09
300	18	0.015104	2.05	10.45
330	16	0.01789	2.05	8.66
Terschelling-west				
0	16	0.016702	2.1	8.86
30	15	0.017765	2.11	7.79
60	16	0.012194	2.28	8.85
90	15	0.016629	2.35	8.97
120	15	0.016517	2.34	8.28
150	16	0.013654	2.48	9.48
180	18	0.012013	2.49	10.65
210	21	0.008737	2.47	12.4
240	21	0.011251	2.37	13.13
270	22	0.01101	2.16	12.19
300	20	0.015104	2.12	11.74
330	18	0.019168	2.18	10.4
Vlissingen				
0	12	0.018327	1.83	5.98
30	12	0.016465	2.03	5.93
60	13	0.011462	2.19	6.81
90	12	0.015306	2.25	6.88
120	11	0.023174	2.38	6.77
150	12	0.016546	2.34	7.11
180	15	0.009782	2.38	8.52
210	18	0.008438	2.24	9.82
240	19	0.008794	2.03	9.73
270	19	0.010925	1.8	8.62
300	16	0.017738	1.91	9.38
330	14	0.017166	1.78	6.59
IJmuiden				
0	14	0.022119	2.03	7.95
30	14	0.016032	2.08	7.08
60	14	0.017768	2.29	8.28
90	14	0.015117	2.27	7.96
120	13	0.025026	2.33	7.63

sector	$\omega$	$\lambda$	$\xi$	$\sigma$
150	14	0.018875	2.42	8.53
180	17	0.010583	2.44	9.82
210	19	0.011496	2.43	11.73
240	20	0.011022	2.23	11.74
270	20	0.013833	2.04	10.94
300	18	0.018977	2.06	11.01
330	17	0.017038	2.07	9.21

#### 4.8.1.2 Sea water level

Sea water level statistics are given per wind direction. They are described by the conditional Weibull function (see section 6.13). Parameters of 12 stations are presented in Table 4.22. For each region, statistics of a subset of these twelve stations is used to describe the sea water level statistics for each dike section in the region. For this purpose, a spatial interpolation routine is used that is described in section 4.8.2.

Table 4.22 Weibull parameters for sea water level stations

sector	$\omega$	$\lambda$	$\xi$	$\sigma$
<b>Delfzijl</b>				
0	2.47	0.0065	1.59	0.8989
30	2.47	0.000087	1.51	0.3875
60	2.47	0.00007	1.51	0.3875
90	2.47	0.000076	1.51	0.3875
120	2.47	0.0001	1.51	0.3875
150	2.47	0.00008	1.51	0.3875
180	2.47	0.000057	1.51	0.3875
210	2.47	0.000345	1.52	0.459
240	2.47	0.006269	1.48	0.7251
270	2.47	0.039804	1.87	1.4973
300	2.47	0.061501	2.12	1.9565
330	2.47	0.031477	1.9	1.5905
<b>Den Oever</b>				
0	1.88	0.004198	1.65	0.9908
30	1.88	0.00013	1.38	0.2881
60	1.88	0.000105	1.38	0.2881
90	1.88	0.000113	1.38	0.2881
120	1.88	0.00015	1.38	0.2881
150	1.88	0.00012	1.38	0.2881
180	1.88	0.000086	1.38	0.2881
210	1.88	0.000161	1.28	0.3105
240	1.88	0.006383	1.49	0.7069
270	1.88	0.039042	1.9	1.2644
300	1.88	0.049572	1.97	1.4011
330	1.88	0.024918	2.12	1.6529
<b>Huibertgat</b>				

# Deltares

1206006-004-ZWS-0001, 2 January 2013, draft

sector	$\omega$	$\lambda$	$\xi$	$\sigma$
0	1.87	0.00325	1.03	0.1852
30	1.87	0.00026	1.2	0.1217
60	1.87	0.000209	1.2	0.1217
90	1.87	0.000227	1.2	0.1217
120	1.87	0.0003	1.2	0.1217
150	1.87	0.000241	1.2	0.1217
180	1.87	0.000172	1.2	0.1217
210	1.87	0.000621	1.17	0.1939
240	1.87	0.016877	1.23	0.4342
270	1.87	0.061286	1.31	0.5793
300	1.87	0.074785	1.31	0.5894
330	1.87	0.034289	1.28	0.5579
Den Helder				
0	1.61	0.002167	1.48	0.7068
30	1.61	0.000043	2.58	0.4276
60	1.61	0.000035	2.58	0.4276
90	1.61	0.000038	2.58	0.4276
120	1.61	0.00005	2.58	0.4276
150	1.61	0.00004	2.58	0.4276
180	1.61	0.000029	2.58	0.4276
210	1.61	0.000414	1.54	0.3277
240	1.61	0.009116	1.39	0.4665
270	1.61	0.038646	1.65	0.8849
300	1.61	0.04051	1.8	1.1206
330	1.61	0.015845	1.83	1.2168
Harlingen				
0	2.09	0.004379	1.73	1.023
30	2.09	0.00013	3.15	0.8338
60	2.09	0.000105	3.15	0.8338
90	2.09	0.000113	3.15	0.8338
120	2.09	0.00015	3.15	0.8338
150	2.09	0.00012	3.15	0.8338
180	2.09	0.000086	3.15	0.8338
210	2.09	0.000506	1.35	0.2546
240	2.09	0.01504	1.8	1.0739
270	2.09	0.0649	2.17	1.5361
300	2.09	0.07188	2.28	1.6958
330	2.09	0.032542	2.33	1.7917
Hansweert				
0	3.39	0.004649	1.15	0.337
30	3.39	0.00052	1.49	0.3207
60	3.39	0.000418	1.49	0.3207
90	3.39	0.000454	1.49	0.3207
120	3.39	0.000601	1.49	0.3207
150	3.39	0.000482	1.49	0.3207
180	3.39	0.000343	1.49	0.3207

sector	$\omega$	$\lambda$	$\xi$	$\sigma$
210	3.39	0.002368	1.24	0.3416
240	3.39	0.014535	1.2	0.3704
270	3.39	0.033847	1.2	0.4626
300	3.39	0.044421	1.16	0.5054
330	3.39	0.024918	1.04	0.3293
Hoek van Holland				
0	1.97	0.015935	0.8	0.0944
30	1.97	0.00078	1.07	0.1203
60	1.97	0.000627	1.07	0.1203
90	1.97	0.00068	1.07	0.1203
120	1.97	0.000901	1.07	0.1203
150	1.97	0.000723	1.07	0.1203
180	1.97	0.000515	1.07	0.1203
210	1.97	0.001311	0.93	0.0848
240	1.97	0.015339	0.84	0.0716
270	1.97	0.071703	0.79	0.0815
300	1.97	0.085706	0.71	0.0654
330	1.97	0.055288	0.74	0.0913
Lauwersoog				
0	2.12	0.006545	1.38	0.5488
30	2.12	0.00013	1.56	0.2196
60	2.12	0.000105	1.56	0.2196
90	2.12	0.000113	1.56	0.2196
120	2.12	0.00015	1.56	0.2196
150	2.12	0.00012	1.56	0.2196
180	2.12	0.000086	1.56	0.2196
210	2.12	0.000276	1.17	0.1527
240	2.12	0.010011	1.45	0.6257
270	2.12	0.059028	1.84	1.2438
300	2.12	0.088611	1.92	1.3652
330	2.12	0.048856	1.88	1.3309
OS11				
0	2.27	0.013903	1.14	0.332
30	2.27	0.001387	1.22	0.2071
60	2.27	0.001115	1.22	0.2071
90	2.27	0.001209	1.22	0.2071
120	2.27	0.001602	1.22	0.2071
150	2.27	0.001285	1.22	0.2071
180	2.27	0.000915	1.22	0.2071
210	2.27	0.002782	1.13	0.2127
240	2.27	0.019334	1.15	0.2709
270	2.27	0.065916	1.13	0.3329
300	2.27	0.082918	1.15	0.4112
330	2.27	0.051113	1.17	0.4555
Terschelling-west				
0	1.9	0.002302	1.75	0.932

sector	$\omega$	$\lambda$	$\xi$	$\sigma$
30	1.9	0.00026	1.51	0.2695
60	1.9	0.000209	1.51	0.2695
90	1.9	0.000227	1.51	0.2695
120	1.9	0.0003	1.51	0.2695
150	1.9	0.000241	1.51	0.2695
180	1.9	0.000172	1.51	0.2695
210	1.9	0.000644	1.48	0.2834
240	1.9	0.0093	1.84	0.8973
270	1.9	0.036501	2.3	1.4191
300	1.9	0.041478	2.47	1.6146
330	1.9	0.018486	2.8	1.9444
Vlissingen				
0	2.97	0.006681	1.2	0.3362
30	2.97	0.001083	1.81	0.4769
60	2.97	0.000871	1.81	0.4769
90	2.97	0.000945	1.81	0.4769
120	2.97	0.001251	1.81	0.4769
150	2.97	0.001004	1.81	0.4769
180	2.97	0.000715	1.81	0.4769
210	2.97	0.002253	1.35	0.3349
240	2.97	0.012813	1.45	0.5152
270	2.97	0.03365	1.27	0.4699
300	2.97	0.044499	1.35	0.6581
330	2.97	0.02611	1.35	0.6742
IJmuiden				
0	1.86	0.006681	0.95	0.1727
30	1.86	0.000607	1.05	0.1382
60	1.86	0.000488	1.05	0.1382
90	1.86	0.000529	1.05	0.1382
120	1.86	0.000701	1.05	0.1382
150	1.86	0.000562	1.05	0.1382
180	1.86	0.0004	1.05	0.1382
210	1.86	0.001334	1.01	0.1411
240	1.86	0.011802	0.97	0.1379
270	1.86	0.059452	0.87	0.1485
300	1.86	0.062237	0.78	0.1153
330	1.86	0.035098	0.81	0.1523

#### 4.8.1.3 Correlation between wind and sea water level

High sea water levels in the North Sea are caused by combinations of high tide and high surge. In the North Sea, potential variations in the surge exceed the potential variations in the tidal levels, which means the surge is the dominant cause of high water levels. High surges in the North Sea are caused by high wind speeds in combination with northerly or westerly directions. Therefore, the mutual correlation between wind (speed and direction) and sea water level needs to be taken into account for these wind directions.

Similar to the tidal river load model, the wind is described as a function of the water level, i.e. the water level is treated as the independent variable. The correlation is described by different correlation models. The PCR-model of section 3.4.2.3 can be used, in which parameter  $\sigma$  is taken equal to the values, as shown in Table 4.23 (based on the analysis of Diermanse et al, 2003). The preferred option is to use the HES-model of section 3.4.2.2, since that model provides more reliable results for the entire range of wind speeds, whereas the PCR-model is only reliable for high (extreme) wind speeds and sea water levels. In order to make the HES model compatible with the PCR-model for extreme sea water levels and wind speeds, the associated distribution function,  $\Lambda$ , for the independent part should be normally distributed with  $\sigma$  taken equal to the values, as shown in Table 4.23. Parameter  $\delta$  should be taken equal to  $-\sigma^2/2$  (see section 3.4.2.3).

Table 4.23  $\sigma$ -values of the correlation model for wind and sea water level for the coastal regions and 6 wind directions. For the other wind directions, there is no correlation between wind speed and water level. Values of -99 indicate wind sectors for which no correlation was observed.

wind sector:	0	210	240	270	300	330
Wadden Sea east	0.94	-99	1.01	1.05	0.86	0.88
Wadden Sea west	1.56	0.58	1.18	1.21	1.05	0.97
North Sea coast; north	2.37	1.54	1.35	1.43	1.41	1.45
North Sea coast; middle	0.65	0.75	1.07	1.02	1.09	1.05
North Sea coast; south	0.65	0.75	1.07	1.02	1.09	1.05
Oosterschelde	1.23	2.52	1.5	0.65	1.29	1.41
Westerschelde	1.40	1.89	1.46	1.05	1.56	1.37

#### 4.8.1.4 Wave period

For the three North sea regions, (north, middle and south) the wave period,  $T_{m-1,0}$ , is added as an additional random variable. The statistics of this variable are described by a conditional Weibull function (see section 6.13). Parameters of 3 stations are presented in Table 4.25. Table 4.24 shows which stations are used for the three regions.

Table 4.24 Selected wave stations per region

Region	station
North Sea coast; north	ELD
North Sea coast; middle	YM6
North Sea coast; south	EUR

Table 4.25 Weibull parameters for wave period  $T_{m-1,0}$ , for three stations

sector	$\omega$	$\lambda$	$\xi$	$\sigma$
ELD				
0	7.95	0.037991	2.63	5.29
30	7.95	0.013456	2.63	5.33
60	7.95	0.003325	2.63	4.11
90	7.95	0.000168	2.63	3.28
120	7.95	0.000063	2.63	3.65
150	7.95	0.00028	2.63	3.79
180	7.95	0.000869	2.63	3.37

210	7.95	0.003926	2.63	3.72
240	7.95	0.012755	2.63	4.33
270	7.95	0.055066	2.63	4.74
300	7.95	0.091134	2.63	5.27
330	7.95	0.085726	2.63	5.30
EUR				
0	7.23	0.052908	3.59	5.46
30	7.23	0.026785	3.59	5.25
60	7.23	0.007558	3.59	5.01
90	7.23	0.00118	3.59	4.46
120	7.23	0.000115	3.59	4.20
150	7.23	0.000222	3.59	4.23
180	7.23	0.002998	3.59	4.22
210	7.23	0.013288	3.59	4.50
240	7.23	0.021147	3.59	4.65
270	7.23	0.034765	3.59	4.64
300	7.23	0.061161	3.59	5.23
330	7.23	0.079766	3.59	5.45
YM6				
0	7.77	0.03527	3.22	6.26
30	7.77	0.018052	3.22	6.05
60	7.77	0.002926	3.22	5.71
90	7.77	0.000199	3.22	4.33
120	7.77	0.000355	3.22	4.83
150	7.77	0.000548	3.22	4.70
180	7.77	0.001769	3.22	4.22
210	7.77	0.009079	3.22	4.60
240	7.77	0.019763	3.22	5.24
270	7.77	0.05363	3.22	5.41
300	7.77	0.076902	3.22	6.14
330	7.77	0.077629	3.22	6.33

## 4.8.2 Hydrodynamic models

### 4.8.2.1 Spatial interpolation of water levels

Section 4.8.1.2 described the statistical distribution functions of twelve water level stations along the Dutch coast. In order to determine water levels statistics at each coastal dike, spatial interpolation is applied on the statistics of the water level stations. Each coastal region is divided into one or more subregions. Within each subregion, triangular linear interpolation is applied between three selected water level stations. This means that the 2-dimensional (hyper-)plane is derived that connects the water levels in the three stations. This plane then automatically describes water levels in other locations.

The local water level in location L is derived according to the following procedure:

- 1 Starting point of the procedure is the standard normally distributed variable,  $u_h$ , that represents the water level,  $h$ . The value of  $u_h$  is generated by the probabilistic method of choice, e.g. FORM, Monte carlo or numerical integration (see section 2.3).

- 2 Translate  $u_h$  to water level  $h$  for the three selected stations, based on equal probability of (non-)exceedance. The result is a realization of the water level in each station:  $h_1$ ,  $h_2$  and  $h_3$ .
- 3 Apply triangular interpolation between the water levels in the three stations to determine the water level at location  $L$ .

So,  $h_1$ ,  $h_2$  and  $h_3$  are the water levels in the three stations *with the same return period*. The return period corresponds to the value of  $u_h$ .

In step 3 above, a mathematical description is derived for the plane that connects the water levels in the 3 stations. The general mathematical expression for the plane is:

$$h = c_1x + c_2y + c_3 \quad (3.12)$$

In which:

$h$  = water level  
 $x, y$  = coördinates  
 $c_1, c_2, c_3$  = coëfficiënts

The coëfficiënts  $c_1$ ,  $c_2$ ,  $c_3$  are derived from the coördinates and water levels of the stations:

$$h_i = c_1x_i + c_2y_i + c_3 \quad ; i = 1..3 \quad (3.13)$$

In which index  $i$  refers to the number of the station. Coëfficiënts  $c_1$ ,  $c_2$ ,  $c_3$  are derived as follows:

$$\begin{aligned} c_1 &= \frac{b_1}{a} \quad ; c_2 = \frac{b_2}{a} \quad ; c_3 = \frac{b_3}{a} \\ a &= x_1(y_2 - y_3) - x_2(y_1 - y_3) + x_3(y_1 - y_2) \\ b_1 &= h_1(y_2 - y_3) - h_2(y_1 - y_3) + h_3(y_1 - y_2) \\ b_2 &= x_1(h_2 - h_3) - x_2(h_1 - h_3) + x_3(h_1 - h_2) \\ b_3 &= x_1(y_2h_3 - y_3h_2) - x_2(y_1h_3 - y_3h_1) + x_3(y_1h_2 - y_2h_1) \end{aligned} \quad (3.14)$$

This formulation provides the he coëfficiënts  $c_1$ ,  $c_2$ ,  $c_3$  that describe the plane according to equation (3.13). Subsequently, the water level at any location  $L$  can be determined by substitution of the  $x, y$  coordinates of location  $L$  in equation (3.13).

For each location  $L$ , three “nearby” stations are required to determine the water level with the procedure above. For some regions, only two “real stations” are available in the procedure. In that case a fictional third station is defined, which has the exact same statistical features as one of the other two stations, but is situated at a different location. Such a fictional station is used to force the orientation of the water level gradient in a certain direction (see the example of the Westerschelde below).

The remainder of this section shows the selection of stations for each region. Note that this is a pre-processing procedure for Hydra-Ring. For a predefined set of coastal locations (i.e. the output locations of the wave model SWAN) the procedure is applied to determine which three stations are used for the interpolation procedure. The ID of the three stations is stored in a database that is input for Hydra-Ring.

### *Westerschelde*

For the Westerschelde estuary, two water level stations are available: Vlissingen and Hansweert. The gradient of the water level in this estuary during periods of high storm surge is from West/northwest to east/southeast (see Figure 4.3) with higher water levels in the east/southeastern part of the estuary. To ensure that this orientation is replicated in the load model, a third, fictitious, station is created. This fictitious station has the same statistics as Vlissingen, but is placed 30 degrees north-northeast of Vlissingen. This means the water level along the line in north-northeastern direction is constant and therefore the gradient of the water level plane has a southeast-easterly direction. The distance of the fictitious station is chosen arbitrarily 10 km from Vlissingen, but this choice has no effect on the results.

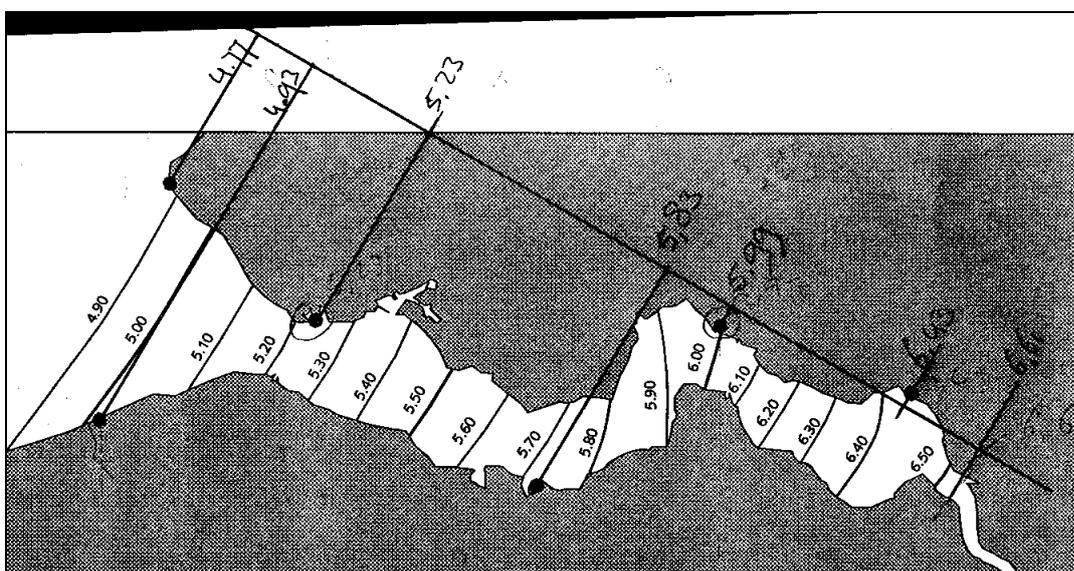


Figure 4.3 Water level gradient in the Westerschelde.

Table 4.26 Water level stations for which statistics are used in the procedure of the Westerschelde.

number	station	x-coordinate	y-coordinate
1	Vlissingen	30480	385220
2	Hansweert	59050	384960
3	fictitious_Vlissingen	35480	393880

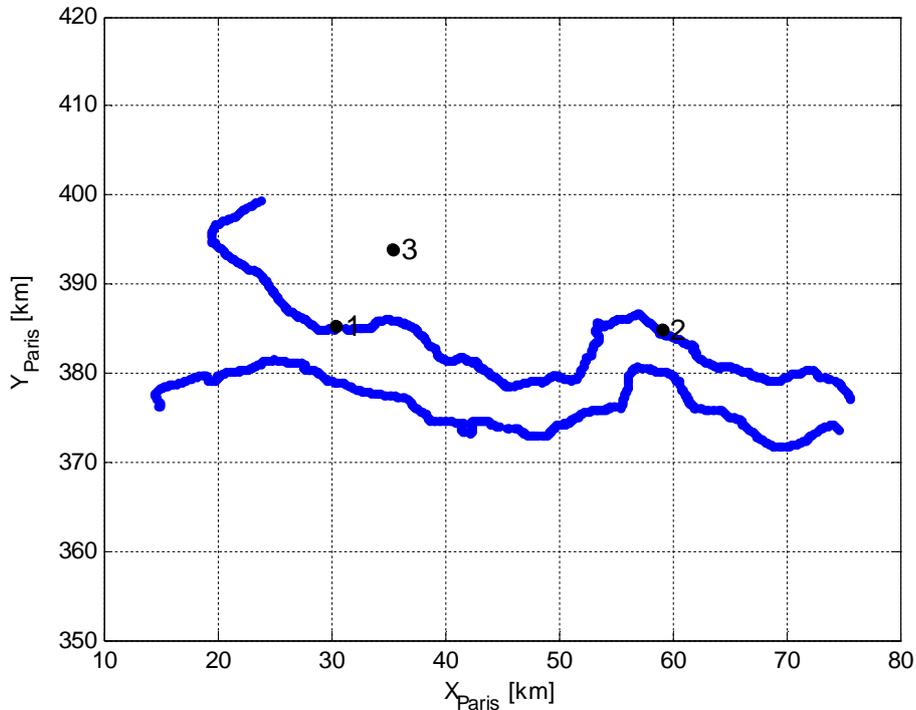


Figure 4.4 The westerschelde and the three stations of Table 4.26

#### North Sea regions

Table 4.27 shows the water level stations that are used for the three North Sea regions (North, Middle and South). Similar to the procedure of the Westerschelde, a fictitious station is introduced to force the orientation of the water level gradient in a desired direction. This station has the same statistical features as station IJmuiden.

The dike locations of the North Sea regions are divided in four subregions. For each subregion, a different set of stations is used for the triangular interpolation method (see Table 4.28). Directly below Table 4.28, the recipe is given for identifying in which subregion a certain location L is situated. The subdivision in subregions is visualized in Figure 4.6.

Table 4.27 Water level stations for which statistics are used in the procedure of the North Sea regions

number	station	x-coordinate	y-coordinate
1	Vlissingen	30480	385220
2	OS 11	23013	407778
3	Hoek van Holland	67930	445000
4	IJmuiden	98430	497500
5	Fictitious_IJmuiden	108133	495081
6	Den Helder	111850	553230

number	station	x-coordinate	y-coordinate
7	Terschelling West	143870	597420
8	Den Oever	132030	549440
9	Harlingen	156480	576550

Table 4.28 Selected water level stations for subregions of the North Sea regions

subregion	Station 1	Station 2	Station 3
I	Den Oever	Harlingen	Terschelling West
II	IJmuiden	Den Helder	fictitious_IJmuiden
III	Hoek van Holland	IJmuiden	Den Helder
IV	Vlissingen	OS11	Hoek Van Holland

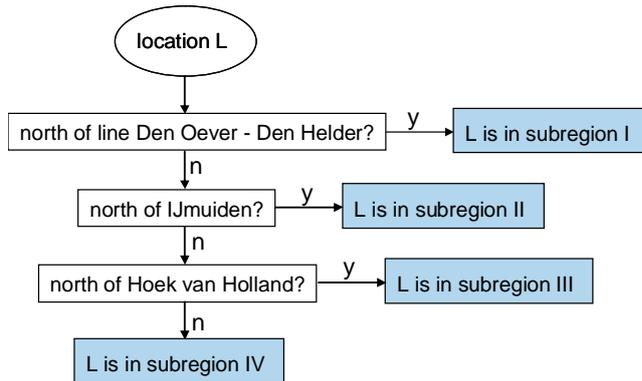


Figure 4.5 Flow diagram for dividing locations along the North Sea coast in different subregions

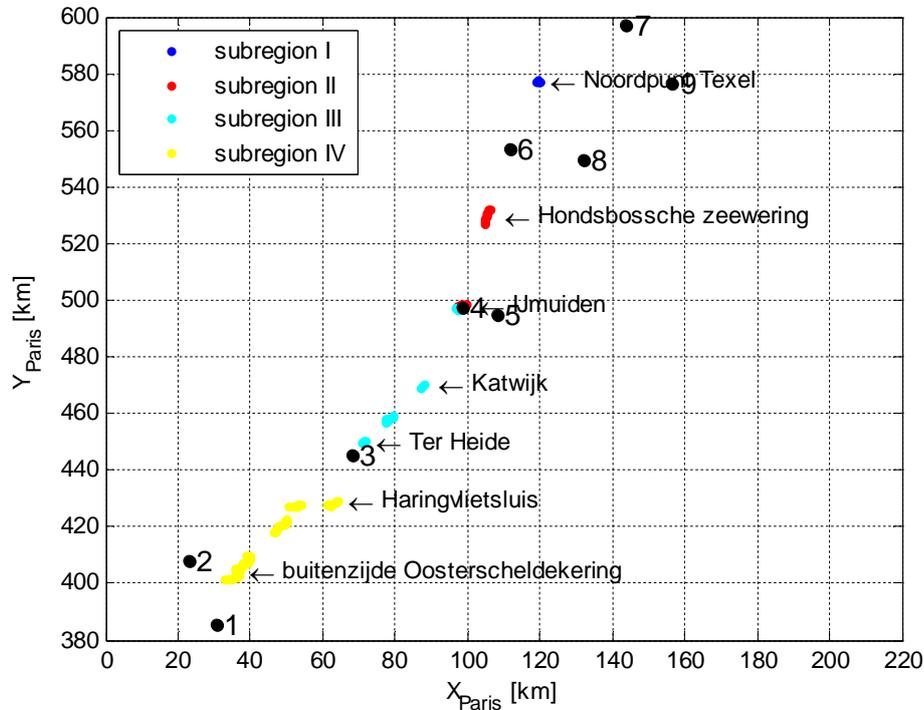


Figure 4.6 Four subregions for triangular interpolation of water levels in the North Sea regions. Station numbers correspond to the numbers shown in Table 4.27.

#### Wadden Sea regions

Table 4.29 shows water level stations that are used for the two Wadden Sea regions (west and east). Similar to the procedure of the Westerschelde and North Sea, two fictitious stations are introduced to force the orientation of the water level gradient in a desired direction. These stations have the same statistical features as station Den Helder.

The dike locations of the Wadden Sea regions are divided in six subregions. For each subregion, a different set of stations is used for the triangular interpolation method (see Table 4.30). Directly below Table 4.28, the recipe is given for identifying in which subregion a certain location L is situated. The subdivision in subregions is visualized in Figure 4.8.

Table 4.29 Water level stations for which statistics are used in the procedure of the Wadden Sea

number	station	x-coordinate	y-coordinate
1	Den Helder	111850	553230
2	Den Oever	132030	549440
3	Harlingen	156480	576550
4	Lauwersoog	208850	602790
5	Delfzijl	258000	594430
6	Huibertgat	221990	621330

number	station	x-coordinate	y-coordinate
7	West Terschelling	143870	597420

Table 4.30 Selected water level stations for subregions of the Wadden Sea

Subregion	Station 1	Station 2	Station 3
I	Lauwersoog	Delfzijl	Huibertgat
II	Huibertgat	Lauwersoog	Terschelling West
III	Harlingen	Lauwersoog	Terschelling West
IV	Harlingen	Den Oever	Terschelling West
V	Den Helder	Den Helder	Den Helder
VI	Harlingen	Den Oever	Terschelling West

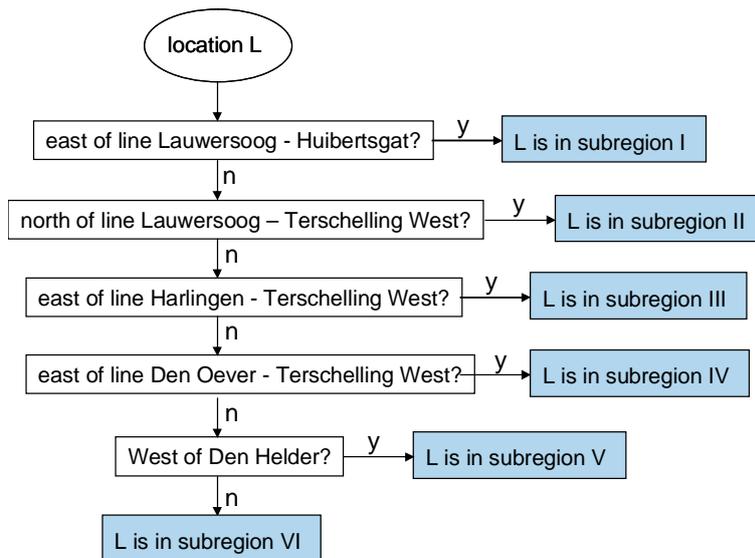


Figure 4.7 Flow diagram for dividing locations along the Wadden Sea in different subregions

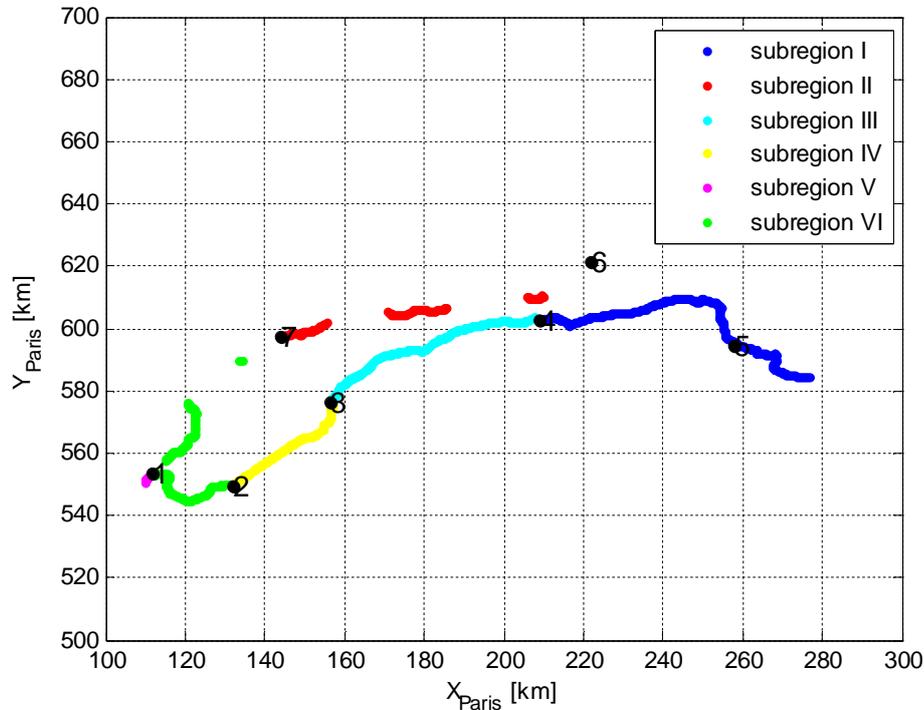


Figure 4.8 Six subregions for triangular interpolation of water levels in the Wadden Sea. Station numbers correspond to the numbers shown in Table 4.29.

#### 4.8.2.2 Water level corrections

With the triangular interpolation method of the previous section, water level  $h$  can be derived for each location and for each frequency of exceedance,  $f$ . Suppose that for a given location the normative frequency is equal to  $f^*$ . The water level,  $h(f^*)$ , can then be considered as the normative water level. However, the actual normative water level,  $H$ , is derived in a slightly different way (see RIKZ, 2006) and can therefore be different from  $h(f^*)$ . The differences, although small, are undesired for practical purposes. Therefore, it was decided to correct the sea water levels as computed in the load model of Hydra-Ring.

$$\text{Correction} = H - h(f^*)$$

The correction is added each time after the water level at the considered location  $L$  has been determined from triangular interpolation. This correction guarantees that the computed water level with exceedance frequency  $f^*$  is equal to  $H$ . Note that the correction is also done for frequencies other than  $f^*$ , to ensure that function  $h(f)$  remain continuous.

The derivation of the corrections is preprocessing for Hydra-Ring and the results are stored in input databases. Due to the complexity of the Oosterschelde model (see section 4.9) no such corrections have been defined for this region.

#### 4.8.2.3 Wave parameters

The wave simulation model SWAN was used to derive wave characteristics (height, period and direction) near all dike sections for selected combinations of wind direction, wind speed

and water level. For other combinations, 3-dimensional interpolation is used to determine the associated derive wave characteristics at the dike sections. For the North Sea regions, SWAN computations were executed for selected combinations of wind direction, wind speed, water level *and wave period*. Consequently, for these regions *4-dimensional* interpolation is used to determine the associated derive wave characteristics at the dike sections.

## 4.9 Load model for the Oosterschelde

Most components of the Oosterschelde load model were already described in section 4.8. However, the Oosterschelde differs from the other coastal areas due to the presense of a storm surge barrier. The barrier closes when high water levels at sea are predicted, to prevent high water levels in the estuary. The effect of the closure of the barrier on water levels in the Oosterschelde complicates the description of hydraulic loads and therefore a separate load model for this area is required.

Similar to the load models of the lake delta and the Rhine-Meuse delta, it is taken into account that the barrier can fail to close due to e.g. mechanical failures or human errors. Furthermore, it is also taken into account that predicted water levels, required to determine if the barrier needs to be closed, contain uncertainties. The manner in which the closure of the barrier influences water levels in the Oosterschelde is further influenced by additional variables, such as the phase difference between the storm surge peak and the peak of the tide and the storm surge duration.

### 4.9.1 Distribution functions of random variables

In the load model of the dikes along the Oosterschelde, the following random variables are relevant:

- Wind (speed and direction)
- Sea water level
- functioning of the barrier
- water level predictions for barrier
- phase tidal peak-surge peak
- storm surge duration

Each variable is considered a “slowly evolving variable” (see section 2.5.2). This means the (peak) values of these variables in subsequent tidal periods are assumed to be independent.

#### 4.9.1.1 *Wind and sea water level*

These variables and their distribution functions have been described in section 4.8.

#### 4.9.1.2 *Functioning of the barrier*

Similar to the the storm surge barriers at Hoek van Holland and Ramspol, the possibility that the Oosterschelde barrier fails to close according to procedure is taken into account. The main difference with the other two barriers is that for the Oosterschelde there is a backup-procedure that is assumed never to fail. However, water levels in the Oosterschelde will be higher if the backup procedure is carried out, compared to the case in which the normal procedure is carried out. Water levels for both situations are computed with the hydrodynamic model IMPLIC (see section 4.9.2). The probability that the Oosterschelde barrier fails to close according to procedure is estimated to be equal to 0.01.

#### 4.9.1.3 Water level predictions for the barrier

The barrier is closed if the *predicted* water level is higher than 3.00 m+NAP. Water level predictions contain uncertainties that need to be taken into account by the probabilistic model. This uncertainty is modelled by a normal distribution function with a mean of 0 m and a standard deviation of 0.25 m.

#### 4.9.1.4 Phase tidal peak-surge peak

The phase difference between the tidal peak and the surge peak is assumed to be uniformly distributed between -6.208 hrs and +6.208 hrs (6.208 hrs = halve a tidal period).

#### 4.9.1.5 Storm surge duration

The storm surge duration is assumed to be lognormally distributed with a mean of 54.3 hours and a standard deviation of 18.8 hours.

### 4.9.2 Hydrodynamic models

#### 4.9.2.1 Water levels

Model simulations were executed with the Hydrodynamic model IMPLIC to compute water levels for a set of combinations of:

- Wind speed;
- Wind direction;
- Sea water level (seaward of the flood barrier);
- functioning of the barrier;
- phase tidal peak-surge peak; and
- storm surge duration.

The output of the model runs has been stored in the input databases of Hydra-Ring for the 9 locations as described in Table 4.31. For other combinations of the six random variables, 6-dimensional interpolation is used to determine the associated water levels in the 9 locations. Note that the seventh random variable, the error in the predicted water level, is not included in the bullet list above. Realisations of this variable are simply added to the realizations of the water level. The resulting water level, "corrected" for prediction uncertainties, is used as input for the 6-dimensional interpolation procedure.

Table 4.31 Water level stations for which statistics are used in the procedure of the Oosterschelde

number	name	x-coordinate	y-coordinate
1	ROOMPOT BUITEN	35133	404133
2	ROOMPOT BINNEN	39067	403000
3	BURGHSLUIS	44067	411533
4	WEMELDINGE	61667	393933
5	MAROLLE GAT	68333	386333
6	STAVENISSE	59533	402333
7	PHILIPSDAM WEST	69533	409933
8	COLIJNSPLAAT	50333	402867
9	ZEELANDBRUG NOORD	51000	405667

For other locations than the 9 mentioned in Table 4.31, the water level is determined by triangular interpolation. For this purpose, the dike locations of the Oosterschelde are divided in six subregions. For each subregion, a different set of stations is used for the triangular interpolation method (see Table 4.32). Directly below Table 4.32, the recipe is given for

identifying in which subregion a certain location L is situated. The subdivision in subregions is visualized in Figure 4.10.

Table 4.32 Selected water level stations for subregions of the Oosterschelde

Subregion	Station 1	Station 2	Station 3
I	Roopot binnen	Burghsluis	Zeelandbrug noord
II	Roopot binnen	Colijnsplaat	Zeelandbrug noord
III	Wemeldinge	Colijnsplaat	Zeelandbrug noord
IV	Wemeldinge	Marollegat	Stavenisse
V	Stavenisse	Philipsdam West	Zeelandbrug noord
VI	Wemeldinge	Stavenisse	Zeelandbrug noord

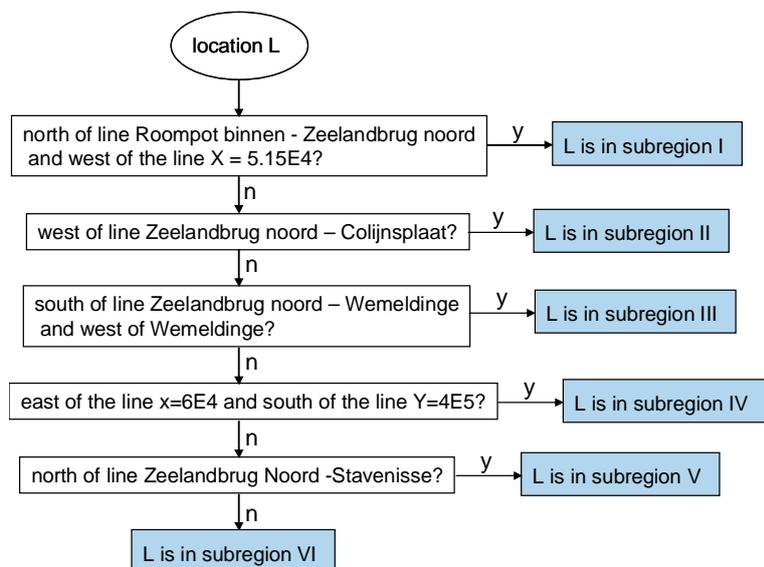


Figure 4.9 Flow diagram for dividing locations along the Oosterschelde in different subregions

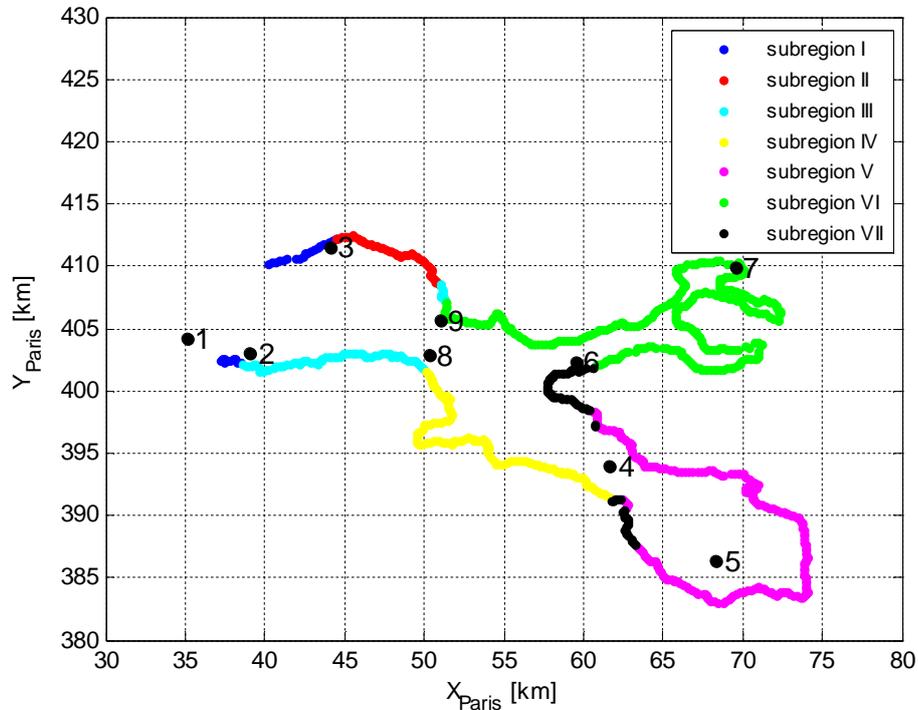


Figure 4.10 Six subregions for triangular interpolation of water levels in the Oosterschelde. Station numbers correspond to the numbers shown in Table 4.31.

#### 4.9.2.2 Waves

The wave simulation model SWAN was used to derive wave characteristics (height, period and direction) near all dike sections for selected combinations of wind direction, wind speed and water level. Note that the water level in this case refers to the water level on the “inward” side of the Oosterschelde barrier, as computed with the IMPLIC model and subsequent 6-dimensional interpolation procedure. For other combinations than the ones computed with IMPLIC, 3-dimensional interpolation is used to determine the associated wave characteristics at the dike sections.

### 4.10 Load model for the coastal dunes

#### 4.10.1 Distribution functions of random variables

In the load model of the coastal dikes, the following random variables are relevant:

- Sea water level
- Waves (height and period)

The load model for dunes is different from the other load models because it only considers omnidirectional year statistics. In other words: statistical distribution functions are not derived/applied separately per wind direction sector.

Each variable is considered a “slowly evolving variable” (see section 2.5.2). This means the (peak) values of these variables in subsequent tidal periods are assumed to be independent.

### 4.10.1.1 Sea water level

The statistics of this variable are described by a conditional Weibull function (see section 6.13). Parameters of six stations are presented in Table 4.33. Section 0 describes how statistics of these six stations are translated to dune locations.

Table 4.33 Weibull parameters for sea water level for five stations

station	$\omega$	$\lambda$	$\xi$	$\sigma$
Vlissingen	2.97	3.907	1.04	0.2796
Hoek van Holland	1.95	7.237	0.57	0.0158
IJmuiden	1.85	5.341	0.63	0.0358
Den Helder	1.60	3.254	1.60	0.9001
Eierland	2.25	0.500	1.86	1.0995
Borkum	1.85	5.781	1.27	0.5350

### 4.10.1.2 Wave height

The wave height,  $H_s$ , has a probability distribution function that is conditional to the water level. The expected value of  $H_s$  is formulated as follows:

$$\begin{aligned} \mu(H_s | h) &= a + bh - c(d - h)^e \quad [m] \quad ; \text{if } h \leq d \\ &= a + bh \quad [m] \quad ; \text{if } h > d \end{aligned} \quad (4.4)$$

where:

- $H_s$  = wave height [m]
- $h$  = water level [m+NAP]
- $\mu$  = expected value (in statistical sense)
- a-e = parameters

a, b, c, d and e are parameters that differ from location to location (see Table 4.34).

Table 4.34 Parameters of equation (4.4) for five locations

	Vlissingen	Hoek van Holland	IJmuiden	Den Helder	Eierland	Borkum
<b>a</b>	2.40	4.35	5.88	9.43	12.19	10.13
<b>b</b>	0.35	0.6	0.6	0.6	0.6	0.6
<b>c</b>	0.0008	0.0008	0.0254	0.68	1.23	0.57
<b>d</b>	7	7	7	7	7	7
<b>e</b>	4.67	4.67	2.77	1.26	1.14	1.58

Furthermore, it is assumed that the wave height is standard normally distributed with a mean according to eq. (4.4) and a standard deviation of  $\sigma(H_s) = 0,6$  m. This standard deviation of 0,6 m is assumed to be valid for all locations.

Figure 4.11 shows the relation between water level and expected wave height for the six locations.

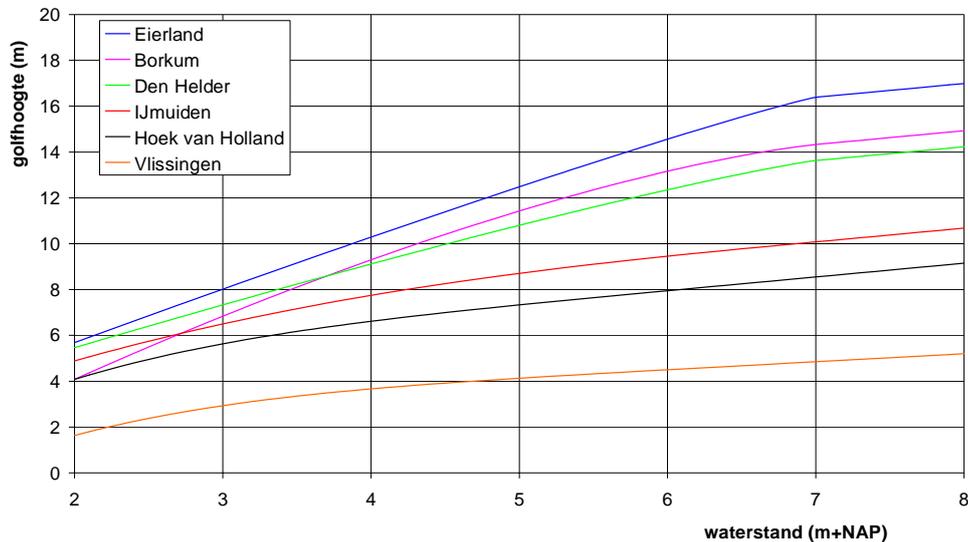


Figure 4.11 Relation between water level and (expected) wave height at 5 locations

In Hydra-Ring, the correlation model as described above is programmed in a preprocessing procedure for practical reasons. As with other correlation models (Figure 3.8), the starting point of the preprocessing procedure is a set of realizations  $u_1$  and  $u_2$  of two mutually independent standard normally distributed variables  $U_1$  and  $U_2$ . Variables  $U_1$  and  $U_2$  represent two “real world” variables  $X_1$  and  $X_2$ . In the example above, variable  $X_1$  is the water level and  $X_2$  is the wave height. The first step in the correlation model is to translate  $u_1$  to  $x_1$ , using the inverse distribution function of variable  $X_1$ :

$$x_1 = F_{X_1}^{-1}(\Phi(u_1)) \quad (3.15)$$

The following step is to write the distribution function of variable  $X_2$  in terms of  $X_1$ .

$$F_{X_2,cor}(x_2) = N(G(x_1), \sigma) \quad (3.16)$$

where  $N$  represents the normal distribution, and where  $G(x_1)$  is the mean value of the normal distribution, and is a function of  $x_1$ . Function  $G$  can be virtually any type of function, although for reliability analysis it is preferable that  $G$  is a continuous function. The value of  $x_2$  is computed from the inverse distribution function:

$$x_2 = F_{X_2,cor}^{-1}(\Phi(u_2)) \quad (3.17)$$

This model is used to correlate:

1. the wave height to the water level; and
2. the wave period to the wave height.

Exacte invulling met tabellen etc nog uitwerken

### 4.10.1.3 Peak wave period

The peak wave period,  $T_p$ , is assumed to have a Gaussian distribution. The mean,  $\mu$ , is derived from the wave height,  $H_s$ , by a deterministic relation that is described by the tables below. The standard deviation,  $\sigma$ , is equal to 1 s. The tables below show the relation between wave height ( $H_s$ ) and wave period ( $T_{m-1,0}$  and  $T_p$ ) for five different stations:

- 1 Borkum
- 2 Eierland
- 3 IJmuiden
- 4 Hoek van Holland
- 5 Vlissingen

For location, Den Helder, no such relation is directly available. The peak wave period for this location is based on the peak wave period of IJmuiden and Eierland, using the following relation:

$$T_{p;Den\ Helder}(H_s) = \lambda T_{p;IJmuiden}(H_s) + (1-\lambda)T_{p;Eierland}(H_s)$$

The value of  $\lambda$  is equal to 0.35, based on the distances from Den Helder to IJmuiden and Den Helder to Eierland respectively.

Table 4.35 Relations between wave height,  $H_s$ , on one hand and wave periods,  $T_{m-1,0}$  and  $T_p$ , on the other hand for location Borkum.

Hs	$T_{m-1,0}$	$T_p$
2.80	7.41	8.59
3.82	8.58	9.97
4.69	9.59	11.18
5.46	10.47	12.28
6.18	11.28	13.29
6.85	12.03	14.23
7.49	12.73	15.11
8.09	13.38	15.96
8.68	14.01	16.76
9.24	14.60	17.53
9.79	15.17	18.27
10.32	15.71	18.99
10.84	16.24	19.69
11.35	16.75	20.36
11.85	17.24	21.01
12.34	17.72	21.65
12.81	18.19	22.28
13.28	18.64	22.88
13.75	19.08	23.48
14.20	19.52	24.06
14.65	19.94	24.63
15.09	20.35	25.19
15.53	20.76	25.74
15.96	21.15	26.28
16.38	21.54	26.82

<b>Hs</b>	<b>T<sub>m-1,0</sub></b>	<b>T<sub>p</sub></b>
16.80	21.93	27.34
17.22	22.30	27.85

Table 4.36 Relations between wave height,  $H_s$ , on one hand and wave periods,  $T_{m-1,0}$  and  $T_p$ , on the other hand for location ELD.

<b>Hs</b>	<b>T<sub>m-1,0</sub></b>	<b>T<sub>p</sub></b>
3.08	7.22	8.22
3.72	7.79	8.87
4.24	8.28	9.43
4.86	8.85	10.10
5.46	9.41	10.76
6.04	9.96	11.42
6.62	10.50	12.07
7.19	11.02	12.72
7.75	11.54	13.35
8.31	12.04	13.98
8.85	12.53	14.59
9.39	13.02	15.20
9.93	13.49	15.80
10.46	13.95	16.39
10.99	14.41	16.97
11.51	14.86	17.55
12.03	15.30	18.12
12.55	15.73	18.68
13.06	16.16	19.23
13.57	16.58	19.78
14.07	16.99	20.32
14.57	17.40	20.85
15.07	17.80	21.38
15.57	18.19	21.90
16.06	18.59	22.42
16.55	18.97	22.93
17.04	19.35	23.44
17.52	19.73	23.94
18.01	20.10	24.44
18.49	20.47	24.93

Table 4.37 Relations between wave height,  $H_s$ , on one hand and wave periods,  $T_{m-1,0}$  and  $T_p$ , on the other hand for location Ijmuiden (IJM).

<b>Hs</b>	<b>T<sub>m-1,0</sub></b>	<b>T<sub>p</sub></b>
3.20	7.11	8.02
4.25	8.33	9.37
5.00	9.21	10.38
5.61	9.91	11.20
6.13	10.50	11.89
6.57	11.00	12.50

Hs	$T_{m-1,0}$	$T_p$
6.98	11.44	13.03
7.34	11.84	13.52
7.67	12.20	13.96
7.98	12.53	14.37
8.26	12.84	14.74
8.53	13.12	15.09
8.78	13.39	15.43
9.02	13.63	15.74
9.25	13.87	16.03
9.47	14.09	16.31
9.67	14.30	16.58
9.87	14.51	16.83
10.06	14.70	17.07
10.24	14.88	17.31
10.42	15.06	17.53
10.59	15.23	17.75
10.75	15.39	17.96
10.91	15.55	18.16
11.07	15.70	18.35
11.22	15.85	18.54
11.36	16.00	18.73

Table 4.38 Relations between wave height,  $H_s$ , on one hand and wave periods,  $T_{m-1,0}$  and  $T_p$ , on the other hand for location Hoek van Holland (EUR).

Hs	$T_{m-1,0}$	$T_p$
3.34	7.02	7.82
4.23	7.76	8.61
4.86	8.32	9.24
5.37	8.77	9.76
5.79	9.16	10.20
6.16	9.50	10.60
6.48	9.80	10.95
6.78	10.07	11.27
7.04	10.32	11.56
7.29	10.54	11.83
7.52	10.75	12.08
7.73	10.95	12.31
7.93	11.13	12.53
8.11	11.30	12.74
8.29	11.46	12.93
8.46	11.61	13.12
8.62	11.75	13.29
8.77	11.89	13.46
8.92	12.02	13.62
9.06	12.15	13.78

Hs	T <sub>m-1,0</sub>	T <sub>p</sub>
9.19	12.27	13.92
9.32	12.38	14.07
9.45	12.50	14.20
9.57	12.60	14.34
9.68	12.71	14.46
9.80	12.81	14.59

Table 4.39 Relations between wave height,  $H_s$ , on one hand and wave periods,  $T_{m-1,0}$  and  $T_p$ , on the other hand for location Vlissingen (SCW).

Hs	T <sub>m-1,0</sub>	T <sub>p</sub>
1.55	5.79	6.57
2.72	6.86	7.77
3.31	7.56	8.58
3.73	8.09	9.22
4.07	8.53	9.76
4.36	8.90	10.22
4.62	9.22	10.62
4.85	9.52	10.99
5.06	9.78	11.32
5.25	10.03	11.63
5.43	10.25	11.92
5.60	10.46	12.18
5.76	10.66	12.44
5.91	10.85	12.67
6.05	11.02	12.90
6.19	11.19	13.12
6.32	11.35	13.32
6.44	11.50	13.52
6.56	11.65	13.71
6.68	11.79	13.89
6.79	11.93	14.07
6.90	12.06	14.24
7.01	12.18	14.40

#### 4.10.2 Hydrodynamic model

##### 4.10.2.1 Interpolation procedure

The load model for the dunes is different from the other load models for multiple reasons. One of them is that no hydrodynamic model is used to translate the available statistics of water levels and waves to the nearshore locations at the flood defense. Instead, the water level and wave statistics are used directly as input for a dune erosion model (to be described in a later phase in chapter 6). First, however, spatial interpolation is applied to make the statistics of the five stations available for each dune transect.

Based on the available statistics of the 6 locations the hydraulic loads for all locations along the Dutch coast north of Hoek van Holland can be derived. Actually, first an additional, 7<sup>th</sup>,

location is defined: “Steunpunt Waddenzee” (see Figure 4.12). This is done for interpolation purposes, i.e. to prevent the line between the two northernmost locations, Eierland and Borkum, to cross the Wadden Islands. The hydraulic loads of the “steunpunt” are derived through application of the following relation:

$$X_{\text{steunpunt}} = 0.57 \times X_{\text{Eierland}} + 0.43 \times X_{\text{Borkum}}$$

Where  $X$  stands for water level, wave height and peak wave period. Table 4.40 and Figure 4.12 show the seven locations for which the hydraulic boundary conditions are available.

Table 4.40 Locations for which boundary conditions are available and their Parisian coordinates.

nr	locatie	X	Y
1	Vlissingen	-7.797	380.645
2	Hoek van Holland	58.748	450.830
3	IJmuiden	79.249	501.800
4	Den Helder	98.372	549.340
5	Eierlandse Gat	106.514	587.985
6	Steunpunt Waddenzee	150.000	621.230
7	Borkum	221.990	621.330

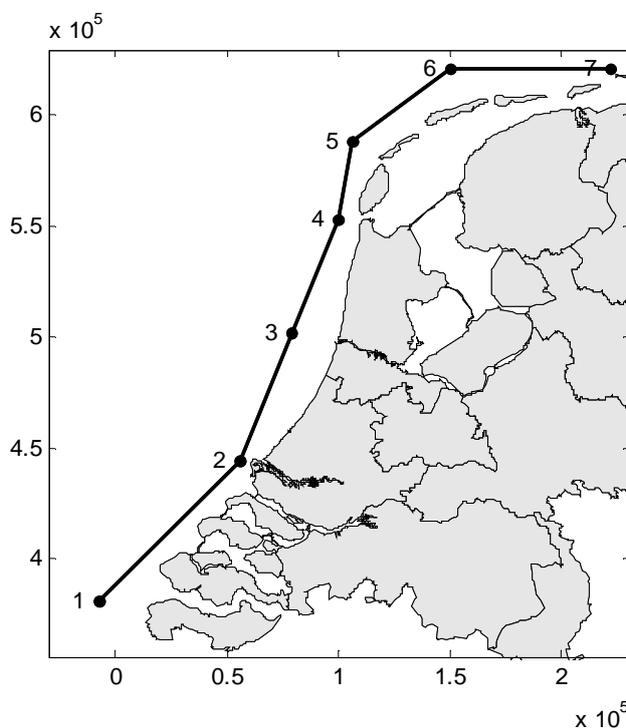


Figure 4.12 Locations for which boundary conditions are available. Numbers correspond with Table 4.40.

The seven locations (steunpunten) are connected by six straight lines in Figure 4.12 and, again, in Figure 4.13. For each dune profile in the Netherlands, a line is drawn, perpendicular

to the coast line (the coloured lines in Figure 4.13). The location,  $L$ , where such a perpendicular line intersects with one of the five straight lines is determinative for the interpolation procedure. The distance between  $L$  and the 2 closest “steunpunten” determines the value of the multiplication factors that are used for the interpolation procedure.

**Example:** Suppose the line, perpendicular to the coast line at a location  $D$ , intersects at location  $L$  with the line between Hoek van Holland (location 2 in Figure 4.13) and IJmuiden (location 3 in Figure 4.13). Furthermore, suppose the distance from  $L$  to Hoek van Holland is 20 km and the distance from  $L$  to IJmuiden is 30 km. This means the relative distances are 0.4 (Hoek van Holland) and 0.6 (IJmuiden). The water level and wave characteristics at location  $D$  are then derived as follows:

$$X_D = 0.6 \times X_{\text{Hoek van Holland}} + 0.4 \times X_{\text{IJmuiden}}$$

Where  $X$  stands for water level, wave height and peak wave period.

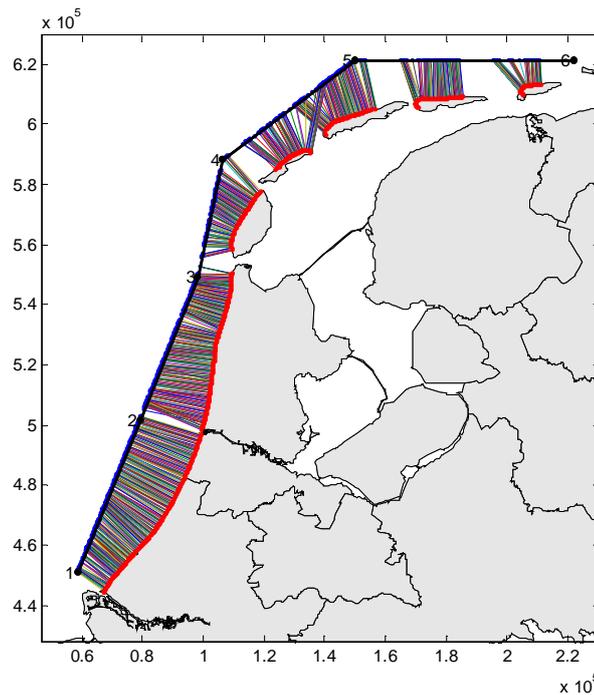


Figure 4.13 Schematic view of the interpolation procedure.

#### 4.10.2.2 Water level corrections

With the interpolatiin method of the previous section, water level  $h$  can be derived for each location and for each frequency of exceedance,  $f$ . Suppose that for a given location the normative frequency is equal to  $f^*$ . The water level,  $h(f^*)$ , can then be considered as the normative water level. However, the actual normative water level,  $H$ , is derived in a slightly different way and can therefore be different from  $h(f^*)$ . The differences, although small, are undesired for practical purposes. Therefore, it was decided to correct the sea water levels as computed in the load model of Hydra-Ring. These corrections are applied in a similar manner

as described in section 4.8.2.2. The derivation of the corrections is preprocessing for Hydra-Ring and the results are stored in input databases.

## 4.11 Spatial correlations between load models of different regions

Some dike ring areas in The Netherlands are bordering different water systems. In order to compute the probability of failure for these dike rings, multiple load models need to be considered. Furthermore, correlations between random variables of different regions need to be taken into account. For instance, the sea is experiencing high wind speeds, the tidal river area is likely to experience high wind speeds as well. Table 4.41 shows the correlation coefficients in Hydra-Ring that are applied to model the statistical dependence between variables of different regions in The Netherlands.

The computation of the combined failure probabilities of two dike sections along different water systems are done in similar style as the computation for two dike sections along the same water system, i.e. using the Hohenbichler method (see section 2.4.2). If a load variable X is only used in water system 1 and not in water system 2, this simply means the  $\alpha$ -value of variable X in system 2 is equal to 0 by definition.

Table 4.41 Correlations between random variables of different regions

variable	Wind speed @ Schiphol/Deelen	Lake level – Marker Lake	Water level @ Den Helder	Water level @ Harlingen	Water level @ Lauwersoog	Water level @ OS11	Water level @ IJmuiden	Wind speed @ Terschelling West	Wind speed @ IJmuiden	Wind speed @ Hoek van Holland	Water level – dunes (zachte kust)
Water level @ Hoek van Holland	0.80									0.40	0.97
Wind speed @ Schiphol/Deelen  MM										0.40	
Wind speed @ Schiphol/Deelen			0.90	0.97				0.40			
Lake level – IJssel Lake		0.97									
Water level @ Den Helder				0.40			0.40				0.40
Water level @ Vlissingen						0.97					0.97
Water level @ Harlingen					0.97						0.97
Water level @ Lauwersoog											0.97
Water level @ OS11											0.97
Wind speed @ de Kooy								0.40			
Wind speed @ Terschelling West									0.40		

Suppose there are two random (load) variables, X and Y, from two different regions, with a mutual correlation coefficient  $\rho$ . The value of  $\rho$  is obtained through the input correlation matrix. The u-value of Y can be written as a function of the u-value of X:

$$u_y = \rho u_x + \sqrt{1 - \rho^2} u_y^*$$

In which  $u_y^*$  is independent of  $u_x$ . The output of Hydra-Ring (i.e. the input of Combin) will contain several  $\alpha$ -values for variable  $y$ , describing the contribution of variable  $Y$  in the linearised  $Z$ -functions for various combinations of sections, submechanisms, layers etc. Consider one such  $Z$ -function in which  $Y$  is involved:

$$Z = \alpha_y u_y + \alpha_2 u_2 + \dots + \alpha_n u_n$$

$u_2 \dots u_n$  represent the other variables that are relevant for the  $Z$ -function. Combining the two equations above provides:

$$\begin{aligned} Z &= \alpha_y \left( \rho u_x + \sqrt{1 - \rho^2} u_y^* \right) + \alpha_2 u_2 + \dots + \alpha_n u_n \\ &= \alpha_y \rho u_x + \alpha_y \sqrt{1 - \rho^2} u_y^* + \alpha_2 u_2 + \dots + \alpha_n u_n \end{aligned}$$

So, the linearised  $Z$ -function is now a function of variables  $u_x$  and  $u_y^*$  instead of variable  $u_y$ . In terms of Hydra-Ring output, this means  $\alpha_y$  is replaced by two  $\alpha$ -values,  $\alpha_{y1}$  and  $\alpha_{y2}$ :

$$\alpha_{y1} = \alpha_y \rho$$

$$\alpha_{y2} = \alpha_y \sqrt{1 - \rho^2}$$

## 4.12 Combination

### 4.12.1 Stappen

- transposeToRingStochasts
- combineMainMechanismSections
  - combineAlternatives
  - combineLayers
  - combineMechanisms
- combinePresentationSections
  - combineSections
  - combineMainMechanisms
- combineAreaMainMechanism
- combineArea

### 4.12.2 Transpose to Ring stochasts

In the HydraRing calculation the stochastic variables are calculated for a specific alternative, layer, section and mechanism. For this combination the section is belonging to only one region. The dike ring may belong to more than one region. In that case the stochastic variables should be transferred to one basis. This is done in the subroutine transposeToRingStochasts.

Hoe:

Zie 4.10 scientific documentation.

Zie PC-Ring en correlatie data per ring.

Dit wordt in PC-Ring verwerkt in InCor.f90.

Dit moet bij voorkeur voor heel Nederland. Dus een grote transformatie matrix.

OK, om te checken of ik dit goed begrijp zie andere bijlage.

Deze actie kan worden uitgevoerd na het inlezen van de data, of pas wanneer over regiogrenzen wordt gestapt. Deze laatste optie is nauwkeuriger, maar lastiger uit te leggen en te programmeren.

Meer in detail: De vertaalslag kan gedaan worden als eerste bewerking op de resultaten uit het mechanisme programma. Alle resultaten van de combinaties zijn dan uitgedrukt in de transposed variabelen.

Als alternatief kan men ook vertaalslag uitvoeren bij het oprollen over de ring als een regiem grens wordt overschreden. De vertaalslag komt dan net voor het oprollen tot de ring/area.

#### 4.12.3 Combine alternatives

Alternatives represent the different possible subsoil scenarios.

Input: The results of HydraRing calculations for:

- Mechanism
- Section
- Layer
- Alternative
- Presentation section

Let op: De berekening in het mechanisme programma berekent voor een combinatie van mechanisme, sectie, laag en alternatief. Hierbij wordt de doorsnede opgeschaald naar de sectielengte, Echter als een sectie tot meerdere presentatie secties behoort, wordt ook uitvoer gegeven voor de verschillende lengten binnen de afzonderlijke presentatiesecties.

In deze subroutine gaat het niet over het oprollen van de secties naar de presentatiesecties.

Procedure: Multiply the different alternatives with the probability of occurrence and combine them ("or") as not correlated.

Note: This procedure differs from PC-Ring, but is more accurate.

Output: The results of HydraRing calculations for:

- Mechanism
- Section
- Layer
- Presentation section (ook over gehele sectie)

#### 4.12.4 Combine layers

Layers represent the different possible sub layers. For instance for revetments, the dike may be riveted with grass, stones, asphalt, etc.

Input: The results of HydraRing calculations for:

Mechanism  
Section  
Layer  
Presentationsection

Procedure: Combine ("or") the different revetments. The different stochastic variables between the different layers are fully correlated if they refer to the same natural variable.

Note: In the definition of stochastic variables account should be made with this assumption. So, e.g. there should only be one variable "grain size diameter" if this variable is thought to be the same for all layers. At the same time, e.g. the width of the revetment layer should be defined as separate variables if they refer to different layers, otherwise they will be incorrectly fully correlated in the model.

Output: The results of HydraRing calculations for:

Mechanism  
Section  
Presentation section

#### 4.12.5 Combine mechanisms

Mechanisms represent the different possible mechanisms which belong to one main mechanism.

Input: The results of HydraRing calculations for:

Mechanism  
Section  
Presentationsection

Procedure: Combine ("or") the different mechanisms. The different stochastic variables are fully correlated if they refer to the same natural variable.

Note: In the definition of stochastic variables account should be made with this assumption.

Output: The results of HydraRing calculations for:

MainMechanism  
Section  
Presentationsection

#### 4.12.6 combineSections

Some sections combined give the presentation section.

Input: The results of HydraRing calculations for:

MainMechanism  
Section  
Presentationsection

Procedure: Combine ("or") the different sections within a presentation section. The different stochastic variables that refer to the same natural variable (e.g. grain size) are partially correlated, since they refer to different locations. The partial correlation is equal to the residual correlation,  $\rho_Z$ , of the spatial correlation model (see section 2.5.5 of the Scientific documentation). The partial correlation is the same for each pair of sections, no matter if they are neighboring sections or more remote sections.

Note: This procedure may be replaced/extended later with a procedure with a different length effect in which correlation depends on the distance between sections. Such a procedure would be more physically correct and opens doors for the use of smaller sections, i.e. smaller than the correlation length of the Z-function. However, this will also lead to an increase of computation times. It needs to be verified if this increase in computation time is acceptable and correct.

Output: The results of HydraRing calculations for:

MainMechanism  
Presentationsection

#### 4.12.7 combineMainMechanisms

Within a presentation section the combination of the main mechanisms give the probability of the presentation section.

Input: The results of HydraRing calculations for:

MainMechanism  
Presentationsection

Procedure: Combine ("or") the different main mechanisms. The different stochastic variables are partially correlated.

Note: This procedure may be replaced/extended later with a procedure with a different length effect. Different correlation dependent on the distance between sections.

Output: The results of HydraRing calculations for:

Presentationsection

The procedures above result in the failure probability of the presentation sections. Now this can be combined to the dike ring (or more general: area).

#### 4.12.8 combineAreaMainMechanism

The combination of the full length sections of a main mechanism gives the failure probability of the area/dike ring due to a given main mechanism.

Input: The results of HydraRing calculations for:

MainMechanism  
Section (full section length of main mechanism, so no presentation section!)

Procedure: Combine ("or") the different sections of the main mechanism. The different stochastic variables are partially correlated.

Note: This procedure may be replaced/extended later with a procedure with a different length effect. Different correlation dependent on the distance between sections.

Output: The results of HydraRing calculations for:

MainMechanism

#### 4.12.9 combineArea

The results of the complete area/dike ring can now be calculated from the results for the different presentation sections or the results for the different main mechanisms. The one with the highest reliability index should be used.

For the correlation the partial correlation should be used.

#### 4.12.10 Consequences for Hydra-Ring

In order to apply the method as described above, some adaptations need to be made in the part of Hydra-Ring that is executed prior to Combin. The main activities are:

[A] upscale results of representative cross sections to presentation sections. Theoretically, a dike segment can be spread out over multiple presentation sections. In practice, however, presentation segments will be at least the scale of the dike segments, which means one dike segment is spread out over 2 presentation segments at most. The representative cross section of a dike segment therefore needs to be upscaled to:

1. The dike section
2. The overlap of the dike section with the first presentation segment
3. The overlap of the dike section with the second presentation segment

Steps 2 and 3 are additional to the current version of Hydra-Ring.

[B] combine all (sub)sections of the presentation section to derive results for the entire presentation section.

[C] store additional alpha-values (and corresponding betas). Values for all combinations of Mechanms, Sections, Layers, Alternatives and Presentation Sections need to be stored.

## 5 Failure mechanisms

### 5.1 Introduction

A flood defence system can fail due to different failure mechanisms. An overview of relevant mechanisms for levees is shown in Figure 5.1.

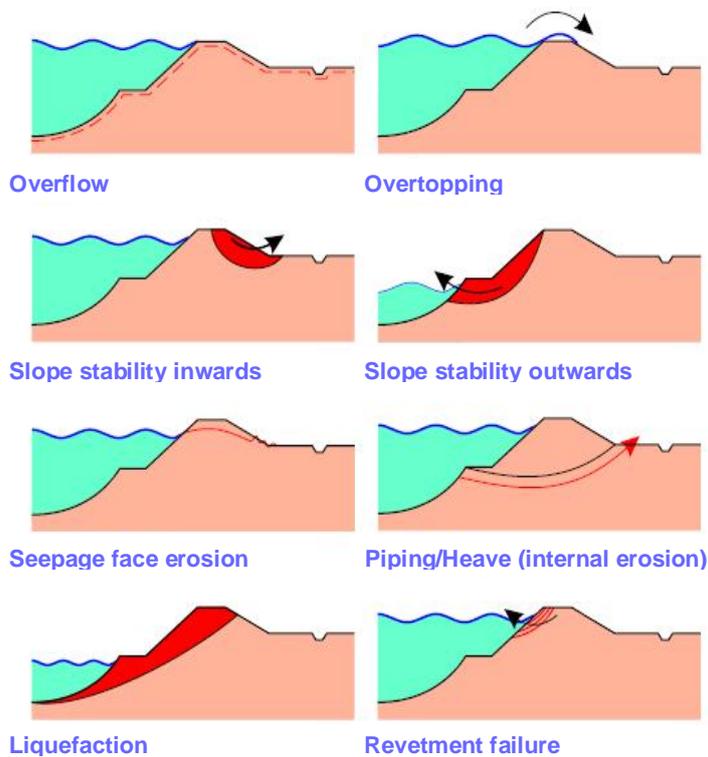


Figure 5.1 Illustration of different failure mechanisms for dikes.

Failure due to a particular failure mechanism can sometimes depend on the occurrence of more than one *sub-mechanism*. An example is failure due to piping (internal erosion of sand underneath the levee or structure), which can only occur if heave (of the protecting cover layer at the land side) first occurs, see Figure 5.2.

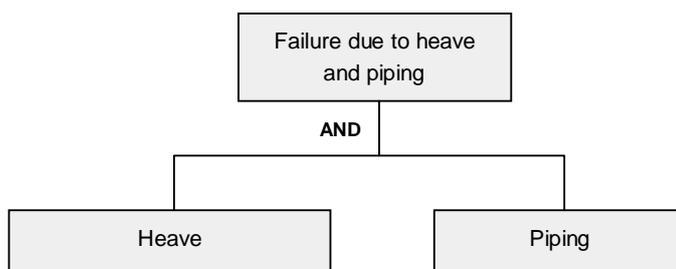


Figure 5.2 Example of the fault tree for failure mechanism heave and piping

The probability of sub-mechanism failure is predicted on the lowest level by *cross sectional models*, where a cross section refers to a (transectional) cross section of a flood defence segment. These models require input data, for example hydrodynamic data, the cross section geometry, the revetment geometry, the subsoil geometry and the physical parameters of soil and revetment.

This chapter describes the failure mechanisms that are programmed in Hydra-Ring. For each mechanism, the formulas that describe them are presented and briefly explained in the following sections:

- overflow/overtopping
- sliding of the inner slope
- heave & piping
- revetment erosion and subsequent internal erosion
- overflow/overtopping of hydraulic structures
- closing failure of hydraulic structures
- piping of hydraulic structures
- structural failure of hydraulic structures
- dune erosion

The models for each of the mechanisms are presented in the form of a limit state function, which is useful for probabilistic computations. A limit state function is defined such that values greater than zero indicate no failure, and values less than zero indicate failure. Together with the statistical load models and local input data, failure probabilities can be computed for each dike section and for each failure mechanism, and can subsequently be combined to determine the failure probability of the dike ring (see Chapter 2 for details on the combination).

## 5.2 Models of failure mechanisms in probabilistic failure computations

### 5.2.1 Introduction

In chapter 2 it was described that Hydra-Ring uses the following two-step procedure to quantify the probability of failure of a dike ring:

1. quantification of the probability of all the individual components of the system, and
2. integration of the failure probabilities of the components to derive the failure probability of the entire system (system analysis).

A single component in Hydra-Ring refers to a combination of one cross section, one failure mechanism, one wind direction, one closure scenario and one relatively small (<1 day) time interval during which load conditions are assumed to be constant.

### 5.3 Overtopping and Overflow

Overflow and overtopping are considered to be one of the main failure mechanisms. For this mechanism, water passing over the crest of the dike either due to overtopping or overflow is the cause of erosion loading on the inside slope. In Hydra-Ring, water discharges due to overflow are assumed to be relevant only in case of offshore wind and wave heights smaller than 1 mm. Water discharges in other situations are assumed to occur due to wave overtopping.

In the following sections, the formulas to compute the failure mechanism overtopping and overflow are presented. For background information on the derivation of these formulas see [Technical Report Wave Run-up and Wave Overtopping at Dikes, TAW, 2002].

Table 5.1 gives an overall view of the steps that have to be followed to compute both failure mechanisms.

Table 5.1 Overview of the general calculation procedure for overtopping and overflow

Calculation procedure	Variables to compute (in order)	Section
<i>Overtopping</i>		
1. Optional: Calculation of the critical overtopping discharge ( $q_c$ ), if it is not direct input	$T_e \rightarrow V_c \rightarrow q_c$	5.3.6
2. Optional: Calculation of the wave height ( $H_s$ ), wave period ( $T_s$ ) and peak period ( $T_p$ ) with Bretschneider, if it is not available from the hydraulic boundary conditions database	$H_s \rightarrow T_s \rightarrow T_p$	3.5.2.2
3. Calculation of the combined reduction factor ( $\gamma$ )	$\gamma_f \rightarrow \gamma_b \rightarrow \gamma_s \rightarrow \gamma_B \rightarrow \gamma$	5.3.4.6
4. Calculation of the overtopping discharge ( $q_o$ )	$R_b \rightarrow Q_b \rightarrow S_{op} \rightarrow q_o$	5.3.4.7
5. Calculation of the failure mechanism overtopping ( $Z$ )	$Z$	5.3.2
<i>Overflow</i>		
1. Optional: Calculation of the critical overflow discharge ( $q_c$ ), if it is not direct input	$T_e \rightarrow V_c \rightarrow q_c$	5.3.6
5. Calculation of the failure mechanism overflow ( $Z$ )	$Z$	5.3.1

The meaning of the variables to be computed is explained in the sections to which is referred.

Table 5.5 and Table 5.6 give an overview of the input parameters for overtopping and overflow.

Table 5.2 Overview of the variables to compute the failure mechanism overtopping

	Description	Unit
$m_{qc}$	Model factor critical discharge	[-]
$q_c$	Critical discharge	[m <sup>3</sup> /s]
$m_{qo}$	Model factor overtopping discharge	[-]
$q_o$	Overtopping discharge	[m <sup>3</sup> /s]
$P_t$	Percentage of time that overtopping/overflow takes place	[-]

Table 5.3 Overview of the variables to compute the failure mechanism overflow

	Description	Unit
$h$	Outer water level	[m+NAP]
$h_d$	Dike height	[m]
$\Delta h_c$	Critical height difference	[m]
$q_c$	Critical overtopping discharge	[m <sup>3</sup> /s]
$h_B$	Outer bank height	[m+NAP]

### 5.3.1 Limit State Function Overflow

The overflow mechanism is considered in the presence of offshore wind and wave heights with less than 1 mm. Figure 5.3 illustrates this mechanism.

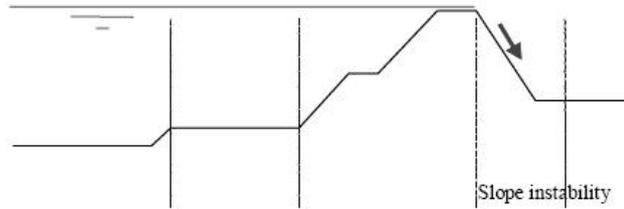


Figure 5.3 Dike failure mechanism overflow

In this mechanism, failure occurs when the water level rises above the critical height of the dike  $h_{kd}$ . The limit state function of this mechanism given by:

$$Z = h_{kd} - h \quad (4.1)$$

where  $h_{kd}$  is the critical height of the dike and  $h$  is the local water level. The critical height of the dike  $h_{kd}$  is the sum of the dike height  $h_d$  and the critical height difference  $\Delta h_c$ .

$$h_{kd} = h_d + \Delta h_c \quad (4.2)$$

The limit state function can be rewritten as:

$$Z = h_d + \Delta h_c - h \quad (4.3)$$

The critical height difference  $\Delta h_c$  is a function of the critical overtopping discharge  $q_c$ . Assuming that overflow occurs, from the critical overtopping discharge  $q_c$  (determined as in section 5.3.6), the critical height difference  $\Delta h_c$  can be computed as:

$$q_c = 0,6 \cdot \sqrt{g \cdot \Delta h_c^3} \quad \Leftrightarrow \quad \Delta h_c = \sqrt[3]{\frac{q_c^2}{0,36 \cdot g}} \quad (4.4)$$

where  $g$  is the gravitational force and the limit state function can be expressed as,

$$Z = h_d + \sqrt[3]{\frac{q_c^2}{0,36 \cdot g}} - h \quad (4.5)$$

### 5.3.2 Limit State Function Overtopping

The overtopping failure mechanism occurs when at a certain location the amount of water, as a result of waves and water level, is higher than the level that the dike crest and inner slope can handle. The water discharge passing over the crest leads to erosion, which if not taken care of on time, can cause a breach in the dike and flooding may occur.

The limit state function that describes this failure mechanism is given by:

$$Z = m_{q_c} q_c - m_{q_o} q_o / P_t \quad (4.6)$$

where  $q_c$  is the critical discharge expressing the allowed discharge,  $q_o$  is the actual occurring overtopping discharge,  $m_{q_c}$  and  $m_{q_o}$  are the model factors describing the uncertainty of the models for calculating  $q_c$  and  $q_o$ . The occurring discharge is divided by the percentage of the time  $P_t$  that overtopping takes place which accounts for the wave periodicity. In the overflow mechanism  $P_t$  is assumed to be one.

In the implementation of this mechanism, the geometry of the dike is taken into account, from toe to toe. Hydra-Ring assumes that this geometry can consist of the following components:

- one toe
- one outer slope between the toe and the berm
- one berm
- one outer slope between the berm and the outer crest line
- one crest
- one inner slope

The available models for calculating the occurring and allowed overtopping discharge are the following:

- Overtopping discharge as described in the manual by van der Meer with some adjustments from the software PC-Overslag [section 5.3.3]
- Formulas from Strickler/Manning to determine the speed of discharge at the inner slope , and CIRIA formulas to determine the strength of the grass revetment (the user can chose other relations) [5.3.6]

### 5.3.3 Introduction to the overtopping discharge models

In Hydra-Ring, two methods are available to compute the overtopping discharge. Section 5.3.4 describes the approach presented in “Technical report wave run-up and wave overtopping at dikes” [REF] of the technical advisory committee on flood defence in The Netherlands. This method is referred to as PC-Overslag (the program that implemented this approach). Section 5.3.5 describes the slightly different implementation in the program PC-Ring [REF], where some simplifications to the standard design have been made. Both implementations share several calculation steps. The description of the PC-Ring approach will therefore be limited to the part that is different.

Both methods base their computation on differentiating between two scenarios: Overtopping discharge for breaking waves and overtopping discharge for non-breaking waves. In the first scenario, overtopping increases by increasing the so-called breaker parameter  $\xi_{op}$  ( $\xi_{op} \gamma_b < \approx 2$ ). The second scenario is developed for the maximum level that can be achieved with non-breaking waves ( $\xi_{op} \gamma_b > \approx 2$ ). The overtopping discharge is assumed to be the minimum between these two cases. In the following subsections the computation of the reduction factor  $\gamma_b$  is and the breaker parameter is explained.

### 5.3.4 Overtopping: PC-overslag approach

In the box below, the computation steps used in PC-overslag are mentioned as well as the sections where these steps are explained. The whole computation starts by initializing the variables  $H_{m0}$  and  $T_{m-1,0}$  at the toe of the dike (as a pre-processing step).

1. Calculate influence factor for angle of wave attack $\gamma_{\beta,z}$ for wave run-up	section 5.3.4.2
2. Adjust wave conditions $H_{m0}$ , $T_{m-1,0}$ if $\beta > 80$ degrees	section 5.3.4.2
3. Iterate until 2% wave run-up reaches an equilibrium	
a. Calculate average slope, $\tan \alpha$	section 5.3.4.2
b. Calculate $Z_{2\%,smooth}$ (smooth: for $\gamma_b = 1$ and $\gamma_f = 1$ )	section 5.3.4.1
c. Calculate influence factor for roughness on slope $\gamma_f$	section 5.3.4.5
d. Calculate $Z_{2\%,rough}$ (rough: for $\gamma_b = 1$ )	section 5.3.4.1
e. Calculate influence factor for berms $\gamma_b$	section 5.3.4.4
f. Calculate 2% wave run-up	section 5.3.4.1
4. Calculate influence factor for roughness on slope $\gamma_f$	section 5.3.4.5
5. Calculate influence factor for angle of wave attack $\gamma_{\beta}$ for wave overtopping	section 5.3.4.2
6. Adjust influence factors in case $\gamma_b \cdot \gamma_f \cdot \gamma_{\beta} < 0.4$ , so that $\gamma_b \cdot \gamma_f \cdot \gamma_{\beta} = 0.4$	section 5.3.4.6
7. Calculate wave overtopping with the computed $\gamma_b$ and $\gamma_f$	section 5.3.4.7

#### 5.3.4.1 Computation of the z2% wave run-up level

The general formulas for calculating the z2% wave run-up level are:

$$z_{2\%} = H_{m0} \cdot f_{run-up1} \cdot \gamma_b \cdot \gamma_f \cdot \gamma_{\beta} \cdot \xi_0 \quad (4.7)$$

with a maximum of:

$$z_{2\%} = H_{m0} \cdot \gamma_f \cdot \gamma_{\beta} \cdot \left( 4.3 - \frac{1.6}{\sqrt{\xi_0}} \right) \quad (4.8)$$

where:

$H_{m0}$	significant wave height
$\gamma_b$	influence factor for a berm
$\gamma_f$	influence factor for roughness components
$\gamma_{\beta}$	influence factor for angle of wave attack
$\xi_0$	breaker parameter

The breaker parameter is computed with the following fomula:

$$\xi_0 = \frac{\tan \alpha_{repr}}{\sqrt{S_{op}}}, \quad (4.9)$$

in which  $S_{op}$  represents the wave steepness:

$$S_{op} = \frac{2\pi H_s}{g T_p^2}, \tag{4.10}$$

A graphical representation of these formulas is shown in Figure 5.4. The first formula applies to situations with low breaker parameters (approx.  $\xi_0 < 1.8$ , in figure tagged with 'formula 3a'), the second formula is valid for higher values of  $\xi_0$ .

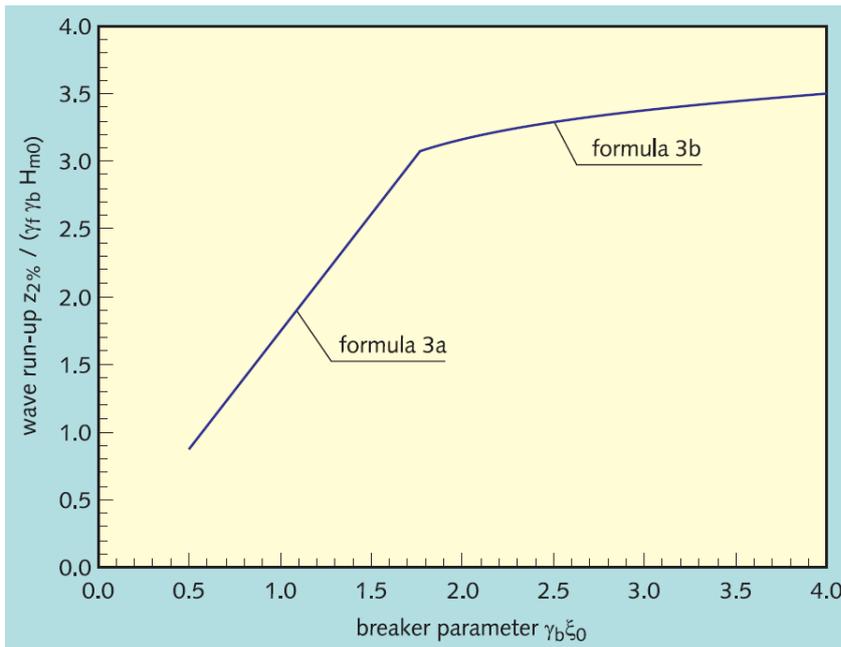


Figure 5.4 Wave run-up as function of break parameter

Within Hydra-Ring both formulas are evaluated, where the final wave run-up level is determined by calculating the minimum wave run-up level of both formulas.

The value for the  $z_{2\%}$  wave run-up level is calculated by means of iteration. The iteration ends and produces an answer if the following condition is met:

$$\left| \frac{z_{i+1} - z_i}{(z_i + z_{i+1})/2} \right| < 5 \cdot 10^{-5} \tag{4.11}$$

where:

- $z_i$   $z_{2\%}$  wave run-up level of iteration step  $i$
- $z_{i+1}$   $z_{2\%}$  wave run-up level of iteration step  $i+1$ .

The value for the average slope angle depends on the  $z_{2\%}$  wave run-up level, which is calculated by an iteration loop. In order to calculate the average slope in the first iteration step, an initial value for the  $z_{2\%}$  is needed. As initial value within Hydra-Ring  $z_{2\%} = 1.5H_{m0}$  is used.

#### 5.3.4.2 Calculate influence factor for angle of wave attack $\gamma_\beta$

Influence factors or correction factors, are quantities that affect the final overtopping discharge. In this case, the angle of incidence of wave attack  $\beta$ , defined as the angle between the direction of wave propagation and the normal to the dike axis  $\psi$ , has a direct influence in the final computation. On average, the waves should have the same direction as the wind. To capture the deviation of the waves from the wind direction, a stochastic variable  $\varphi$  is added to the wind direction. The angle of wave attack is described as follows:

$$\beta = |\varphi - \psi|, \quad \text{if } \beta > 180^\circ \text{ then } \beta = 360^\circ - \beta. \quad (4.12)$$

With onshore waves ( $0^\circ \leq \beta \leq 180^\circ$ ) the influence factors for wave overtopping are computed as:

$$\begin{aligned} \gamma_\beta &= 1 - 0,0033 \beta & 0^\circ \leq \beta \leq 80^\circ \\ \gamma_\beta &= 1 - 0,0033 \cdot 80 & \beta > 80^\circ \end{aligned} \quad (4.13)$$

To compute the 2% wave run-up the reduction is determined as,

$$\begin{aligned} \gamma_{\beta;z} &= 1 - 0,0022 \beta & 0^\circ \leq \beta \leq 80^\circ \\ \gamma_{\beta;z} &= 1 - 0,0022 \cdot 80 & \beta > 80^\circ \end{aligned} \quad (4.14)$$

For  $80^\circ < \beta \leq 110^\circ$  the wave height  $H_{m0}$  and the wave period  $T_{m-1,0}$  are adjusted as follows:

- $H_{m0}$  is multiplied by  $\frac{110 - \beta}{30}$  (4.15)

- $T_{m-1,0}$  is multiplied by  $\sqrt{\frac{110 - \beta}{30}}$  (4.16)

For  $110^\circ < \beta \leq 180^\circ$  then  $H_{m0} = 0$ , which results in wave run-up  $z_{2\%} = 0$  and wave overtopping  $q_0 = 0$ .

#### 5.3.4.3 Computation of average slope

In most cases a dike slope does not consist of an entirely straight slope, but of sections with various slopes and often with one or more berms. The wave run-up formula requires a characteristic slope. A representative slope angle is used for this slope, ignoring any berms. The influence of berms is considered separately (in section 5.3.4.4).

The representative slope angle is computed with the following formula:

$$\tan \alpha_{repr} = \frac{1.5H_{m0} + z_{2\%}}{L_{slope} - B} \quad (4.17)$$

where:

- $H_{m0}$  significant wave height at the toe of the dike;
- $z_{2\%}$  wave run-up height exceeded by 2% of the incoming waves
- $L_{slope}$  horizontal length between two points on the slope,  $z_{2\%}$  above and  $1.5H_{m0}$  below SWL (Still Water Line)
- $B$  width of the berm measured horizontally.

In Figure 5.5 (below) these variables are shown.

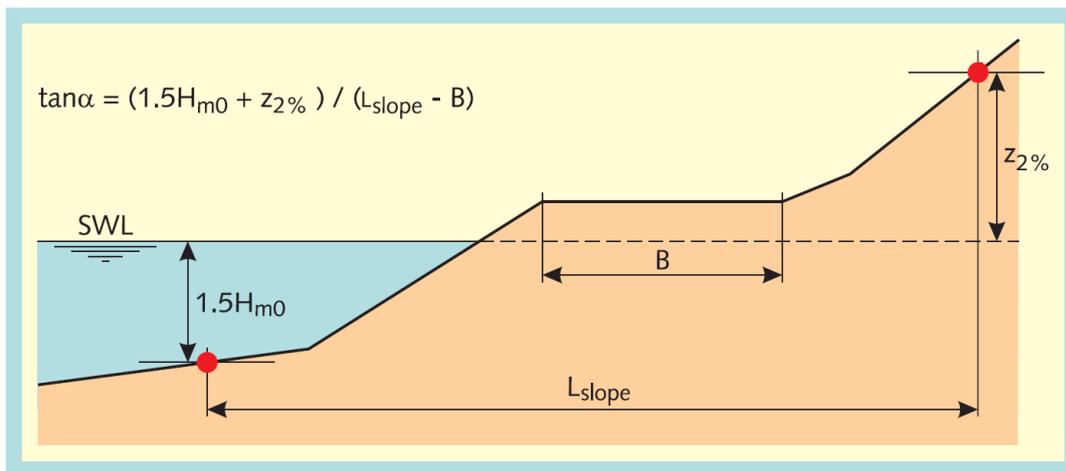


Figure 5.5 Determination of the characteristic slope for a cross-section consisting of various slope sections, excluding any berm influence

The dike cross sections in Hydra-Ring can only consist of 3 sections: a lower slope, berm and upper slope. With this schematization three different situations can be distinguished:

1. Only the lower slope is relevant in determining the average slope angle. This situation occurs if the level  $SWL+z_{2\%}$  stays below the height of the berm. The characteristic slope angle is represented by the lower slope angle alone. The formula for  $\tan \alpha_{repr}$  then reduces to:  $\tan \alpha_{repr} = \tan \alpha_{lower\ slope}$
2. Only the upper slope is relevant in determining the average slope angle. This situation occurs if the level  $SWL-1.5H_{m0}$  stays above the height of the berm. The characteristic slope angle is represented by the upper slope angle alone. The formula for  $\tan \alpha_{repr}$  then reduces to:  $\tan \alpha_{repr} = \tan \alpha_{upper\ slope}$
3. The level  $SWL-1.5H_{m0}$  is on the lower slope and the level  $SWL+z_{2\%}$  is on the upper slope. The.

The first two cases are straightforward. The average slope angle is equal to either lower or upper slope. The procedure for determining the average slope angle in the third case is more complicated. This procedure is described below.

First the  $SWL-1.5H_{m0}$  and  $SWL+Z_{2\%}$  levels are examined. If the  $SWL-1.5H_{m0}$  level lies below the toe of the dike, the toe of the dike is used as representative lower boundary.

$$\text{lower boundary} = \max(SWL-1.5H_{m0}, \text{height of toe})$$

For the upper boundary the minimum of the  $SWL+Z_{2\%}$  and dike height is used:

$$\text{upper boundary} = \min(SWL+Z_{2\%}, \text{height of crest})$$

The average slope angle is equal to the vertical height difference divided by the horizontal distance minus the horizontal berm length. As the profiles can only consist of three sections the average slope angle can also be calculated with:

$$\tan \alpha_{repr} = \frac{\Delta h_{lower\ slope} + \Delta h_{upper\ slope}}{b_{lower\ slope} + b_{upper\ slope}} \quad (4.18)$$

where:

$\Delta h_{lower\ slope}$	height difference between $SWL-1.5H_{m0}$ and berm
$\Delta h_{upper\ slope}$	height difference between $SWL+Z_{2\%}$ and berm
$b_{lower\ slope}$	horizontal distance between $SWL-1.5H_{m0}$ and intersection between berm and lower slope
$b_{upper\ slope}$	horizontal distance between $SWL+Z_{2\%}$ and intersection between berm and upper slope.

In case that  $SWL-1.5H_{m0}$  is below the toe, the height of the toe should be used for calculating  $\Delta h_{lower\ slope}$  (instead of  $SWL-1.5H_{m0}$ ). This also applies to the upper slope (use the minimum of  $SWL+Z_{2\%}$  and the dike height). These distances are shown in Figure 5.6.

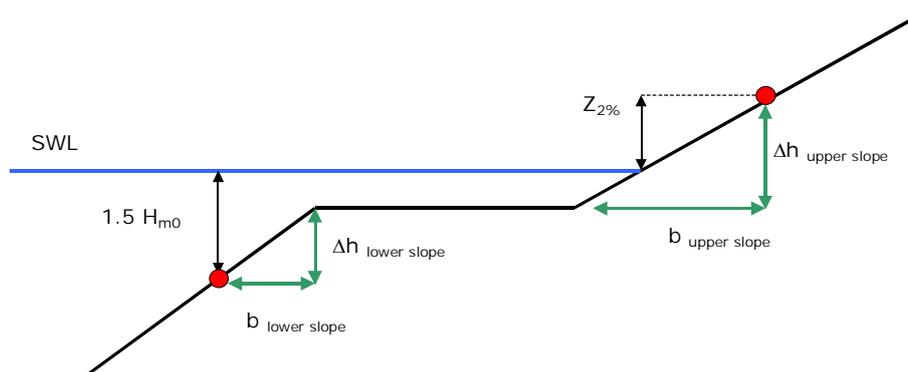


Figure 5.6 Relevant distances for determining average slope

In wave run-up and overtopping calculations the berms need to be schematized as horizontal sections. This means that sloping berms need to be rotated until a horizontal berm is created. The advantage of this rotating is that the above describes procedure is also valid in case the berm is not exactly horizontal (for computation of the average slope).

### 5.3.4.4 Influence of berms

The position of berms in relation to the still water line (SWL) has influence on wave run-up and wave overtopping. In wave run-up formulas, the influence of berms is calculated with a separate berm factor, in which two influences are taken into account:

- Width of the berm (parameter  $r_b$ )
- Depth of the berm in relation to the water level (parameter  $r_{dh}$ ).

These parameters are used in the following formula:

$$\gamma_b = 1 - r_b(1 - r_{dh}) \quad (4.19)$$

The influence of the width of the berm can be calculated with:

$$r_b = 1 - \frac{2H_{m0} / L_{berm}}{2H_{m0} / (L_{berm} - B)} = \frac{B}{L_{berm}} \quad (4.20)$$

This factor determines which part of the cross section is horizontal and is considered as berm (see Figure 5.7).

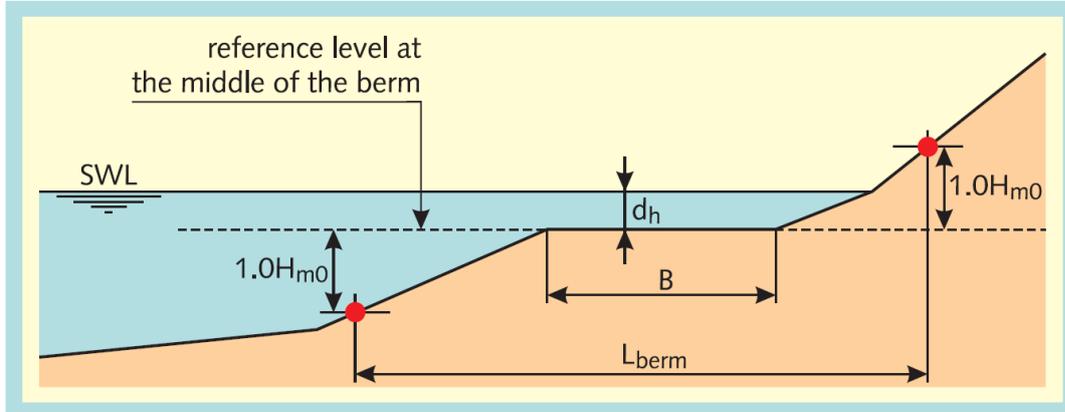


Figure 5.7 Influence of berm in wave run-up calculations

In Hydra-Ring the factor  $r_b$  is calculated with the following steps:

1. Determine the berm width
2. Determine the lower influence level the berm
3. Determine the upper influence level the berm
4. Determine the horizontal length between the height level of the berm minus  $H_{m0}$  and plus  $H_{m0}$ .

#### Step 1:

Only horizontal berm can be used in wave run-up calculations. Sloping berms need to be rotated. The average height level of the berm is determined with the average of lower and upper boundary of the berm. This will be the 'new' schematised berm level. The upper and

lower slope need to be connected to this berm, which means the width of the original berm width will decrease (also see Figure 5.8).

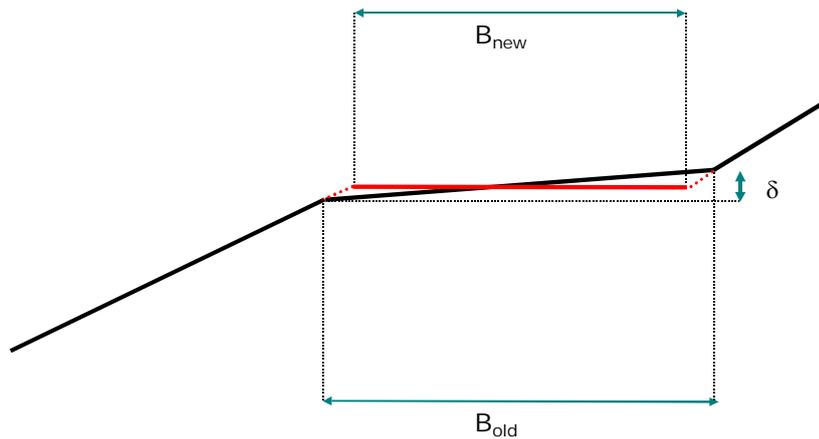


Figure 5.8 Conversion of sloping berm to horizontal berm,  $\delta$  is the height difference in the berm

In this figure the new horizontal berm width is calculated with:

$$B_{new} = B_{old} - \frac{\delta / 2}{\alpha_{lower\ slope}} - \frac{\delta / 2}{\alpha_{upper\ slope}} \quad (4.21)$$

Step 2:

In order to determine  $L_{berm}$  two levels are needed: the height of the berm minus  $H_{m0}$  and height of the berm plus  $H_{m0}$ . In case the toe of the dike or crest of the dike lies outside this interval, either the toe of the dike or crest level of the dike is used. The lower influence level of the berm can therefore be calculated as:

$$\max(h_{berm} - H_{m0}, h_{toe}) \quad (4.22)$$

Step 3:

The upper influence level can then be determined with:

$$\min(h_{berm} + H_{m0}, h_{crest\ of\ dike}) \quad (4.23)$$

Step 4:

The value for  $L_{berm}$  can now be calculated with.

$$L_{berm} = B_{new} + \frac{\min(H_{m0}, h_{berm} - h_{toe})}{\tan \alpha_{lower\ slope}} + \frac{\min(H_{m0}, h_{crest} - h_{berm})}{\tan \alpha_{upper\ slope}} \quad (4.24)$$

where:

$B_{new}$	horizontal berm width;
$h_{berm}$	height level of horizontal berm;
$h_{toe}$	height level of the toe of the dike;
$\alpha_{lower\ slope}$	slope angle of the lower slope;

$h_{\text{crest}}$  height level of the crest of the dike;  
 $\alpha_{\text{upper slope}}$  slope angle of the upper slope.

The influence of the depth of the berm is represented by the factor  $r_b$ :

$$r_{dh} = 0.5 - 0.5 \cos\left(\pi \frac{d_h}{x}\right) \quad (4.25)$$

where:

$d_h$  berm depth in relation to SWL (negative values indicated berm is above SWL)

and where:

$$\begin{aligned} x &= z_{2\%} && \text{if } z_{2\%} > -d_h > 0 \\ x &= 2 H_{m0} && \text{if } 2H_{m0} > d_h \geq 0 \\ r_{dh} &= 1 && \text{if } -d_h \geq z_{2\%} \text{ or } d_h \geq 2H_{m0} \end{aligned} \quad (4.26)$$

The influence of the depth of the berm can be divided in three parts:

- The berm level lies below SWL ( $x = z_{2\%}$ ).
- The berm level lies above SWL ( $x = 2 H_{m0}$ ).
- The berm level lies far beneath or above the berm level and has no influence on wave run-up ( $r_{dh} = 1$ ).

In Hydra-Ring this division is implemented, where the value for  $z_{2\%}$  is evaluated before being used in this formula. If the crest height is lower than the  $z_{2\%}$  wave run-up level, the difference between the crest height and the water level is used. This evaluating is only used for cases, where the SWL is below the berm. On the other hand, if the SWL is above the berm and 2 times  $H_{m0}$  is below the toe of the dike no correction is used for  $x$ . This parameter is still equal to  $2H_{m0}$ .

After calculation of  $\gamma_b$ , a final evaluation is performed, as the value of  $\gamma_b$  should satisfy:

$$0.6 \leq \gamma_b \leq 1.0 \quad (4.27)$$

If  $\gamma_b$  is less than 0.6,  $\gamma_b$  is set equal to 0.6.

#### 5.3.4.5 Influence of roughness

The influence of the roughness of a revetment on a dike for wave run-up is given by the influence factor  $\gamma_b$ . For numerous types of revetment standard roughness factors exist. However the relation between wave height and roughness of the slope is also relevant in wave run-up calculations.

The roughness factors that are known are valid for situation where  $\gamma_b \cdot \xi_0 < 1.8$ . For values larger than 1.8 the influence factor  $\gamma_f$  increases linear to 1 for  $\gamma_b \cdot \xi_0 = 10$ .

The mentioned value 1.8 is the value where the product  $\gamma_b \cdot \xi_0$  gives the point where the two lines of Figure 5.4 intersect. This boundary point can be computed using the two functions for wave run-up:

$$H_{m0} \cdot f_{\text{run-up1}} \cdot \gamma_b \cdot \gamma_f \cdot \gamma_\beta \cdot \xi_0 = H_{m0} \cdot \gamma_f \cdot \gamma_\beta \cdot \left( f_{\text{run-up2}} - \frac{f_{\text{run-up3}}}{\sqrt{\xi_0}} \right) \quad (4.28)$$

which results in:

$$f_{\text{run-up1}} \cdot \gamma_b \cdot \xi_0 = f_{\text{run-up2}} - \frac{f_{\text{run-up3}}}{\sqrt{\xi_0}} \quad (4.29)$$

The boundary value for  $\gamma_b \cdot \xi_0$  depends on values for  $\gamma_b$ . In Hydra-Ring the boundary value is approximated by

$$\xi_{\text{boundary}} = f_2 \cdot \gamma_b^2 + f_1 \cdot \gamma_b + f_0 \quad (4.30)$$

where

$\gamma_b$	influence factor for a berm
$f_0, f_1, f_2$	model factors.

This boundary values is used to determine if the roughness factor is used for the influence factor for roughness or that a higher influence factor should be used. If the value for  $\xi_0$  is larger than the boundary breaker parameter  $\xi_{\text{boundary}}$ , the following formula is used:

$$\gamma_f = \gamma_{f,\text{ref}} + (1 - \gamma_{f,\text{ref}}) \cdot \frac{\gamma_b \cdot \xi_0 - \xi_{\text{boundary}}}{10 - \xi_{\text{boundary}}} \quad (4.31)$$

where

$\gamma_{f,\text{ref}}$	original roughness factor (as valid for $\gamma_b \cdot \xi_0 < 1.8$ )
$\gamma_b$	influence factor for a berm
$\xi_0$	breaker parameter
$\xi_{\text{boundary}}$	boundary breaker parameter

A final evaluation of the value of  $\gamma_f$  is performed, so that the value of the influence factor for roughness equals 1 for  $\gamma_b \cdot \xi_0 \geq 10$ .

#### 5.3.4.6 Computation for the total reduction factor

A requirement of the total reduction factor ( $\gamma_b \cdot \gamma_f \cdot \gamma_\beta$ ) is that the value of this total factor is larger than 0.4. If this is not the case, then a minimum factor of 0.4 should be used. There is however a difference in wave run-up calculation for situations with low and high values for the

breaker parameter. As the actual wave run-up is calculated by determining the minimum of both formulas, a difference in the partial reduction factors can occur. In the formula for high values of the breaker parameter, the influence factor for berms is not present. In these cases the partial influence parameters are adjusted with the following routine. A new variable representing the total reduction factor is introduced,  $\gamma_t$ :

$$\begin{aligned}\gamma_t &= \gamma_f \cdot \gamma_b \cdot \gamma_\beta \\ f_f &= \frac{1 - \gamma_f}{1 - \gamma_{f,\min}} \\ f_b &= \frac{1 - \gamma_b}{1 - \gamma_{b,\min}} \\ f_\beta &= \frac{1 - \gamma_\beta}{1 - \gamma_{\beta,\min}} \\ f_t &= f_f + f_b + f_\beta\end{aligned}\tag{4.32}$$

where:

$$\begin{aligned}\gamma_{f,\min} &= 0.55 && \text{the minimal influence factor for roughness,} \\ \gamma_{b,\min} &= 0.6 && \text{the minimal influence factor for berms,} \\ \gamma_{\beta,\min} &= 0.736 && \text{the minimal influence factor for angle of wave attack.}\end{aligned}$$

These factors are used to calculate adjusted partial influence factors.

$$\begin{aligned}\gamma_f' &= \gamma_f \cdot \exp\left(\frac{f_f}{f_t} \cdot \ln\left(\frac{0.4}{\gamma_t}\right)\right) \\ \gamma_b' &= \gamma_b \cdot \exp\left(\frac{f_b}{f_t} \cdot \ln\left(\frac{0.4}{\gamma_t}\right)\right) \\ \gamma_\beta' &= \gamma_\beta \cdot \exp\left(\frac{f_\beta}{f_t} \cdot \ln\left(\frac{0.4}{\gamma_t}\right)\right)\end{aligned}\tag{4.33}$$

It is easy to verify with the above formulas that the total influence factor is equal to 0.4 if the total influence factor was originally less than 0.4:

$$\gamma_t' = \gamma_f' \cdot \gamma_b' \cdot \gamma_\beta' = 0.4\tag{4.34}$$

#### 5.3.4.7 Computation of the overtopping discharge $q_o$

The overtopping discharge  $q_o$  is computed with the following formula, but only if the breaker parameter is less or equal to 5:

$$q_o = \min(Q_b, Q_n) \sqrt{gH_{m0}^3} \quad \xi_0 \leq 5\tag{4.35}$$

where:

$Q_b$	dimensionless overtopping discharge for breaking waves;
$Q_n$	dimensionless overtopping discharge for non-breaking waves;
$g$	gravity force;
$H_{m0}$	significant wave height;
$\xi_0$	breaker parameter.

In shallow water (i.e. breaker parameter is greater than or equal to 7) an adapted formula is used:

$$q_o = Q_{shallow} \sqrt{gH_{m0}^3} \quad \xi_0 \geq 7 \quad (4.36)$$

where:

$Q_{shallow}$  dimensionless overtopping discharge for shallow water.

Between the breaker parameters 5 and 7 logarithmic linear interpolation is used:

$$q_o = \exp\left(\ln(Q_n|_{\xi_0=5}) + \left(\ln(Q_{shallow}|_{\xi_0=7}) - \ln(Q_n|_{\xi_0=5})\right) \frac{\xi_0 - 5}{7 - 5}\right) \sqrt{gH_{m0}^3} \quad 5 < \xi_0 < 7 \quad (4.37)$$

In this last formula it is sufficient to only use  $Q_n$  for non-breaking waves. The transition between breaking and non-breaking waves is below  $\xi_0 \gamma_b = 2$ . For  $\xi_0 \gamma_b > 2$  only non-breaking waves are relevant. The minimum value for  $\gamma_b$  is equal to 0.6, therefore the scenario with a breaker parameter between 5 and 7 only applies to non-breaking waves.

#### Dimensionless overtopping discharge for breaking waves $Q_b$

The dimensionless overtopping discharge for breaking waves ( $Q_b$ ) is computed with the following formula:

$$Q_b = \frac{0.067}{\sqrt{\tan \alpha_{repr}}} \gamma_b \xi_0 \exp\left(-f_b \frac{h_k - h}{H_{m0}} \frac{1}{\xi_0 \gamma_\beta \gamma_b \gamma_f}\right) \quad (4.38)$$

where:

$h_k$	height of the outer crest line;
$h$	local water level;
$H_{m0}$	significant wave height at the toe of the dike;
$\alpha_{repr}$	average angle of the outer slope;
$f_b$	model factor for breaking waves (= -4.3);
$\gamma_\beta$	reduction factor for the angle of wave attack;
$\gamma_f$	reduction factor for roughness of the slope;
$\gamma_b$	reduction factor for the influence of the berm;
$\xi_0$	breaker parameter.

#### Dimensionless overtopping discharge for non-breaking waves $Q_n$

The dimensionless overtopping discharge for non-breaking waves ( $Q_n$ ) is computed with the following formula:

$$Q_n = 0.2 \exp\left(-f_n \frac{h_k - h}{H_{m0}} \frac{1}{\gamma_\beta \gamma_f}\right) \quad (4.39)$$

where:

- $h_k$  height of the outer crest line;
- $h$  local water level;
- $H_{m0}$  significant wave height at the toe of the dike;
- $f_n$  model factor for non-breaking waves (= -2.3);
- $\gamma_\beta$  reduction factor for the angle of wave attack;
- $\gamma_f$  reduction factor for roughness of the slope.

Dimensionless overtopping discharge for shallow water  $Q_{shallow}$

The dimensionless overtopping discharge for shallow water ( $Q_{shallow}$ ) is computed with the following formula:

$$Q_{shallow} = f_{shallow} \exp\left(-\frac{h_k - h}{\gamma_\beta \gamma_f H_{m0} \cdot (0.33 + 0.022 \xi_0)}\right) \quad (4.40)$$

where:

- $h_k$  height of the outer crest line;
- $h$  local water level;
- $H_{m0}$  significant wave height at the toe of the dike;
- $f_{shallow}$  model factor for shallow waves(= 0.21);
- $\gamma_\beta$  reduction factor for the angle of wave attack;
- $\gamma_f$  reduction factor for roughness of the slope;
- $\xi_0$  breaker parameter

### 5.3.5 Overtopping: PC-Ring approach

The PC-Ring implementation of the van der Meer model [REF] is almost identical to the PC-Overslag approach. There are only some subtle differences which are here mentioned. The reason for supporting this implementation is only to be able to reproduce PC-Ring results. In the box below, the computation steps used by PC-Ring are mentioned as well as the sections where these steps are explained. Note that there references to the procedure used in PC-overslag as many computation steps of this approach are still used by PC-Ring.

1. Calculate influence factor for angle of wave attack $\gamma_{\beta,z}$ for wave run-up	section 5.3.4.2
2. Adjust wave conditions $H_{m0}$ , $T_{m-1,0}$ if $\beta > 80$ degrees	section 5.3.4.2
3. Iterate until 2% wave run-up reaches an equilibrium	
a. Calculate average slope, $\tan \alpha$	section 5.3.4.2
b. Calculate $Z_{2\%,smooth}$ (smooth: for $\gamma_b = 1$ and $\gamma_f = 1$ )	section 5.3.4.1
c. Calculate influence factor for roughness on slope $\gamma_f$	section 5.3.5.1
d. Calculate $Z_{2\%,rough}$ (rough: for $\gamma_b = 1$ )	section 5.3.4.1
e. Calculate influence factor for berms $\gamma_b$	section 5.3.5.2
f. Calculate 2% wave run-up	section 5.3.4.1

4. Calculate influence factor for roughness on slope $\gamma_f$	section 5.3.5.1
5. Calculate influence factor for angle of wave attack $\gamma_\beta$ for wave overtopping	section 5.3.4.2
6. Adjust influence factors in case $\gamma_b \cdot \gamma_f \cdot \gamma_\beta < 0.4$ , so that $\gamma_b \cdot \gamma_f \cdot \gamma_\beta = 0.4$	section 5.3.4.6
7. Calculate wave overtopping with the computed $\gamma_b$ and $\gamma_f$	section 5.3.4.7

#### 5.3.5.1 Influence of roughness

In Table 5.4, the roughness values for different types of slopes are presented. The reduction factor  $\gamma_f$  for the roughness of the slope is equal to the roughness value for  $\xi_{op} < 3$ . From  $\xi_{op} = 3$ , the roughness factor a linear function with 1 as limit for  $\xi_{op} = 5$ . For higher values, the reduction factor remains equal to 1.

Table 5.4 Roughness coefficients

Reference type	$\gamma_f$ (roughness coefficients)
Concrete	1.0
Asphalt	1.0
Closed concrete block	1.0
Grass	1.0
Vilvoorden stone	0.85
Basalt	0.90
Haringman	0.90
Fixtone-open stone asphalt	0.90
Armoflex	0.90
Small blocks over 1/25 of surface	0.85
Small blocks over 1/9 of surface	0.80
¼ of block revetment 10 cm higher	0.90
Ribs (optimum dimensions)	0.75
Armour rock - two layers thick	0.55
Armour rock – single layer	0.70

#### 5.3.5.2 Influence of berms

##### Geometry

In the PC-Ring approach, a simple dike geometry is employed. The dike profile must be described as follows:

- one straight outer slope between the toe and the berm
- one horizontal berm
- one straight outer slope between the berm and the crest of the dike
- de height of the outer crest line

In this approach, the inner and outer slope, de berm height and width and the height of the outer crest line are considered to be stochastic variables. Additionally, the roughness factor is chosen as the roughness of the outer slope between the berm and the crest of the dike.

##### Berm factor $\gamma_b$

The influence of the berm width is described by the change in the slope. The influence in the berm location must be described between the space  $2H_s$  under the still water line up to  $Z_{2\%}$  on

the lower slope. In this model, it is assumed that the berm is horizontal. Thus, the factor  $r_B$  in equation (4.20) for a horizontal berm with width  $B$  is written as:

$$r_B = \frac{1}{1 + H_s / (B \tan \alpha_{u,b}) + H_s / (B \tan \alpha_{u,o})} \quad (4.41)$$

The influence of the berm location with respect to the water still line is defined as,

$$r_{dh} = 0,5 \left( \frac{d_h}{H_s} \right)^2 \quad \text{for} \quad 0 \leq r_{dh} \leq 1 \quad (4.42)$$

where the middle of the berm lies at a depth  $d_h$  below the still water line. The water depth at the berm location is the difference between the local water level  $h$  and the height of the berm  $h_B$ :

$$d_h = h - h_B \quad (4.43)$$

These formulas for the reduction factor  $\gamma_b$  can only be applied if the local water level  $h$  is similar to berm height. The bounds for its application are estimated to be:

$$-0,5 H_s \leq d_h \leq +0,5 H_s \quad (4.44)$$

### 5.3.6 Critical overtopping discharge (CIRIA model)

The critical overtopping discharge  $q_c$  for Hydra-Ring is in general a direct user input.

As an alternative, Hydra-Ring can also determine the critical overtopping discharge based on the grass strength, using a model from the CIRIA research [8].

$$q_c = \frac{v_c^3}{\tan \alpha_i C^2} \quad (4.45)$$

where  $v_c$  is the critical discharge speed,  $\alpha_i$  is the angle of the inner slope and  $C$  is the roughness factor of Chézy. The roughness factor of the inner slope is determined with the Strickler relation as,

$$C = 25 \left( \frac{q_c}{k v_c} \right)^{1/6} \quad (4.46)$$

where  $k$  is the roughness factor in the inner slope (see Table 5.4). An alternative approach is the Manning relation with roughness factor  $n$ . These two expressions are coupled by the formula,

$$n = \frac{k^6}{25} \quad (4.47)$$

Combining the formulas of Chézy and Strickler, the critical overtopping discharge is computed as,

$$q_c = \frac{v_c^{5/2} k^{1/4}}{125 \tan \alpha_i^{3/4}} \quad (4.48)$$

In The Netherlands, Chézy roughness factor is computed with the relation of White-Colebrook. In Hydra-Ring, Strickler expression is used as this is the standard approach in other countries. Additionally, Strickler expression offers the possibility to solve analytically for the critical discharge  $q_c$ .

The critical discharge speed  $v_c$ , that after the time period  $t_e$ , causes the grass layer to fail is given by,

$$v_c = f_g \frac{3,8}{(1 + 0,8^{10} \log t_e)} \quad (4.49)$$

where  $f_g$  is a factor determined by the quality of the grass layer (its ranges from between 0.7 for bad quality to 1.4 for good quality of grass),  $t_e$  is the time period measured in hours and  $v_c$  is the critical discharge speed measured in m/s.

The quality factor  $f_g$  (see table Table 5.5) and the erosion resistance  $c_g$  of the grass layer are properties of the grass layer with a strong correlation. For this reason, in Hydra-Ring the following relations are used for these two variables:

$$c_g = 6 \cdot 10^5 f_g^{1,5} \quad \text{or:} \quad f_g = \left( \frac{c_g}{6 \cdot 10^5} \right)^{2/3} \quad (4.50)$$

For the residual strength of the protective clay layer (blanket layer) it is assumed that the erosion of grass ( $t_{RT, \text{outside}}$ ) given the clay layer ( $t_{RK, \text{outside}}$ ) in the outer talud (see mechanism damage revetment) behaves in the same way as the erosion of grass ( $t_{RT, \text{inside}}$ ) given the clay layer ( $t_{RK, \text{inside}}$ ) in the inner slope,

$$\frac{t_{RT, \text{outside}}}{t_{RK, \text{outside}}} = \frac{t_{RT, \text{inside}}}{t_{RK, \text{inside}}} \quad (4.51)$$

Making use of the model to determine the damage in the revetment in the outer slope, it is possible to compute the relation between the erosion of the grass revetment and the protective clay layer in the outer slope:

$$\frac{t_{RT,outside}}{t_{RK,outside}} = \frac{c_g d_w}{0,4 c_{RK} L_K} \quad (4.52)$$

where  $d_w$  is the depth of the rooting of the grass,  $c_g$  is the erosion resistance of the grass layer,  $c_{RK}$  is a coefficient determined by the erosion resistance of the protective layer and  $L_K$  is the width of the protective clay layer. Values of  $d_w = 0,1$  m and  $c_{RK} = 23000$  ms [7], are assumed for the depth of the grass rooting and for the erosion resistance.

In order for the dike to fail, the grass revetment and the protective clay layer have to erode within the storm period  $t_s$ ,

$$t_s \geq t_{RT,inside} + t_{RK,inside} \quad (4.53)$$

As input for the grass strength model in CIRIA, the relevant time period  $t_e$  for grass erosion is needed. The relevant storm period can be derived from the expressions above as,

$$t_s = t_{RT,inside} + \frac{(0,4 c_{RK} L_{K,inside})}{c_g d_w} t_{RT,inside} = \frac{c_g d_w + (0,4 c_{RK} L_{K,inside})}{c_g d_w} t_{RT,inside} \quad (4.54)$$

$$\Leftrightarrow t_{RT,inside} = \frac{c_g d_w}{c_g d_w + (0,4 c_{RK} L_{K,inside})} t_s$$

The value of  $L_K$  is assumed to be zero when no residual strength or liquefaction is taken into account. Otherwise,  $L_K$  can be determined making use of the thickness of the protective clay layer and the angle of the slope as shown in (insert figure or equation).

Taking into account the amount of time  $P_t$  that overtopping or overflow take place, it is possible to determine the strength of the grass layer applying the CIRIA relation,

$$t_e = P_t t_{RT,inside} \quad (4.55)$$

For overflow,  $P_t$  is equal to 1. For overtopping, the value of  $P_t$  can be provided or computed with an overtopping model. As a remark, in order to compute  $P_t$ , the change in the wave pulse should be taken into account.

## 5.3.7 Relevant variables for overtopping and overflow failure mechanisms

In Table 5.5 and Table 5.6 an overview of the variables that are used to compute the failure mechanism overtopping and overflow are presented.

Table 5.5 Overview of the variables to compute the failure mechanism overtopping

	Description	Unit	ID
$m_{qc}$	Model factor critical discharge	[-]	
$q_c$	Critical discharge	[m <sup>3</sup> /s]	
$m_{qo}$	Model factor overtopping discharge	[-]	
$q_o$	Overtopping discharge (PC-overslag/PC-Ring)	[m <sup>3</sup> /s]	
$P_t$	Percentage of time that overtopping/overflow takes place	[-]	

Table 5.6 Overview of the variables to compute the failure mechanism overflow

	Description	Unit	ID
$h$	Outer water level	[m+NAP]	
$h_d$	Dike height	[m]	
$\Delta h_c$	Critical height difference	[m]	
$q_c$	Critical overtopping discharge	[m <sup>3</sup> /s]	
$h_B$	Outer bank height	[m+NAP]	

## 5.4 Failure mechanism ‘Macro stability’

### 5.4.1 Introduction

Dike failure due to the mechanism macro-instability inner slope (also referred to as e.g. sliding of the inner slope) occurs when the inner slope of the dike becomes unstable and begins to slide due to the development of water pressures inside and/or below the dike. Although it can occur both at the inner side and at the outer side, we generally only consider macro instability of the inner slope. Macro instability of the outer slope usually occurs and quickly dropping water levels; Macro instability of the inner slope is considered the more dangerous type. Failure is defined as a breach that occurs as the result of macro instability of the inner slope, see Figure 5.9. Failure occurs after instability of the inner slope (Stab) and a breach following the macro instability (StPf). The MPROSTAB model is used to calculate the probability of this failure mechanism conditional to outside water level. With the aid of the results of the MPROSTAB calculations, PC-Ring can subsequently determine the total failure probability for the failure mechanism macro instability of the inner slope.

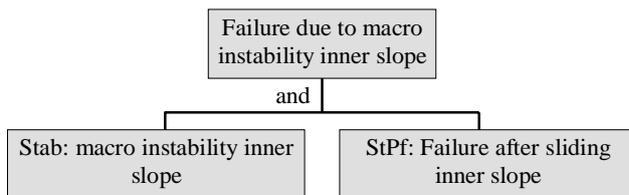


Figure 5.9 Fault tree analysis

### 5.4.2 Sub-mechanism Stab: Macro instability inner slope

#### 5.4.2.1 MPROSTAB

##### General

The Bishop method of moments is used to calculate the mechanism of macro instability of the inner slope. The method assumes moment equilibrium of the whole failure plane and vertical forces equilibrium of the separate slices. The horizontal forces equilibrium is not guaranteed. In addition, the model assumes that macro instability will result in immediate failure without any kind of residual strength.

Bishop's method is a 2D consideration in the plane perpendicular to the longitudinal axis of the slope. It is assumed that the failure plane has an infinitely long cylindrical shape. As it actually has a finite length, contributions to the resisting moment of the failure plane's edge can be of influence. However, this has not been taken into account in the (deterministic) stability analysis.

For the probabilistic stability analysis conform MPROSTAB, Bishop's method has been used as standard computational model. In the probabilistic stability analysis, it is assumed that the failure plane has a finite length. The contributions to the resisting moment of the failure plane's edge can, therefore, be taken into account. In addition, both the length of the failure plane and the contribution of the failure plane's edges to the resisting moment are considered stochastic quantities.

##### Shear strength

The shear strength computation is based on the drained parameters  $c'$  (cohesion) and  $\varphi'$  (friction angle), according to the formula:

$$\tau = c' + (\sigma - u) \tan(\varphi') \quad (4.56)$$

in which  $\tau$  is the effective shear strength along the underside of a slice,  $\sigma$  the total normal stress on the underside of a slice and  $u$  the water(over)pressure. The shear strength parameters  $c'$  and  $\varphi'$  can be determined using triaxial tests, cell tests, direct or simple shear tests, etc.

The computed values of the shear strength parameters from the test data depend on the strain level of the tested soil sample with which the shear strength parameters are determined.

In triaxial tests, a strain level of 2%, 5% or 10% is used, or peak values of the shear strength are used to determine the shear strength parameters. The used strain level, or peak level, determines the found  $c'$  en  $\varphi'$ .

In cell tests, the strain level at which the shear strength parameters are determined is not set explicitly, but it ranges somewhere between 1% and 2%. This influences the interpretation of the physical appearance of limit state exceedance with respect to stability. If we use peak shear strength, the limit state exceedance can be identified with the actual instability of the inner slope. When we use small strain levels, the limit state exceedance has to be interpreted as *the development of excessive* deformations that will eventually result in macro instability of the inner slope. In other words, the physical appearance of the limit state exceedance with regard to stability is not univocal, but it depends on the test and testing procedure chosen to determine the shear strength parameters. A complication is that the shear strength parameters at a certain strain level may not be used unambiguously to determine the shear strength parameters at a higher strain level or the peak shear strength parameters. In probabilistic stability analysis this can be compensated by using a computational model uncertainty factor depending on the type and performance of the test.

#### *Water pressure*

Bishop's method of moments, on which MPROSTAB is based, is a so-called "effective stress analysis". This implicates that the water pressure in the soil body is given. Within MPROSTAB this is done by specifying piezometric lines. For the top layer - or at least for the first partly-saturated layer - one piezometric line is defined. The programme interprets this line as the phreatic head. The water pressures in the underlying layers are assumed hydrostatical, unless piezometric lines have been defined for these layers (always two per layer, the first for the top and de second for the bottom layer). The hydraulic head in a vertical of such a layer is linearly interpolated between the hydraulic head at the top (first piezometric line) and the bottom of the layer (second piezometric line). This way, water over pressure can be specified. The bottom layer (usually deep sand layer) *is not restricted downwards*. The hydraulic head in a vertical of such a bottom layer is considered constant, and the water level therefore hydrostatical, and equals the hydraulic head at the top of that layer (first piezometric line for this layer).

A special situation arises during construction phases in which the layers are not fully adjusted to the layers above (incomplete consolidation). Water over pressure can then be specified

using adjustment percentages. For flood risk computations with Hydra Ring, a complete adjustment of water pressures (a consolidated or long-term situation) is assumed.

Water pressures occurring in the soil body given a certain outer water level should be determined conform the usual procedures, such as calculating the groundwater flow, geohydrological calculations while taking into account uplift or uplift head difference, any (extrapolated) measurements, etc. (see Guidelines for the design of River Dikes and “Technisch Rapport Waterspanningen in Dijken”).

The instability probability computed with MPROSTAB is, in principle, a conditional instability probability given the outer water level and water pressure in the soil body that has been determined as a function of the outer water level. PC-Ring also includes a method to calculate the probability of failure while taking into account the probability distribution function of the outside water level.

### *Equilibrium analysis, stability factor*

The local stability factor (that is the stability factor belonging to a single cross section, calculated according to Bishop’s method) can be determined based on the modelling of shear stress and water (over) pressure. The deterministic stability factor of a slip circle  $\Gamma$  comes from the limit equilibrium:

$$M_R = M_A$$

$$M_R = \frac{R \sum (c_i + (\sigma_i - u_i) \tan \phi_i) \Delta b_i}{\Gamma} \quad (4.57)$$

In which  $M_A$  is the driving moment of the slip circle concerned (with radius  $R$ ) and  $M_R$  the resisting moment at  $M_r$  is calculated by summing the mobilised shear strength along the slip circle and multiplying it with the radius of the circle. The slip circle has been divided in  $N$  slices to calculate the mobilised shear strength, with lengths  $\Delta b_i$  for the bottom of the slices along the slip circle. Perpendicular to the bottom of the slices, the water pressure  $u_i$  and soil pressure  $\sigma_i$  are acting. The soil pressures follow from a vertical equilibrium of forces affecting the slices, which are: the force due to its own weight, the force due to stress perpendicular to and the shear stress along the bottom of the slice. The stability factor is determined iteratively.

The critical slip circle is the slip circle with the smallest stability factor, to be determined iteratively. The stability factor of the slope equals the stability factor of the critical slip circle.

### *Reliability function probabilistic analysis*

The limit state function of the probabilistic stability analysis can thus be written as:

$$Z = \Gamma - q \quad (4.58)$$

in which the threshold value  $q$  must be equal to 1,0, provided that there is no computational model uncertainty. In the probabilistic stability analysis, we use a stochastic threshold value of which the mean and standard deviation depend on the type of tests used to assess the shear stress parameters (as mentioned above).

#### 5.4.2.2 Influence stochastic research outside water level

With the program MPROSTAB, the failure probability of macro instability of the inner slope of a dike ring is calculated conditional to an outside water level and during the total computational period. By doing this calculation for a number of water levels, it is possible to generate an input file with which the failure probability of the mechanism of macro instability can be computed accurately by integrating over the probability density of the water level.

The statistics needed to determine the local water level is processed using the mechanism of macro instability of the inner slope in Hydra Ring (including closing regime). The limit state function thus becomes:

$$Z = \beta(h) + \sum_{i=1}^{n_{MPROSTAB}} \alpha_i(h) u_i \quad (4.59)$$

in which  $\beta$  is the reliability index according to MPROSTAB conditional to water level  $h$ ,  $\alpha_i$  are the FORM influence coefficients given this water level resulting from the MPROSTAB calculations and  $u_i$  the variables with a standard normal distribution.

Table 5.7 Overview variables sub-mechanism STAB: Macroinstability inner slope.

	Description	unit	ID
	variability cohesion	[kN/m <sup>2</sup> ]	2001
	uncertainty mean value cohesion	[kN/m <sup>2</sup> ]	2002
	variability $\tan(\varphi)$	[m]	2003
	uncertainty mean value $\tan(\varphi)$	[-]	2004
	correlations cohesion and $\tan(\varphi)$	[-]	2005
	uncertainty water (over) pressure	[m+NAP]	2006
	uncertainty phreatic line	[m+NAP]	2007
	model uncertainty	[-]	2008

#### 5.4.3 Sub-mechanism StPf: Failure probability after macro instability inner slope

When macro instability of the inner slope occurs, it does not always result in dike failure. This is expressed in the failure probability given the occurrence of the mechanism macro instability: StPf. The standard value for this probability is  $P_{stab\ StPf} = 1,0$ , which implies dike failure give instability of the inner slope. Lower values may be used in case it can be shown the water retaining function remains intact after macro instability of the inner slope.

Given the failure probability, the following reliability index can be found:

$$\beta_{stab} = -\Phi^{-1}(P(StPf)) \quad (4.60)$$

The following limit state function can then be applied:

$$Z = \beta_{stab} - u \quad (4.61)$$

in which  $u$  is the standard normal distributed variable with mean 0 and a standard deviation 1,0.

Table 5.8 Overview variables sub-mechanism StPf: failure probability after macro instability inner slope.

	description	unit	ID
Pstpf	Failure probability after macro instability inner slope	[-]	2009

5.4.4 Variables for the mechanism of macro instability inner slope  
**Error! Reference source not found.** presents an overview of the distributions of variables in the sliding models. For the variables related to the hydraulic load, please refer to chapter 3 and 4.

Table 5.9 Variables for the mechanism macro instability inner slope

ID		description	unit	type	Parameters		spatial distribution		Variation over time	
					location	distribution	dx	ρx	Δt	ρt
2001		variability cohesion	[kN/m <sup>2</sup> ]	-	-	-	-	0	-	1
2002		uncertainty mean value cohesion	[kN/m <sup>2</sup> ]	-	-	-	-	0	-	1
2003		variability tan(φ)	[m]	-	-	-	-	0	-	1
2004		uncertainty mean value tan(φ)	[-]	-	-	-	-	0	-	1
2005		correlations cohesion and tan(φ)	[-]	-	-	-	-	0	-	1
2006		uncertainty water (over) pressure	[m+NAP]	-	-	-	-	0	-	1
2007		uncertainty phreatic line	[m+NAP]	-	-	-	-	0	-	1
2008		model uncertainty	[-]	-	-	-	-	0	-	1
2009	Pstpf	failure probability after macro instability inner slope	[-]	det	nom	-	-	-	-	-

## 5.5 Piping and heave

### 5.5.1 Description

When the mechanism of uplift and piping occur, the dike fails because soil particles are being washed out. The water pressure first leads to bursting (uplift) of the cohesive blanket layers of the dike (or aquitard). The resulting water flow causes the development of pipe shaped erosion channel (piping) in non-cohesive sand layer below the dike (aquifer). The complete failure mechanism is thus composed of two sub-mechanisms. Failure only occurs if for both mechanisms the resistance drops below the load, or, formally:

$$\{\text{Failure}\} = \{\text{uplift AND piping}\}$$

Figure 5.10 provides a schematic description in the shape of a fault tree.

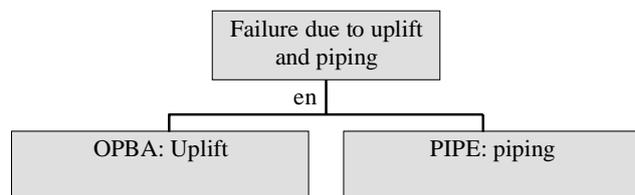


Figure 5.10: Fault tree uplift and piping

The guidelines TAW (1994) and the Technical Report Sand Boils (TAW, 2002) have been used to describe this mechanism. In addition, the following figure presents the definition of a number of variables. For the complete definitions of the various sub-mechanisms, please refer to the subsequent sections.

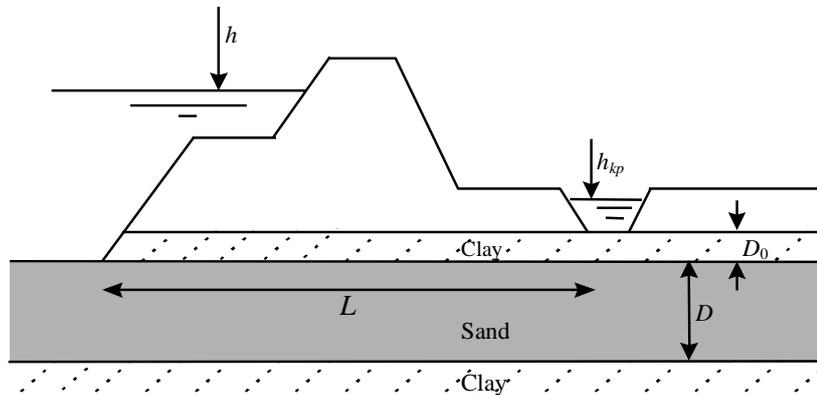


Figure 5.11 Definitions of variables of uplift and piping in a typical cross-section of a dike

in which:

$D_0$  is the thickness of the cohesive blanket layer [m];

$D$  is the thickness of the permeable sand layer [m];

$h$  is the local outer water level [m+NAP];

$h_{kp}$  is the inner water level [m+NAP];

$L$  is the length of the seepage (the length between the entrance and exit point, which is usually minimally the width of the flood defence) [m].

Further instructions to determine the length of the seepage path are given in the Safety Assessment Regulation (Ministerie Verkeer en Waterstaat, 2006)

The thickness of the cohesive blanket layer is used to compute the mechanism of uplift (see section 5.5.2). This value is equal to the vertical seepage length with piping (see section 5.5.3).

The inner water level is the same as the water level in the ditch or, in case there is no ditch, the surface level. In addition to the various models on uplift and piping, it is also possible to compute several piping scenarios for one location. The probability of uplift and piping for each scenario is computed separately and multiplied with the piping scenario probability to obtain the total failure probability of uplift and piping.

#### 5.5.2 Sub-mechanism ZfOPBA: Uplift

When the mechanism of uplift occurs, the water pressures in the sand layer leads to bursting (uplift) of the cohesive blanket layer (aquitarde). The maximum pressure in this layer can be expressed as a critical water level. The blanket layer will burst if the difference between the local water level  $h$  and the inside water level  $h_{kp}$  exceeds the critical water level difference  $h_c$ . This gives the following limit state function:

$$Z = m_o h_c - m_h (h - h_{kp}) \quad (4.62)$$

This limit state function also includes two model factors,  $m_o$  and  $m_h$ . The model factor  $m_o$  represents the model uncertainty with which to determine the critical water level during uplift. The damping factor,  $m_h$ , indicates the level of damping (i.e. head difference reduction).

The critical water level difference  $h_c$  is a function of the wet volumetric weight of the cohesive blanket layer  $\gamma_{nat}$ , the volumetric weight of water  $\gamma_w$ , and the thickness of the cohesive blanket layer  $D_0$ :

$$h_c = \frac{\gamma_{nat} - \gamma_w}{\gamma_w} D_0 > 0 \quad (4.63)$$

If  $D_0$  has a mean value of 0, it is assumed that there is no cohesive blanket layer present. The probability of failure of the mechanism uplift then equals 1.

Table 5.10 Overview variables sub-mechanism ZfOPBA: uplift.

	Description	Unit	ID
$m_o$	Model factor uplift	[-]	3012
$h_c$	critical water level difference	[m]	-
$m_h$	Damping factor	[-]	3014
$h$	outside water level	[m+NAP]	-
$h_{kp}$	inside water level	[m+NAP]	99007
$\gamma_{nat}$	Wet volumetric weight cohesive blanket layer	[kN/m <sup>3</sup> ]	3010
$\gamma_w$	volumetric weight water	[kN/m <sup>3</sup> ]	99008
$D_0$	thickness cohesive blanket layer	[m]	3001

### 5.5.3 Sub-mechanism ZfPIPE: Piping

When the mechanism of piping occurs, the increasing outside water pressure leads to the development of a pipe shaped erosion channel from the exit point of the water to the entrance point. The maximum resistance against piping in the sand layer can be expressed as a critical water level. The dike will fail as a consequence of piping if the difference between the local water level  $h$  and the inside water level  $h_{kp}$ , reduced with a part of the vertical length of the seepage path exceeds the critical water level  $h_p$ . A smaller difference may lead to the development of piping but will not lead to dike failure.

For the sub-mechanism piping, the following two models five models have been implemented in Hydra Ring:

1. The calculation rule of Sellmeijer (2007) with  $c$  determined with the specific permeability  $k$ ;
2. The calculation rule of Sellmeijer (2007) with  $c$  determined with Bear's constant

We will further discuss these models in the following section.

#### 5.5.3.1 Sellmeijer's rule

TAW (2002) is used to describe the mechanism of piping, which, is based on Sellmeijer's rule.

The limit state function can be written as:

$$Z = m_s h_p - (h - 0,3D_0 - h_{kp}) \quad (4.64)$$

This limit state function includes the model factor  $m_s$ . This model factor  $m_s$  is the model uncertainty with which, for piping, the critical water level over the dike is determined using Sellmeijer's rule.

The critical water level  $h_p$  is presented in Sellmeijer's rule by the following relation:

$$h_p = \alpha c L \left( \frac{\gamma_k - \gamma_w}{\gamma_w} \right) (0,68 - 0,1 \ln c) \tan \theta > 0 \quad (4.65)$$

The critical gradient is determined by the factor  $\alpha$ , the coefficient  $c$ , the length of the seepage path  $L$ , the volumetric weight of the sand grains  $\gamma_k$  (27 kN/m<sup>3</sup>), the volumetric weight of water  $\gamma_w$  and the bedding angle  $\theta$  of the sand.

The factor  $\alpha$  reflects the effect of the limited thickness of the aquifer:

$$\alpha = \left( \frac{D_1}{L} \right)^{\frac{0,28}{\left( \left( \frac{D_1}{L} \right)^{2,8} - 1 \right)}} \quad (4.66)$$

The coefficient  $c$  is determined using the properties of the sand in the aquifer subject to erosion.

$$c = \eta d_{70} \left( \frac{1}{\kappa L} \right)^{\frac{1}{3}} \quad (4.67)$$

in which  $\eta$  is the drag force factor (White's constant),  $d_{70}$  represents the 70% percentile grain size of the sand in the aquifer (the mesh size of a fictitious sieve through which 70 percent of the weight of the sand samples passes),  $\kappa$  the intrinsic permeability and  $L$  the seepage length.

The intrinsic permeability can be determined in a number of ways. For instance, the specific permeability  $k_{z;b}$  [m/s] results in the following relation:

$$\kappa = \frac{\nu}{g} k_{z;b} \quad (4.68)$$

in which  $\nu$  is the kinematic viscosity ( $1,33 \cdot 10^{-6}$  m<sup>2</sup>/s for water at 10°C) and  $g$  the gravitational constant 9,81 m/s<sup>2</sup>.

If we use this term in the equation for the coefficient  $c$ , we find:

$$c = \eta d_{70} \left( \frac{g}{\nu k_{z;b} L} \right)^{\frac{1}{3}} \quad (\text{model 1}) \quad (4.69)$$

if the permeability is unknown,  $\kappa$  can be estimated based on the representative of the small grain size  $d_{10}$  [m] of the sand. This relation is as follows:

$$\kappa = C_{Bear} d_{10}^2 \quad (4.70)$$

in which  $C_{Bear}$  is the constant of Bear ( $C_{Bear}$  ranges between  $0,5 \cdot 10^{-3}$  for silty sand and  $1,2 \cdot 10^{-3}$  for clean sand).

If we use this term in the equation for the coefficient  $c$ , we find:

$$c = \eta d_{70} \left( \frac{1}{C_{Bear} d_{10}^2 L} \right)^{\frac{1}{3}} = \eta \left( \frac{1}{C_{Bear}} \left( \frac{d_{70}}{d_{10}} \right)^2 \frac{d_{70}}{L} \right)^{\frac{1}{3}} \quad (\text{model 2}) \quad (4.71)$$

in which  $(d_{70} / d_{10})$  is a measure for the uniformity of the sand. Both statements for the coefficient  $c$  have been implemented in PC-Ring. Note that for common piping problem, the  $d_{70}$  of the upper part of the sand layer is used, while the permeability of the aquifer should be used, in which case the  $d_{70} / d_{10}$  measure may not be used

Table 5.11 Overview variables sub-mechanism ZfPIPE: Piping with the calculation rule of Sellmeijer

	description	unit	ID
$m_s$	modelfactor Sellmeijer	[-]	3013
$h_p$	critical gradient	[m]	-
$h$	outside water level	[m+NAP]	-
$D_0$	thickness cohesive blanket layer	[m]	3001
$h_{kp}$	inside water level	[m+NAP]	99007
$\alpha$	Geometry factor	[-]	-
$c$	Erosion coefficient	[-]	-
$L$	length of the seepage path	[m]	3004
$\gamma_k$	volumetric weight sand grain material	[kN/m <sup>3</sup> ]	3011
$\gamma_w$	volumetric weight water	[kN/m <sup>3</sup> ]	99008
$\theta$	bedding angle	[°]	3005
$D_1$	thickness upper sand layer	[m]	3002
$\eta$	White's constant	[-]	3009
$d_{70}$	70% percentile grain size upper sand layer	[m]	3007
$\kappa$	intrinsic permeability	[m <sup>2</sup> ]	-
$\nu$	kinematic viscosity	[m <sup>2</sup> /s]	-
$g$	gravitational constant	[m/s <sup>2</sup> ]	-
$k_{z,b}$	permeability upper sand layer	[m/s]	3015
$C_{Bear}$	Factor Bear	[-]	3006
$d_{10}$	10% grain size sand	[m]	-
$d_{70}/d_{10}$	uniformity	[-]	3008

### Variables for the mechanism of uplift/piping

Table 5.12 presents an overview of the distribution of variables in piping models.

Table 5.12 Variables for the mechanism uplift/piping.

ID		description	Unit	Type	Parameters		Spatial distribution		Variation over time	
					location	distributio	$d_x$	$\rho_x$	$\Delta_t$	$\rho_t$
3001	$D_0$	thickness cohesive blanket layer	[m]	log	nom	V = 0,3	200 m	0	-	1
3002	$D_1$	thickness sand layer 1	[m]	log	nom	V = 0,1	200 m	0	-	1
3003	$D_2$	thickness sand layer 2	[m]	log	nom	V = 0,1	200 m	0	-	1
3004	$L$	Length of the seepage path	[m]	log	nom	V = 0,1	3000 m	0	-	1

3005	$\theta$	bedding angle	[°]	log	43°	$\sigma = 3^\circ$	600 m	0	-	1
3006	$C_{\text{Bear}}$	Factor $C_{\text{Bear}}$	[-]	log	nom	$V = 0,15$	450 m	0	-	1
3007	$d_{70}$		[m]	log	nom	$V = 0,15$	180 m	0	-	1
3008	$d_{70}/d_{10}$	Uniformity	[-]	log	nom	$V = 0,15$	750 m	0	-	1
3009	$\eta$	Constant of White	[-]	log	0,3	$V = 0,15$	-	1	-	1
3010	$\gamma_{\text{nat}}$	Wet volumetric weight cohesive cover layer	[kN/m <sup>3</sup> ]	nor	nom	$V = 0,05$	300 m	0	-	1
3011	$\gamma_k$	Volumetric weight sandgrainmaterial	[kN/m <sup>3</sup> ]	nor	27	$V = 0,01$	300 m	0	-	1
3012	$m_o$	Modelfactor uplift	[-]	log	1	$V = 0,1$	-	1	-	1
3013	$m_s$	Modelfactor Sellmeijer	[-]	log	1	$V = 0,08$	-	1	-	1
3014	$m_h$	Damping factor	[-]	log	nom	$V = 0,1$	-	1	-	1
3015	$k_{z,b}$	Permeability upper sand layer 1	[m/s]	log	nom	$V = 1$	600 m	0	-	1
3016	$k_{z,o}$	Permeability lower sand layer 2	[m/s]	log	nom	$V = 1$	600 m	0	-	1
3017	$k_{z,a}$	Permeability river side sand layer 3	[m/s]	log	nom	$V = 1$	600 m	0	-	1
3018	$C_{\text{creep}}$	Creep factor Bligh	[-]	log	nom	$V = 0,1$	-	1	-	1
3019	$m_B$	Modelfactor Bligh	[-]	log	1,5	$V = 0,15$	-	1	-	1
3020	$m_M$	Modelfactor Mpiping	[-]	log	1	$V = 0,08$	-	1	-	1
3021	$B$	Width areaprone to piping (see NB2)	[m]	det	nom	-	-	-	-	-
3022	$P_{\text{sc}}$	Probability pipingscenario	[-]	det	nom	-	-	-	-	-
99007	$h_{\text{kp}}$	inside water level	[m+NAP]	nor	nom	$\sigma = 0,1$ m	vak	1	12 hours	0
99008	$\gamma_w$	Volumetric weight water	[kN/m <sup>3</sup> ]	det	10	-	-	-	-	-

NB 1: For the lognormal distribution, the mean and standard deviation of the real lognormal distribution are given and not of the underlying normal distribution.

NB 2: The width of piping-prone area  $B$  is only used to determine the length-effect.

#### 5.5.4 Subsoil scenarios

Hydra-Ring uses subsoil scenarios to incorporate the uncertainty regarding the composition/layering of the subsoil. Usually, 10 scenarios are defined for each 250 m dike section, based on regional geological data and local measurement; 5 scenarios with blanket layer and 5 scenarios without blanket layer. Each scenario is assigned a probability of occurrence. For each scenario, the probability of piping is calculated. The probability of failure of a dike section is calculated by summing the products of piping probability of scenario probability.

## 6 Statistical distribution functions

N.B This chapter is planned to be move to an appendix

### 6.1 Introduction

This chapter describes the supported distributions and correlation models for load and strength variables. Strength and load variables can follow a number of probability distributions. Table 6.1 gives an overview of the probability distributions supported in Hydra-Ring.

Table 6.1 Probability distributions of load variables, supported in Hydra-Ring.

Probability distributions
Uniform
Normal
(Shifted) lognormal
(Shifted) exponential
Gumbel
Weibull
Rayleigh
Pareto
Triangular
Multi-linear interpolation
Modified Gumbel
Conditional Weibull

Each of these distributions and the required input parameters will be described in the following subsections.

### 6.2 General method to avoid influence of rounding errors

As described in section 3.3.1, statistics are generally described with inverse distribution functions:  $x = F^{-1}(p)$ . For extremely high values of a variable  $x$ , the probability of non-exceedance,  $p$ , will be close to 1. It may occur that  $p$  is so close to 1, that the computer program rounds it off to 1. This will lead to an inaccurate result of  $F^{-1}(p)$  and in some cases even an error because the function is not defined for  $p=1$ . To prevent this from happening,  $p$  is replaced in those cases with  $\exp(-q)$ , where  $q=1-p$ . The functional description of  $F^{-1}(p)$  is thus replaced (only for  $p$  close to 1!) by  $F^{-1}(\exp(-q))$ . The value of  $q$  is close to 0, since  $p$  is close to 1. Values close to 0 can be represented with much higher accuracy than values close to 1. Therefore,  $F^{-1}(\exp(-q))$  will give a higher accuracy than  $F^{-1}(p)$ .

### 6.3 Uniform distribution

The probability density function for the uniform distribution is given by:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \\ 0 & \text{for } x < a \text{ or } x > b \end{cases} \quad (4.72)$$

The corresponding cumulative distribution function is given by:

$$F(x) = \begin{cases} 0 & \text{for } x < a \\ \frac{x-a}{b-a} & \text{for } a \leq x < b, \\ 1 & \text{for } x \geq b \end{cases} \quad (4.73)$$

The inverse of the uniform distribution is given by:

$$F^{-1}(p) = a + p(b-a), \quad p \in (0,1), \quad (4.74)$$

The inverse,  $F^{-1}$ , gives the value of  $x$  associated for a given value of the probability of non-exceedance,  $p$ .

The distribution parameters  $a$  and  $b$  indicate the range over which the probability density function is non-zero, with  $a$  indicating the starting point and  $b$  indicating the ending point. Figure 6.1 and Figure 6.2 show the uniform probability density and cumulative probability distribution, respectively, as a function of  $a$  and  $b$ .

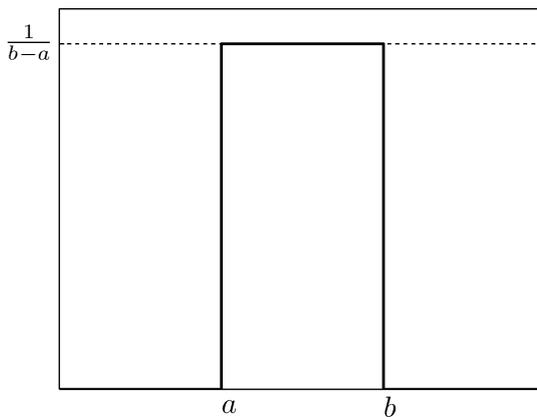


Figure 6.1 Uniform probability density function, with parameters  $a$  and  $b$  indicated

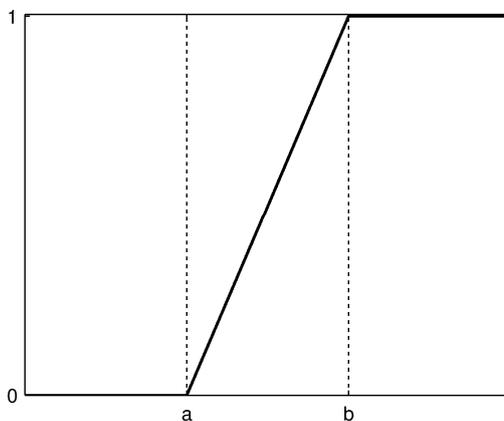


Figure 6.2 Uniform cumulative distribution function, with parameters  $a$  and  $b$  indicated

## 6.4 Normal distribution

The probability density function for the normal distribution is given by:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \quad (4.75)$$

The corresponding cumulative distribution function is given by:

$$F(x) = \frac{1}{2} \left[ 1 + \operatorname{erf}\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right) \right]. \quad (4.76)$$

where erf refers to the error function, which is expressed as follows:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (4.77)$$

The standard normal distribution, denoted by  $\Phi$ , is the special case of the normal distribution for which the mean is equal to zero and the standard deviation is equal to one. For the special case of the standard normal distribution, the inverse is given as follows:

$$\Phi^{-1}(p) = \sqrt{2} \cdot \operatorname{erf}^{-1}(2p-1), \quad p \in (0,1). \quad (4.78)$$

For the general case of a normal distribution with mean  $\mu$  and standard deviation  $\sigma$ , the inverse is given as follows:

$$F^{-1}(p; \mu, \sigma^2) = \mu + \sigma \cdot \Phi^{-1}(p), \quad p \in (0,1). \quad (4.79)$$

There is no explicit form for the normal distribution or inverse normal distribution; they need to be approximated numerically. The method employed in Hydra-Ring is the one described in the Handbook of Mathematical Functions (Abramowitz & Stegun, YEAR, page 392).

Figure 6.3 shows the probability density of the normal distribution, with the parameters  $\mu$  and  $\sigma$  indicated. Figure 6.4 shows the variation in the density function for different choices of  $\mu$  and  $\sigma$ .

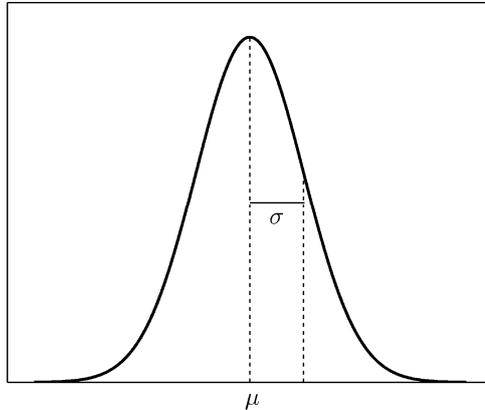


Figure 6.3 Normal probability density function, with parameters  $\mu$  and  $\sigma$  indicated

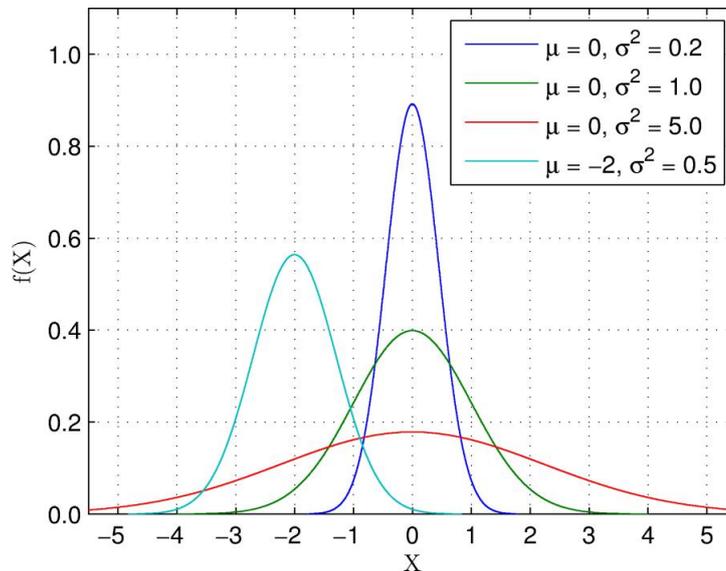


Figure 6.4 Illustration of the effect of parameters  $\mu$  and  $\sigma$

### 6.5 (Shifted) Lognormal distribution

The lognormal distribution is a probability distribution of a random variable whose logarithm is normally distributed. That is, if  $x$  is a random variable with a lognormal distribution, then  $Y = \log(x)$  is normally distributed, and similarly if  $y$  is a random variable with a normal distribution,  $\exp(y)$  is lognormally distributed.

The lognormal density function is given as follows:

$$f(x; \mu, \sigma, \varepsilon) = \frac{1}{(x - \varepsilon)\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln(x - \varepsilon) - \mu)^2}{2\sigma^2}\right], \quad x > 0. \quad (4.80)$$

The corresponding cumulative distribution function is given by:

$$F(x) = \Phi\left(\frac{\ln(x - \varepsilon) - \mu}{\sigma}\right). \tag{4.81}$$

where  $\Phi$  is the standard normal distribution function described in section 6.4.

The parameters of the lognormal distribution,  $\mu$  and  $\sigma$ , are the mean and standard deviation of the associated normally distributed variable. That is, if  $x$  is lognormally distributed, and  $y = \log(x)$  is normally distributed, the lognormal parameters  $\mu$  and  $\sigma$  are the mean and standard deviation of  $y$ . The parameter  $\varepsilon$  is the shift parameter and serves to horizontally translate the distribution. Figure 6.5 shows the effect of different choices of the parameters.

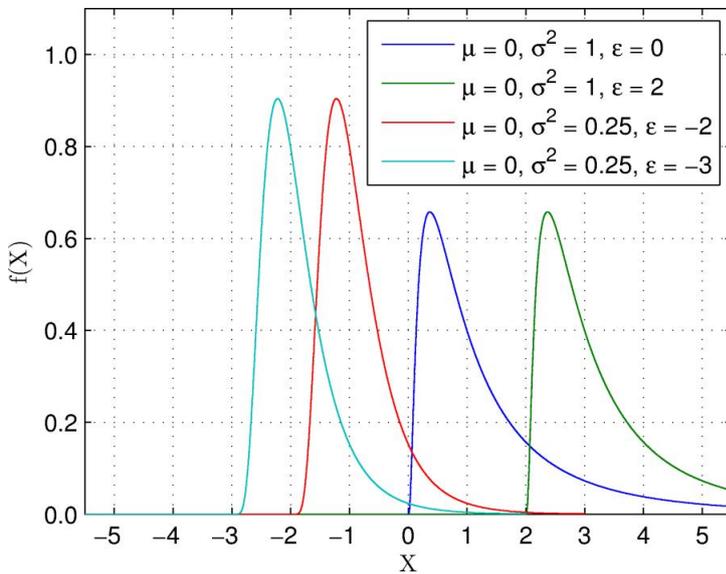


Figure 6.5 Illustration of the effect of parameters  $\mu$ ,  $\sigma$ , and  $\varepsilon$  on the lognormal density function

To get the inverse distribution, the following steps are taken. The first step is, for a given probability  $p$ , the inverse standard normal variable is computed by numerically solving equation (4.78). The associated normal variable with mean  $\mu$  and standard deviation  $\sigma$  is then computed analytically using equation (4.79). The relationship between the normal and lognormal distributions is then exploited:

$$F^{-1}(p; \mu, \sigma) = \exp[\mu + \sigma \cdot \Phi^{-1}(p)] + \varepsilon, \quad p \in (0, 1) \tag{4.82}$$

## 6.6 (Shifted) Exponential distribution

The density function for the shifted exponential distribution is given as follows:

$$f(x; \lambda, \varepsilon) = \begin{cases} \lambda \exp[-\lambda \cdot (x - \varepsilon)], & x > \varepsilon \\ 0, & x < \varepsilon \end{cases}. \tag{4.83}$$

The corresponding cumulative distribution function is given by:

$$F(x; \lambda, \varepsilon) = \begin{cases} 1 - \exp[-\lambda \cdot (x - \varepsilon)], & x \geq \varepsilon \\ 0, & x < \varepsilon \end{cases}. \quad (4.84)$$

The inverse distribution can be analytically computed from the distribution function:

$$F^{-1}(p; \lambda, \varepsilon) = \varepsilon - \frac{\ln(1-p)}{\lambda}. \quad (4.85)$$

The parameter of the exponential distribution,  $\lambda$ , is referred to as a rate parameter, and determines how quickly the density function goes to zero. The parameter  $\varepsilon$  is the shift parameter and serves to horizontally translate the density. Figure 6.5 shows the effect of different choices of the parameters.

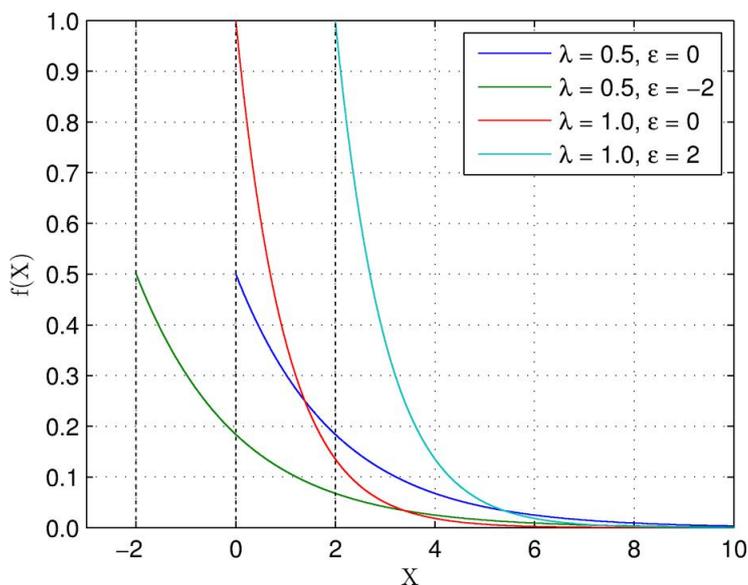


Figure 6.6 Illustration of the effect of parameters  $\lambda$  and  $\varepsilon$  on the exponential density function

## 6.7 Gumbel distribution

The Gumbel distribution is one of three cases of the generalized extreme value distribution. It is often used to model the distribution of block maxima.

The probability density function is expressed as follows:

$$f(x; \alpha, \xi) = \frac{1}{\alpha} \exp\left[-\frac{1}{\alpha}(x - \xi)\right] \cdot \exp\left[-\exp\left[-\frac{1}{\alpha}(x - \xi)\right]\right]. \quad (4.86)$$

The corresponding cumulative distribution function is given by:

$$F(x; \alpha, \xi) = \exp \left[ -\exp \left[ -\frac{1}{\alpha}(x - \xi) \right] \right] \quad (4.87)$$

The inverse distribution can be analytically computed from the distribution function:

$$F^{-1}(p; \alpha, \xi) = \xi - \alpha \ln(\ln(p)) \quad (4.88)$$

The parameters of the gumbel distribution,  $\alpha$  and  $\xi$ , are referred to as the scale and location parameters, respectively. These two parameters can be written in terms of the mean ( $E(x)$ ) and standard deviation ( $\sigma(x)$ ), or in the case of a sample selection, of the sample. The mean and standard deviation of the gumbel distribution are given as follows:

$$E(x) = \xi + \gamma\alpha \quad (4.89)$$

$$\sigma(x) = \frac{\alpha\pi}{\sqrt{6}} \quad (4.90)$$

Equations (4.89) and (4.90) can be used to solve for the parameters  $\alpha$  and  $\xi$ . First,  $\alpha$  is solved in terms of the standard deviation (equation (4.91)), and subsequently, equations (4.89) and (4.91) are used in combination to solve  $\xi$  in terms of the mean and standard deviation.

$$\alpha = \frac{\sigma(x) \cdot \sqrt{6}}{\pi} \quad (4.91)$$

$$\xi = E(x) - \gamma \cdot \frac{\sigma(x) \cdot \sqrt{6}}{\pi} \quad (4.92)$$

Hydra-Ring supports two types of input: either the set of parameters, or the mean and standard deviation. An important note is that Hydra-Ring works with the rate parameter instead of the scale parameter. The rate parameter is equal to the reciprocal of the scale parameter; that is:

$$\text{rate parameter} = \frac{1}{\text{scale parameter}} \quad (4.93)$$

The two parameters that should be input into Hydra-Ring are therefore the rate parameter and the location parameter. Figure 6.7 shows the effect of different choices of the parameters.

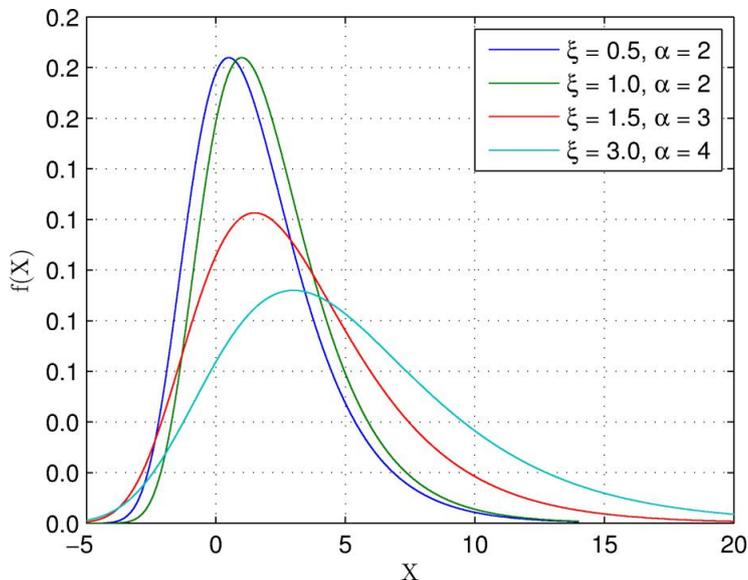


Figure 6.7 Effect of the scale and location parameters,  $\alpha$  and  $\xi$ , on the Gumbel density function

#### 6.8 Weibull distribution

To be filled in

#### 6.9 Rayleigh distribution

To be filled in

#### 6.10 Pareto distribution

To be filled in

#### 6.11 Triangular distribution

To be filled in

#### 6.12 Linear and log-linear interpolation

This section describes the method of interpolation to determine the variable associated with a given non-exceedance probability, using log-linear interpolation. But first we recall (ordinary) linear interpolation. The input to such function is a table with two columns: the first gives probabilities ( $p$  values) and the second gives the values of the variable ( $x$  values) associated with the probabilities in the first column. This function can be useful when the distribution of the variable is best described by a piece-wise linear function, or in the case that the distribution that describes the variable is not included in the library of distribution functions in Hydra-Ring. Essentially any distribution function can be handled using interpolation given sufficient entries in the input table.

The idea of interpolation is that given known values at discrete points, values at positions between those points can be estimated. There are a number of techniques to accomplish that. Hydra-Ring offers the function of log-linear interpolation.

Log-linear interpolation is essentially the same as linear interpolation, only the interpolation takes place over the logarithm of the probability. Linear interpolation is a simple method of estimating values at positions in between known support points. Support points are  $p$ - $x$  pairs in the input table, where  $p$  is the probability and  $x$  is the value associated with that probability.

The points are simply joined by straight line segments in the  $\log(p)$ - $x$  space. The first step in linear interpolation is, for a given input value, to determine the two bounding points  $(\log(p)_1, x_1)$  and  $(\log(p)_2, x_2)$ . Once that is determined, the value of  $x$  can be interpolated between the two bounding points as follows:

$$x_{interpol} = x_1 \cdot (1 - \alpha(p)) + x_2 \cdot \alpha(p) \tag{4.94}$$

where  $\alpha$ , a value bounded between 0 and 1, represents the position of the value  $\log(p)$  between  $\log(p)_1$  and  $\log(p)_2$ :

$$\alpha(p) = \frac{\log(p) - \log(p)_1}{\log(p)_2 - \log(p)_1} \tag{4.95}$$

Figure 6.8 illustrates the concept of log-linear interpolation between two support points  $(\log(p)_1, x_1)$  and  $(\log(p)_2, x_2)$ .

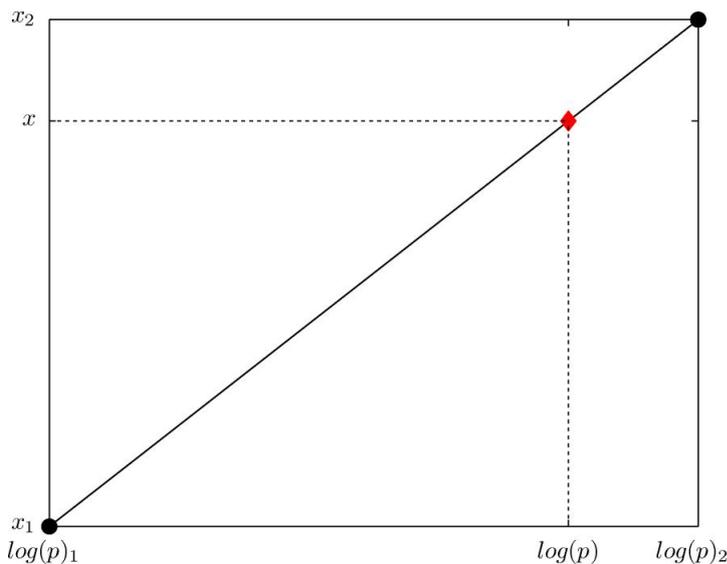


Figure 6.8 Concept of linear interpolation between support points. The black solid circles represent the support points, and the red diamond shows the interpolated value  $(\log(p), x)$

### 6.13 Conditional Weibull distribution

The conditional Weibull distribution gives the probability that  $X \leq x$  conditionally on the event that  $X > \omega$ , where  $\omega$  represents a threshold. It is used in when considering a peaks-over-threshold (POT) method. The distribution function of the conditional Weibull is given as follows:

$$P(X \leq x | X > \omega) = 1 - \exp\left[-\left(\frac{\omega}{\sigma}\right)^\xi - \left(\frac{x}{\sigma}\right)^\xi\right] \tag{4.96}$$

where the parameters  $\omega$ ,  $\sigma$ , and  $\xi$  refer to the threshold, scale, and shape parameter, respectively. The conditional Weibull distribution is often described in terms of exceedance frequencies rather than probabilities. The exceedance frequency of  $x$  can be described as follows:

$$Fr(X > x) = \lambda \cdot P(X > x | X > \omega) \quad (4.97)$$

where  $Fr$  refers to 'frequency', and  $\lambda$  is the frequency with which the selected threshold  $\omega$  is exceeded:

$$\lambda = Fr(X > \omega) \quad (4.98)$$

In practice,  $\lambda$  is determined by counting the number of independent peaks above the threshold and dividing by the number of years of record.

Expanding equation (4.97) so that the probability is full written out gives the following form of the exceedance frequency distribution for the condition Weibull:

$$Fr(X > x) = \lambda \exp\left[\left(\frac{\omega}{\sigma}\right)^\xi - \left(\frac{x}{\sigma}\right)^\xi\right] \quad (4.99)$$

#### 6.14 Modified Gumbel

The modified Gumbel function is applied in the Dutch load model of Hydra-Ring to describe wind statistics (see section 4.5.1.4). The form of the modified Gumbel is identical to the Gumbel distribution (see equation (4.87)), only the argument:

$$\frac{1}{\alpha}(x - \xi)$$

is replaced with a 2<sup>nd</sup>-degree polynomial. The function for the modified Gumbel is given by:

$$F(x; K_r) = \exp\left[-\exp\left[-K_r(x; a, b, c)\right]\right] \quad (4.100)$$

The polynomial  $K_r$  is given as:

$$K_r(x; a, b, c) = ax^2 + bx + c \quad (4.101)$$

The parameters of the modified Gumbel are the coefficients of  $K_r$ :  $a$ ,  $b$ , and  $c$ .

The modified Gumbel is not a probability distribution; equation (4.100) does not necessarily span the range from 0 to 1, and the area under the density does not sum to 1. Rather it is a function that approximates the distribution of the wind speed. The parameters of the function ( $a$ ,  $b$ , and  $c$ ) were chosen to ensure this agreement. However, the agreement is only valid for particular ranges of the wind speed. Figure 6.9 and Figure 6.10 show the modified Gumbel for the wind station Deelen, shown over two different ranges of wind speeds. The former shows the functions over the range 0 to 40 m/s, and the probability density and distribution function appear to be reasonable approximations. The latter shows that for negative wind speeds, rather than the density and the non-exceedance probability being zero, they display peculiar behavior. This illustrates that these functions are not distributions/density functions, and

should therefore be used with care, such that only wind speeds in the valid range are considered.

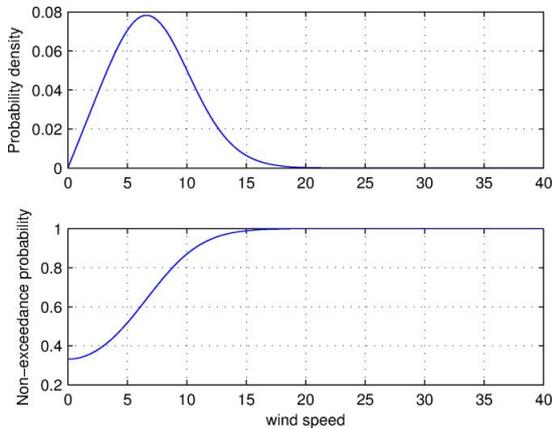


Figure 6.9 Modified Gumbel for wind station Deelen, shown over the range of wind speeds 0 to 40 m/s

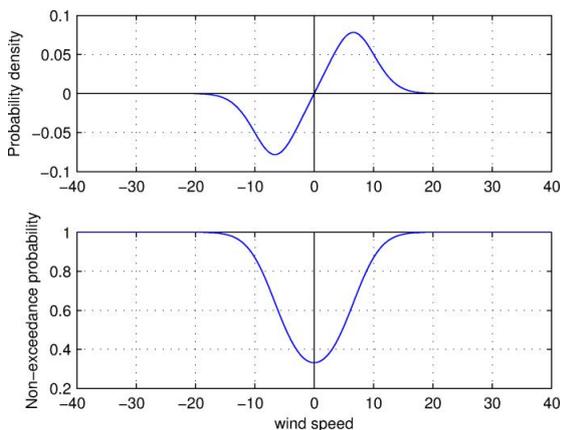


Figure 6.10 Modified Gumbel for wind station Deelen, shown over the range of wind speeds -40 to 40 m/s

To solve for the inverse of the modified Gumbel, equation (4.100) is rearranged:

$$K_r = \ln[-\ln(p)] \tag{4.102}$$

where  $p$  is the non-exceedance probability. Substituting equation (4.101) for  $K_r$  gives the following:

$$ax^2 + bx + c = \ln[-\ln(p)] \rightarrow ax^2 + bx + c' = 0 \tag{4.103}$$

where  $c'$  is given as:

$$c' = c + \ln[-\ln(p)] \tag{4.104}$$

The inverse of the modified Gumbel can then be solved using the quadratic equation:

$$F^{-1}(p; a, b, c) = \begin{cases} \frac{-b + \sqrt{b^2 - 4ac'}}{2a}, & (b^2 - 4ac') > 0 \\ \frac{-b}{2a}, & (b^2 - 4ac') < 0 \end{cases} \quad (4.105)$$

Note that this solution is pragmatic in the sense that only one half of the quadratic solution is used, based on the knowledge that wind speeds are not negative (note that typically the quadratic solution is  $-b \pm \sqrt{b^2 - 4ac}$ , and here we ignore the minus). Also, in cases where the square root term is negative, there are no real solutions; what is done in this case is to take the value at the peak (or trough) of the parabola, as this point will be the closest to the  $x = 0$  axis. To illustrate this, Figure 6.11 shows the quadratic equation given by formula (4.103) for six different values of the non-exceedance probability  $p$ . In Figure 6.10 it was shown that for non-exceedance probabilities less than about 0.35, there is no solution to the inverse. It is for these cases that the square-root term in formula (4.105) is negative, and the second term  $(-b/2a)$  is used as an approximate solution. This is evident in the first two subplots of Figure 6.11, where the parabola does not cross the  $x = 0$  line.

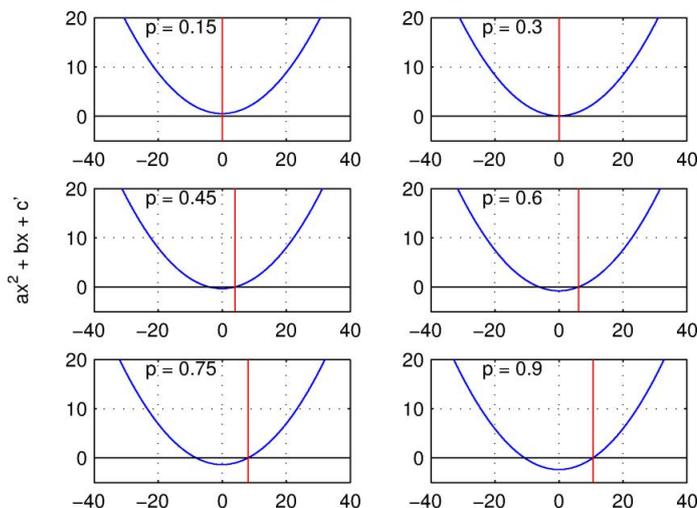


Figure 6.11 Quadratic equation (equation (4.103)) for wind station Deelen, shown for six different non-exceedance probabilities (indicated in figure). The red vertical line indicates the solution to the inverse (equation (4.105))

### 6.15 Truncated Gumbel

The truncated Gumbel distribution is used in the correlation model Volker, and is therefore described in this section. Recall the form of the Gumbel distribution function (equation (4.87)), shown below for the case of a location parameter equal to zero and a scale parameter equal to 1:

$$F(x) = \exp[-\exp(-x)] \tag{4.106}$$

A truncated Gumbel distribution has the following form:

$$F(x;d) = \begin{cases} \frac{1}{1-d} \exp[-\exp(-x)], & x \leq x_d \\ 1 & x > x_d \end{cases} \tag{4.107}$$

Setting the non-exceedance probability to 1 for values of  $x$  greater than  $x_d$  essentially truncates the probability that values higher than  $x_d$  will be observed. That is, the probability density of  $x$  values higher than  $x_d$  will be equal to zero. The ratio  $1/(1 - d)$  normalizes the distribution and that ensures that the total probability is equal to 1. The value of  $d$  can be thought of as the fraction of the distribution that will be truncated. That is, if  $d = 0.02$ , the highest 2% of the distribution will be truncated.

The endpoint of the distribution,  $x_d$ , can be derived from equation (4.107) and is given as follows:

$$x_d = -\ln[-\ln(1-d)] \tag{4.108}$$

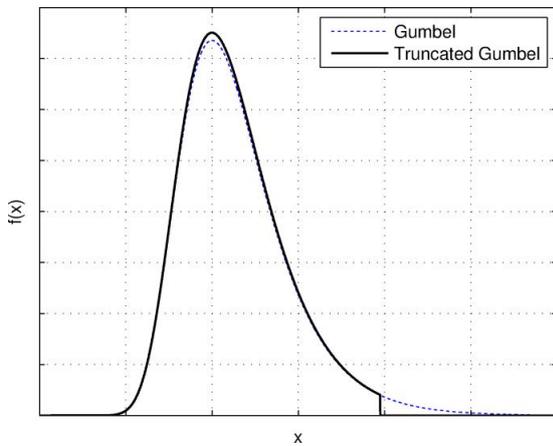


Figure 6.12 Illustration of the effect of including the truncation factor  $d > 0$  in equation(4.107), here shown for  $d = 0.02$

## 7 References

- Berger, H.E.J., 2007: Achtergrondrapport HR2006 voor de zoete wateren. Het samenstellen van de Hydraulische Randvoorwaarden 2006 en de Thermometerrandvoorwaarden 2006. RWS RIZA rapport 2007.026 (In Dutch).
- Bjerager P, 1988. On computation methods for structural reliability analysis. In: Frangopol DM, editor. New directions, in structural system reliability. Boulder (CO): University of Colorado, 1988. p. 52±67.
- Den Heijer, F., R.J. Vos, F.L.M. Diermanse, J. Groeneweg, R.Tonis, 2007: Achtergrondrapport HR2006 voor de Zee en Estuaria, Hydraulische Randvoorwaarden 2006. Rapport RIKZ/2006.029., 30 mei 2007 (In Dutch)..
- De Waal, J.P. 1999: Achtergronden belastingen dijken IJsselmeergebied, deelrapport 9, modelleren dammen, voorlanden en golfoploop, RIZA rapport, maart 1999.
- Diermanse, 2004: HR2006 herberekening werklijn Maas, rapport WL | Delft Hydraulics in opdracht van Rijkswaterstaat.
- Diermanse, F., Thonus, B. Lammers, I. en Den Heijer, F., 2003: De veiligheid van Nederland in kaart – hydraulische randvoorwaarden. Fase 1 – inventarisatie en gegevensverzameling, WL | Delft Hydraulics en HKV | Lijn in water in opdracht van Rijkswaterstaat (in Dutch).
- Diermanse, F en Geerse, C, 2011: Correlatiemodellen in Hydra-Ring. Memo juni 2011, Deltares en HKV in opdracht van Rijkswaterstaat.
- Diermanse, F.L.M. and Geerse, C.P.M , 2012: Correlation models in flood risk analysis, Reliability engineering and system safety, 105 (2012) 64–72
- Ditlevsen O, Melchers RE, Gluwer H, 1990. General multi-dimensional probability integration by directional simulation, Computers And Structures 1990;36(2):355±68.
- Ferry Borges, J. & M. Castanheta (1971). Structural Safety, course 101, 2nd edition. Lisbon: Laboratorio Nacional de Engenharia Civil.
- Geerse, 2006: Hydraulische Randvoorwaarden 2006 Vecht- en IJsseldelta - Statistiek IJsselmeerpeil, afvoeren en stormverlopen voor Hydra-VIJ. C.P.M. Geerse. RIZA-werkdocument 2006.036x. Rijkswaterstaat-RIZA. Lelystad, januari 2006.
- Geerse, 2008. Hydra-VIJ invoer Markermeer (herziene en uitgebreide versie van 9 oktber). C.P.M. Geerse. Memo PR1371.30. HKV Lijn in Water, Lelystad, 28 augustus 2008.
- Geerse, 2010. Overzichtsdocument probabilistische modellen zoete wateren. Hydra-VIJ, Hydra-B en Hydra-Zoet. C.P.M. Geerse (HKV), met medewerking van Herbert Berger en Robert Slomp (Waterdienst). HKV Lijn in Water, Lelystad, juli 2010. In opdracht van de Waterdienst.
- Geerse, C.P.M, 2011: Hydra-Zoet for the fresh water systems in the Netherlands, probabilistic model for the assessment of dike heights, report HKV for Rijkswaterstaat.
- Geerse, 2011b: Correlatiemodel VSU, Uitgebreidere versie van model VS. Memo, versie 2. Chris Geerse. HKV, 6 juli 2011.
- Geerse en Van Veen, 2007. Statistische afhankelijkheid wind voor de bovenrivieren. Chris Geerse en Nelle van Veen. HKV Lijn in Water, Lelystad, november 2007. (in Dutch)
- Gautier, C. en Groeneweg, J., 2011: Achtergrondrapportage HR2011 voor zee en estuaria, Deltares in opdracht van Rijkswaterstaat, 1204143-002-HYE-0037, 22 maart 2012 (in Dutch)
- Grooteman, F., 2011: An adaptive directional importance sampling method for structural reliability, Probabilistic Engineering Mechanics 26 (2011) 134–141.
- Grimmett, G.R. and Stirzaker, D.R., 1983: Probability and random processes, Oxford science publications, isbn 0-19-853185-0.

- Jongejan, R., 2012: Het lengte-effect in een statistisch homogeen vak (in Dutch), memo februari 2012
- Melchers, 2002: Structural reliability analysis and prediction, Robert E. Melchers, John Wiley and sons, ISBN0471983241
- Ministerie Verkeer en Waterstaat (2006). Voorschrift Toetsen op Veiligheid (in English: Safety Assessment Regulation) Bligh, W.G. (1915). Submerged weirs founded on sand. In: Dams and weirs, 1915.
- RIKZ, 2006. Waterstanden Nederlandse kust en estuaria. Rapport RIKZ/2006.012, 27 juni 2006.
- Rossenblatt, M., 1952: Remarks on a multivariate transformation, Ann. Matht Stat., 23, 470-472
- Steenbergen, H. and A. Vrouwenfelder, User's manual PC-Ring Version 3.0 (in Dutch), 2003.
- Van Gelder, 2002: Probabilistic design in hydraulic engineering, lecture notes Delft University of technology, temporary issue, august 2002,
- Vrouwenfelder, A, Courage, W, Diermanse, F. en Geerse, C., 2011: belastingmodellen en rekentechnieken voor TOI, TNO report.
- Waarts, P. H., 2000: Structural reliability using finite component analysis; an appraisal of DARS: Directional Adaptive Responce Surface Sampling, PH. D. Thesis, delft University of Technology
- TAW, 1994: Technical Advisory committee on Water defences - TAW (1994), Technisch rapport voorcontrole op het mechanisme piping. Delft, by Calle, E., Weijers, J.
- TAW, 2002: Technical Advisory committee on Water defences -. Technical Report on Sand Boils (Piping). Published at [www.enwinfo.nl](http://www.enwinfo.nl)
- Thonus, B., Diermanse, F. en Lammers, I, 2003: De veiligheid van Nederland in kaart – hydraulische randvoorwaarden. Fase 2 – uitwerken en implementeren van belastingmodellen. WL | Delft Hydraulics en HKV | Lijn in water in opdracht van Rijkswaterstaat (in Dutch).
- Volker, 1987: Statistiek van wind en waterstanden in Hoek van Holland DWW, tweede concept, 20 mei 1987.

<C:\Data\Dijken\Hydra-Ring\Trunk\doc\Hydra Ring Glossary.doc>







## A Dutch Hydraulic Boundary Conditions – version TMR2006

### A.1 Overview of stochastic input variables and regions

This annex describes the definition of the different regions with associated stochastic variables and their distributions and correlations, according to the Dutch TMR2006 boundary conditions set.

Hydraulic boundary conditions are typically the water level and the different wave parameters at a certain location. Hydra-Ring determines these values by interpolation in data tables, using the values of certain stochastic variables as input. The values of the water level and wave parameters in the data tables are a result from hydraulic models. Hydra-Ring supports the use of different data tables for different regions. A region is in this context defined as an area with a unique set of stochastic input parameters, with associated distributions and correlations.

The Dutch TMR2006 boundary conditions set consists of sixteen regions. Table 7.1 presents these regions.

Table 7.1 - Overview of regions in the TMR2006 database

Region number	Region Description
1	Upstream river region (no tidal influence) - Rhine dominated
2	Upstream river region (no tidal influence) - Meuse dominated
3	Downstream river region (tidal influence) – Rhine dominated
4	Downstream river region (tidal influence) – Meuse dominated
5	IJssel delta
6	Vecht delta
7	IJsselmeer
8	Markermeer
9	Wadden Sea east
10	Wadden Sea west
11	Dutch coast north
12	Dutch coast middle
13	Dutch coast south
14	Oosterschelde
15	Westerschelde

Table 7.2 lists (in the vertical) the independent load variables that are present in Hydra-Ring and in the horizontal the sixteen regions. Every dike location in the Netherlands belongs to one region. For each region, Table 7.2 can be used to determine which independent load variables are active there (a cross (x) indicates the variable is active in the given region). The value of the parameters of the load variables can be obtained from the TMR2006 table at the end of this appendix (Table 7.3).

Table 7.2- Inclusion of load variables at the different regions in the Netherlands; the sixteen regions are shown on the horizontal axis and the load variables are shown on the vertical; an 'x' indicates that the variable is active in the given region

Load variables Name	Region															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Discharge Lobith	x	x	x	x												
Discharge Lith   Lobith		x		x												
Discharge Olst					x											
Discharge Borgharen		x		x												
Discharge Dalfsen   Olst						x										
Water level Maasmond			x	x												
Water level IJssel lake   windspeed Schiphol12					x	x	x									
Water level Marker lake   windspeed Schiphol12								x								
Water level Lauwersoog									x							
Water level Harlingen										x						
Water level Den Helder											x					
Water level IJmuiden												x				
Water level Hoek van Holland													x			
Water level OS11														x		
Water level Vlissingen															x	
Wind speed Deelen	x	x														
Wind speed Schiphol 12**							x	x								
Wind speed Schiphol 16**					x	x										
Wind speed Schiphol   MM**			x	x												
Wind speed Terschelling West   Water level									x							



Variable	Probability distribution and peak duration		Duration curve (FBC model)		Correlation	
	Type	Data sources	Type	Data sources	Type	Data sources
Discharge Lobith	Interpolation (loglinear for distribution and linear for peak duration)	VNK_2003_belastingmodellen_fase1 [App. 3A,3N] TMR2006/WWdatabasis/WenWtzV HydraRing\Trunk\data\Afvoerstatistiek Hydra-Zoet	polynomial	PC-Ring manual-Statistical models [Sec. 4.2.2] TMR2006/WWdatabasis/WenWtzV	See Lith	
Discharge Lith   Lobith	Interpolation (loglinear for distribution and linear for peak duration)	VNK_2003_belastingmodellen_fase1 [App. 3B] TMR2006/WWdatabasis/WenWtzV HydraRing\Trunk\data\Afvoerstatistiek Hydra-Zoet	polynomial	PC-Ring manual-Statistical models [Sec. 4.2.3] TMR2006/WWdatabasis/WenWtzV	PCR/NL	Wind/waterlevel VNK_2003_belastingmodellen_fase1 [sec. 6.2.3] PC-Ring manual-Statistical models [Sec. 4.2.6]
Discharge Olst	Interpolation (loglinear for distribution and linear for peak duration)	VNK_2003_belastingmodellen_fase1 [App. 4A] HydraRing\Trunk\data\Afvoerstatistiek Hydra-Zoet	polynomial	PC-Ring manual-Statistical models [Sec. 4.2.4]		
Discharge Dalfsen	Interpolation (loglinear for distribution and linear for peak duration)	VNK_2003_belastingmodellen_fase1 [App. 4A] HydraRing\Trunk\data\Afvoerstatistiek Hydra-Zoet	polynomial	PC-Ring manual-Statistical models [Sec. 4.2.6.2]		
Water level Maasmond	Cond. Weibull	VNK_2003_belastingmodellen_fase2 [Section 2.2.4]			----	----
Water level IJssel lake	Interpolation (loglinear for	VNK_2003_belastingmodellen_fase1 [App. 5A,5B]	exponential	PC-Ring manual-Statistical models [Sec. 4.3.2]		VNK_2003_belastingmodellen_fase1 [sec. 4.2.3]

	distribution and linear for peak duration)	TMR2006/WWdatabasis/WenWtzV HydraRing\Trunk\data\Afvoerstatistiek Hydra-Zoet				
Water level Marker lake   wind speed Schiphol 12	Interpolation (loglinear for distribution and linear for peak duration)	VNK_2003_belastingmodellen_fase1 [App. 5C,5D] TMR2006/WWdatabasis/WenWtzV HydraRing\Trunk\data\Afvoerstatistiek Hydra-Zoet	exponential	PC-Ring manual-Statistical models [Sec. 4.3.3]	Volker	VNK_2003_belastingmodellen_fase1 [sec. 4.2.3]
Water level Lauwersoog	Cond. Weibull	VNK_2003_belastingmodellen_fase1 [App. 6B] PC-Ring manual-Statistical models [App. C]	-----	-----	-----	-----
Water level Harlingen	Cond. Weibull	VNK_2003_belastingmodellen_fase1 [App. 6B] PC-Ring manual-Statistical models [App. C]	-----	-----	See Wind speed Terschelling West	-----
Water level Den Helder	Cond. Weibull	VNK_2003_belastingmodellen_fase1 [App. 6B] PC-Ring manual-Statistical models [App. C]	-----	-----	See Wind speed de Kooy	-----
Water level IJmuiden	Cond. Weibull	TMR2006/WWdatabasis/WenWtzV	-----	-----	See wind speed IJmuiden	-----
Water level Hoek van Holland	Cond. Weibull				See Wind speed Hoek van Holland	
Water level OS11	Cond. Weibull	VNK_2003_belastingmodellen_fase1 [App. 6B] PC-Ring manual-Statistical models [App. C]	-----	-----	See Wind speed Vlissingen	-----
Water level Vlissingen	Cond. Weibull	VNK_2003_belastingmodellen_fase1 [App. 6B] PC-Ring manual-Statistical models [App. C]	-----	-----	See Wind speed Vlissingen	-----
Wind speed Deelen	Modif. Gumbel	VNK_2003_belastingmodellen_fase1 [App. 3J] PC-Ring manual-Statistical models [App. B] TMR2006/WWdatabasis/WenWtzV	-----	-----	-----	-----

Wind speed Schiphol 12	Modif. Gumbel	TMR2006/WWdatabasis/WenWtzV	----	----	<i>See water level lakes</i>	
Wind speed Schiphol 16	Modif. Gumbel	TMR2006/WWdatabasis/WenWtzV	----	----	----	----
Wind speed Schiphol   MM	----	TMR2006/WWdatabasis/WenWtzV	----	----	Volker	TMR2006/WWdatabasis/WenWtzV
Wind speed Terschelling West   Water level Lauwersoog	----	----	----	----	PCR	Appendix B [B.2.5.2] TMR2006/WWdatabasis/WenWtzV
Wind speed Terschelling West   Water level Harlingen	----	----	----	----	PCR	Appendix B [B.2.5.2] TMR2006/WWdatabasis/WenWtzV
Wind speed de Kooy   Water level Den Helder	----	----	----	----	PCR <i>See Wave period ELD</i>	Appendix B [B.2.5.2] TMR2006/WWdatabasis/WenWtzV
Wind speed IJmuiden   Water level IJmuiden	----	----	----	----	Volker <i>See Wave period YM6</i>	TMR2006/WWdatabasis/WenWtzV
Wind speed Hoek van Holland   Water level Hoek van Holland	----	----	----	----	Volker <i>See Wave period EUR</i>	TMR2006/WWdatabasis/WenWtzV
Wind speed Vlissingen   Water level OS 11	----	----	----	----	PCR	Appendix B [B.2.5.2] TMR2006/WWdatabasis/WenWtzV
Wind speed Vlissingen   Water	----	----	----	----	PCR	Appendix B [B.2.5.2]

level Vlissingen						TMR2006/WWdatabasis/WenWtzV
Wave period ELD   Wind speed de Kooy	----	----	----	----	Volker	TMR2006/WWdatabasis/WenWtzV ?
Wave period YM6   Wind speed IJmuiden	----	----	----	----	Volker	TMR2006/WWdatabasis/WenWtzV ?
Wave period EUR   Wind speed Hoek van Holland	----	----	----	----	Volker	TMR2006/WWdatabasis/WenWtzV ?
Prediction error water level MM	Normal[cr-ref]	PC-Ring manual-Statistical models [Section 4.8]	----	----	----	----
Duration of the wind setup	????	????	----	----	----	----
Phase difference	????	????	----	----	----	----
Water level dunes	Cond. Weibull	Belastingmodel duinen [Section 2.3.2]	----	----	See Wave height Dunes	----
Wave height dunes   water level dunes	----	Belastingmodel duinen [Section 2.3.3]	----	----	NL-Dunes See Peak Period Dunes	Belastingmodel duinen [Section ?]
Peak period dunes   wave height dunes	----	Belastingmodel duinen [Section 2.3.4, App. A]	----	----	NL-Dunes	Belastingmodel duinen [Section ?]
Wind direction	---	TMR2006/WWdatabasis/WenWtzV	----	----		----

The parameters for autocorrelation in space and time can be found in this datasource: TMR2006/belasting