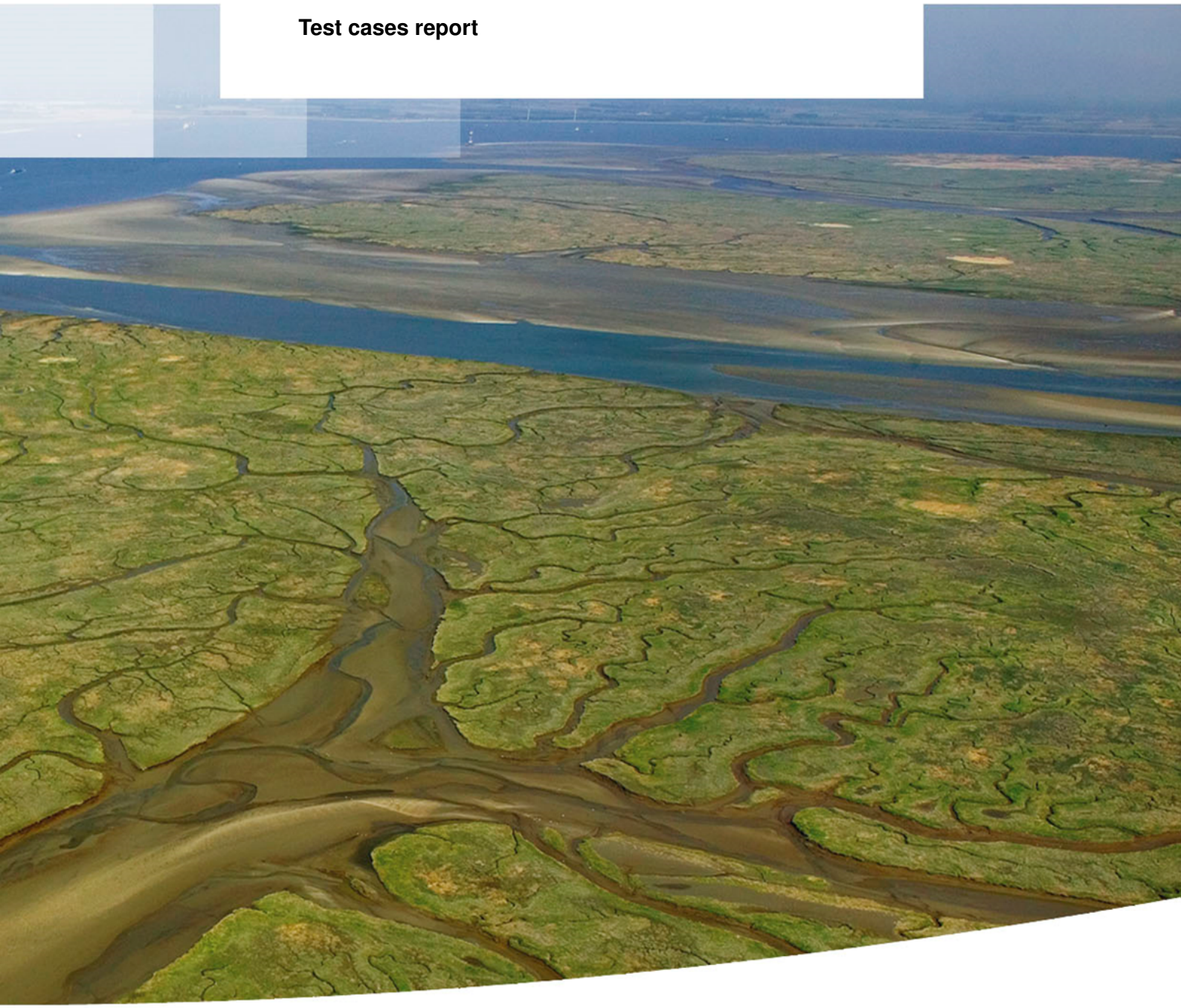
Deltares

Enabling Delta Life

Reliability updating for slope stability of dikes

Test cases report



Deltares

Reliability updating for slope stability of dikes

Test cases report

Dr. ir. T. Schweckendiek
Dr. A. Teixeira
Ir. M.G. van der Krogt
Dr. ir. W. Kanning

with contributions from:
Ir. K. Rippi
Dr. ir. C. Zwanenburg

1230090-037

Deltares

Title
Reliability updating for slope stability of dikes

Client	Project	Reference	Pages
Rijkswaterstaat WVL	1230090-037	1230090-037-GEO-0003	98

Classification
none

Keywords
Bewezen sterkte, dike safety, macrostability, Markermeerdijken, past performance, reliability updating, test cases

Summary
see chapter "Summary and conclusions"

References
see chapter "References"

Version	Date	Author	Initials	Review	Initials	Approval	Initials
01	31 May 2016	T. Schweckendiek		A.P.C. Rozing		M. Sule	
02	26 Aug 2016	T. Schweckendiek		A.P.C. Rozing		M. Sule	
03	28 Nov 2016	T. Schweckendiek		A.P.C. Rozing		M. Sule	

Status
FINAL

Contents

List of Figures

List of Tables

List of symbols

Latin symbols

CoV	coefficient of variation
$g(\cdot)$	performance function
h	water level (load)
h_p	polder water level
h^*	observed water level
H	house load
H*	observed house load
IL	geohydrological intrusion length
m	shear strength increase exponent
m_d	model uncertainty
n	number of MCS-realizations
PL	phreatic level
$P(\cdot)$	probability operator
POP	pre-overburden pressure
p_f	probability of failure
S	shear strength ratio (normally consolidated)
t	time duration
T	traffic load
T*	observed traffic load

Greek symbols

α	influence coefficient or importance factor (FORM)
β	reliability index
γ	volumetric weight
ϕ	friction angle sand
λ	geohydrological leakage length
μ	mean value
ρ	linear correlation coefficient
σ	standard deviation
σ'_y	yield stress

Abbreviations

CDF	cumulative distribution function
CSSM	critical state soil mechanics
DoV	Dijken op Veen (dikes on peat research project)
DS	Directional sampling
FC	fragility curves (approximation method)
FORM	First-order reliability method
MCS	Monte Carlo simulation (crude)
NAP	Normaal Amsterdams Peil (Dutch reference level)
PDF	probability density function
PMF	probability mass function
RUPP	reliability updating with past performance
SF	stability factor
WBI	wettelijk beoordelingsinstrumentarium (Dutch safety assessment framework)

1 Summary and conclusions

The present report provides two case studies testing and illustrating the application of the reliability updating approach described in the accompanying background report by Schweckendiek and Kanning (2016). This summary chapter discusses the main findings and conclusions. The analyses and discussions of results are documented in the remaining chapters.

1.1 Project context

Rijkswaterstaat seeks to operationalize the concept of *Reliability Updating with Past Performance* (RUPP; in Dutch often refered to as *bewezen sterkte*) for advanced safety assessments and reinforcement designs of the primary Dutch flood defenses. Reliability updating means to incorporate past performance observations in estimates of the probability of failure, specifically observations of survived load conditions. The focus in this first phase of the project is on the failure mode of instability of the inner slope, as many dikes were found not to meet the safety criteria for this failure mode in the statutory safety assessment of the Dutch primary flood defenses (Inspectie Verkeer en Waterstaat, 2011). This applies specifically for the Markermeerdijken which have been planned to be reinforced, mainly for inner slope stability. A proof-of-concept study by Schweckendiek and Van der Krogt (2015) suggested that reliability updating with past performance could potentially reduce the envisaged dike reinforcement efforts for the Markermeerdijken. Therefore, the development efforts so far have focused on typical Markermeerdijken conditions and related challenges.

1.2 Objectives

Aim of the long-term development project

The aim of the envisaged development efforts for the long-term project (development time roughly 3 years, see WVL (2016)) is to enable practitioners to use reliability updating in 'advanced safety assessments' (in Dutch: *toets op maat*) and reinforcement designs of the primary Dutch flood defenses. This implies the following sub-objectives:

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

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

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

Notice that the accompanying background report (Schweckendiek and Kanning, 2016) (1) and this test case report (2) are primarily aimed at an expert reader in order to assess the soundness of the approach and the envisaged application, while the manual and software (3 & 4) will address practitioners who are not necessarily experts in reliability analysis.

Objectives of the current phase

The **main objective** of the current development efforts is to operationalize the reliability updating method for advanced safety assessments regarding the failure mode 'slope instability', specifically the method using fragility curves. As for the overall project, this involves both (a) development and documentation of the method, and (b) demonstration of the applicability with realistic test cases.

The **secondary objective** of this study is to generate insights to help estimating the potential effect on the reliability estimates for typical Markermeerdijken conditions.

For the reasons mentioned above, the intermediate phase of the project has focused exclusively on the Markermeerdijken conditions. For a broader acceptance of the reliability updating approach, more case studies from a wider range of conditions need to be investigated.

1.3 Scope

In the current phase of the project, the development efforts are limited to the failure mode of instability of the inner slope. The accompanying background report (Schweckendiek and Kanning, 2016) provides the description of the general reliability updating method, specifically the approach using fragility curves, and its application to dike stability. For the testing and illustration of the applicability, two test cases have been analyzed. In order to address the secondary objective of generating insights for the potential impact in Markermeerdijken conditions, the two cases were taken from the Markermeerdijken project area.

In the Dutch safety assessment framework, there are different assessment levels in terms of complexity and data requirements. The present reliability updating analyses have the character of an advanced assessment (in Dutch: *toets op maat*). The essence of an advanced assessment is that any state-of-the-art models and methods can be used to substantiate the dike safety and in this case the reliability estimate, which from 2017 will be in terms of an annual probability of flooding for a dike segment. That implies that one is not necessarily bound to common conservative assumptions made in standard assessments (in Dutch: *gedetailleerde toets*) or designs.

1.4 Test cases and a-priori reliability

For the reasons discussed, the two tests cases are located in the Markermeerdijken area. One of the cross sections is a regular so-called "green" dike without buildings or special objects in the cross section ('case green'; Figure 1.1); the other cross section does contain a building ('case house'; Figure 1.2). The stratification in both cross sections is rather common for the Markermeerdijken, though in some parts there is less peat (brown) and more clay (grey) in the subsoil profile.

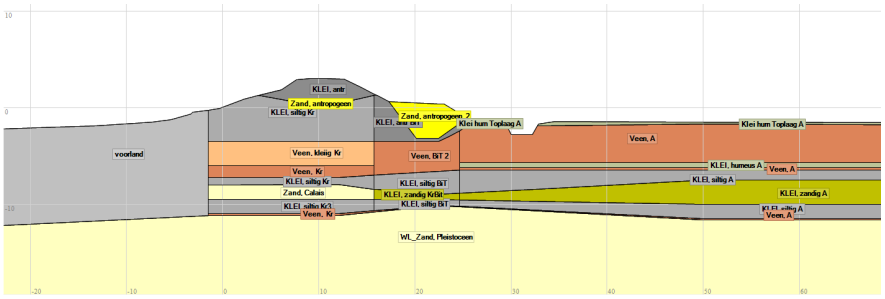


Figure 1.1: *Case green: Geometry and soil layering (the Markermeer is on the left-hand side).*

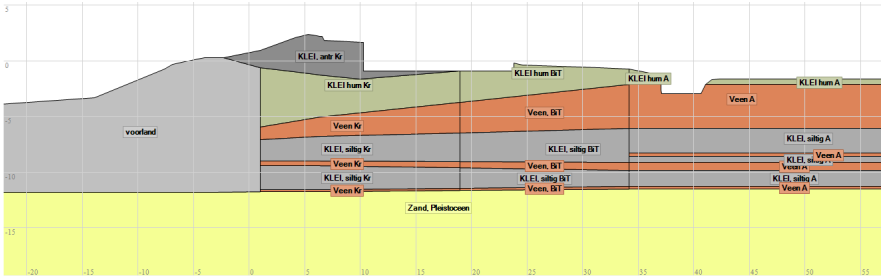


Figure 1.2: *Case house: Geometry and soil layering (the house is itself not depicted).*

For both cases first prior reliability analyses have been carried out (i.e. without past performance information) before incorporating the past performance (survival) information in the so-called 'posterior analysis'. Table 1.1 contains the results of the prior reliability analyses, including the safety factors obtained with mean and design values as defined in WBI-2017, the safety assessment framework which will be in force from 2017. The results illustrate that, although the ground conditions are similar, quite large differences in stability and reliability can be conceived between the cases, depending on the specific geometric profile and the local pore water pressure conditions. The findings confirm earlier experiences that with the low effective stress levels as present in the Markermeerdijken the stability is rather sensitive to apparently minor changes in the cross sections.

Table 1.1: *Summary of the prior results for 'case green' and 'case house'. The prior reliability indices for the considered so-called base cases are 0.91 and 5.62 respectively ("all water levels" refers to the entire probability distribution of the water level).*

CASE GREEN				
Water level h	**Traffic load**	**Factor of Safety SF**		**Reliability index**
[m] + NAP	[kN/m^2]	**mean values**	**design values**	**(annual)**
-0.40	13.3	1.13	0.69	0.95
1.15	13.3	1.11	0.67	0.74
-0.40	0.0	1.18	0.74	1.53
1.15	0.0	1.16	0.72	1.31
all water levels	**13.3**	n/a	n/a	**0.91**
all water levels	0.0	n/a	n/a	1.50

CASE HOUSE				
Water level h	**Traffic load**	**Factor of Safety SF**		**Reliability index**
[m] + NAP	[kN/m^2]	**mean values**	**design values**	**(annual)**
-0.40	13.3	1.83	1.02	5.65
1.15	13.3	1.78	1.00	5.45
-0.40	0.0	1.93	1.08	6.29
1.15	0.0	1.86	1.06	5.84
all water levels	**13.3**	n/a	n/a	**5.62**
all water levels	0.0	n/a	n/a	6.25

We want to emphasize at this stage that, though the soil parameters and other parameters were chosen as close as possible to the values used by the *Hoogheemraadschap Hollands Noorderkwartier (HHNK)* at the beginning of this project, further research by HHNK may have lead to changes in these values in the meantime. Hence, the current results can only be used to obtain an impression of the approximate safety levels and the relative effects of reliability updating. They cannot be directly used to conclude whether the investigated dike sections meet the reliability targets.

Since most practitioners have limited experience yet with the outcomes of a probabilistic analysis in terms of the reliability index, we can place the results in the context of the safety requirements. The target reliability as stated by the current design guidelines (OI, 2015) is roughly 4.6 in the Markermeerdijken area. Moreover, studies in the WBI-2017 project have investigated the relation between the reliability index and factor of safety in order to substantiate the required factors of safety for dike stability (Kanning *et al.*, 2015).

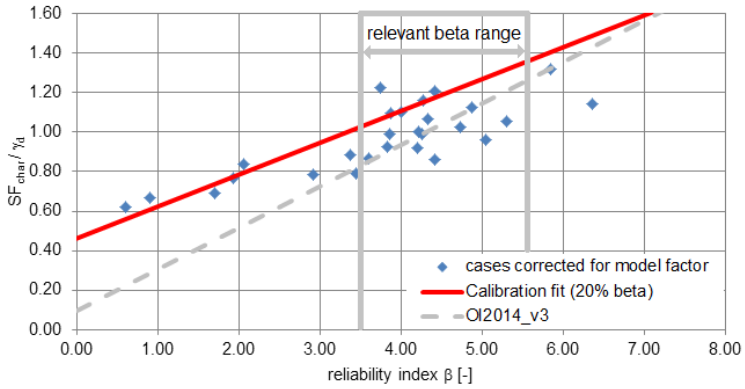


Figure 1.3: *Comparison of safety factors for Dutch dikes (inner slope stability) computed with design values versus the corresponding reliability index according to Kanning* et al. *(2015).*

According to Figure 1.3 safety factors of roughly 1 imply a reliability index of at least 3.5 ranging up to 5.5. The prior reliability index of 5.6 found for 'case house' lies within the upper range of this interval. This comparison illustrates that probabilistic analyses tailored to the local conditions avoid the conservatism implied in semi-probabilistic assessments (i.e. required factors of safety), necessarily introduced to cover all conceivable conditions in the domain of application (i.e. the required safety factors need to ensure sufficient reliability for the entire Netherlands in this case).

Furthermore, the results show that especially 'case green' is very sensitive to the traffic load. In fact, it is the dominant load when considering Figure 1.4; the change of safety factor and reliability index is much more pronounced for variations in the traffic load than for variations in the external water level (lake Marken). That means that the traffic load requires careful treatment also in the reliability updating for 'case green', as we know that the impact of reliability updating largely depends on the observation of significant (dominant) loads.

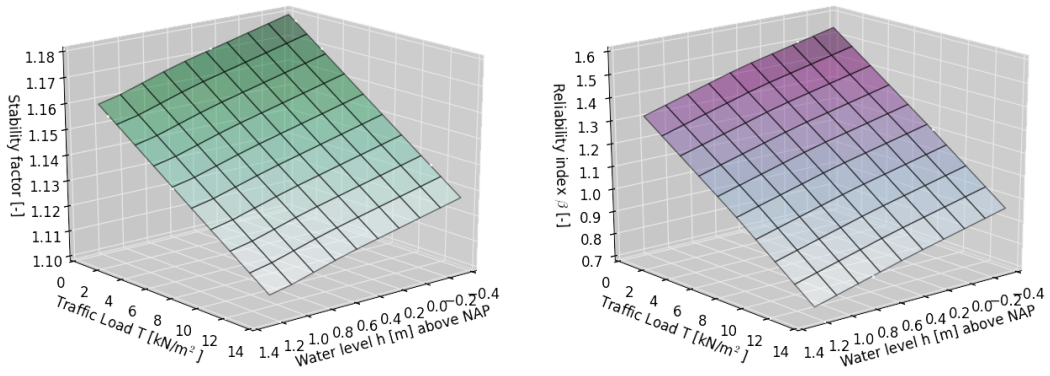


Figure 1.4: *Case green: Illustration of the sensitivity of the safety factor (left) and the reliability index (right) to the water level h and the traffic load T.*

1.5 Reliability updating with past performance (RUPP)

The current study has focused on reliability updating using the information provided by survived load conditions. A survived high external water level is the most commonly observed load that can be used in such analyses, but also observations of extreme rainfall (and the effect on pore water pressures) or other external loads such as traffic loads can provide information with a significant impact on the probability of failure.

For both test cases, we first contemplated a base case with conservative assumptions using the daily conditions as survival observation. In the specific case of the Markermeerdijken, the dike stability is relatively insensitive to the external water level, mainly due to the low permeability of the dike material itself and the relatively thick underlying low-permeability substrata. At the same time, the dike is constantly loaded by the lake and by the constantly high phreatic surface levels inside the dike. Consequently, we find a significant effect even from incorporating survival of the daily loading conditions, which is rather uncommon for other types of dikes such as river dikes. In the daily conditions, there are also relatively limited uncertainties (e.g. phreatic level) compared to historic events with higher water levels. Subsequent to the daily conditions, we also explored the sensitivity to other load observations such as higher water levels, traffic loads and other aspects.

Below we summarize the main results and conclusions from both test cases.

1.5.1 Case green

As discussed above with the prior reliability, the dominant load for 'case green' is not the external water level but the traffic load. Figure 1.5 depicts the sensitivity of the posterior reliability index, a measure of the updated probability of failure, to the assumptions made for the traffic load, both for the assessment conditions as well as for the observed and survived traffic load (with daily conditions for the lake water level). If we assume a traffic load of 13.3 kN/m^2 for the assessment conditions (red line), a common assumption in standard designs and assessments, and no observed traffic load we obtain the lower bound for the posterior reliability index of about 1.14. For these conditions, an increase of the observed traffic load leads to a rather limited increase in the posterior reliability index, because of other known and uncertain differences between assessment and observation conditions such as subsidence leading to a increasing head difference over the structure in time.

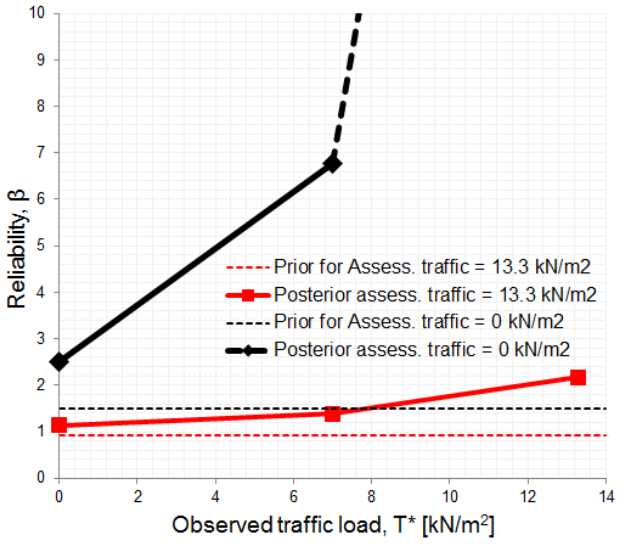


Figure 1.5: *Sensitivity of the posterior reliability index to assumptions in the traffic load in the assessment and the observed traffic load (for daily water level conditions).*

In contrast, assuming no traffic load for the assessment (black line) leads to a much more pronounced effect of reliability updating. Even for no traffic load at the observation the posterior reliability index increases to 2.51, while for higher observed traffic loads the posterior reliability increases drastically. For example, for values of 3 to 5 kN/m^2 for the observed traffic load, the reliability index increases to values ranging from about 4.5 to 5.5, implying a change in probability of several orders of magnitude compared to the prior. Moreover, the assumption of no traffic load (or at least a load with a low probability) for the assessment conditions is conceivable for 'case green', where there is no road on the crest but on the berm.

1.5.2 Case house

'Case house' with a building in the cross section showed a much higher reliability index of 5.6 already in the prior analysis. As typical for cases with a low probability of failure, the information of survived loads does not further increase the reliability estimate significantly. Loosely speaking, we are not surprised by the observation of survival and, hence, have no grounds to adjust the probability of failure on.

The sensitivity analyses gave some valuable insights into the effect of traffic load, the weight of the buildings and typically neglected resistance contributions, though. For all these effects there was only an effect on the prior reliability index, which was increased further to values of almost 7. The effect of reliability updating remained negligible.

As stated earlier, the 'case house' did illustrate that the reliability estimate can be substantially higher than would be assumed from semi-probabilistic relationships based on the factor of safety alone.

1.6 Limitations of the approach

The general reliability updating method has virtually no limitations in terms of applicability except for practical reasons like modeling and computational efforts. However, the approximation with fragility curves does have additional limitations through the simplifications it entails. In practical terms for slope stability, these boils down to the requirement that the failure surface (slip plane) needs to be essentially the same for observation and assessment (Figure 1.6), as is typically the case in safety assessments.

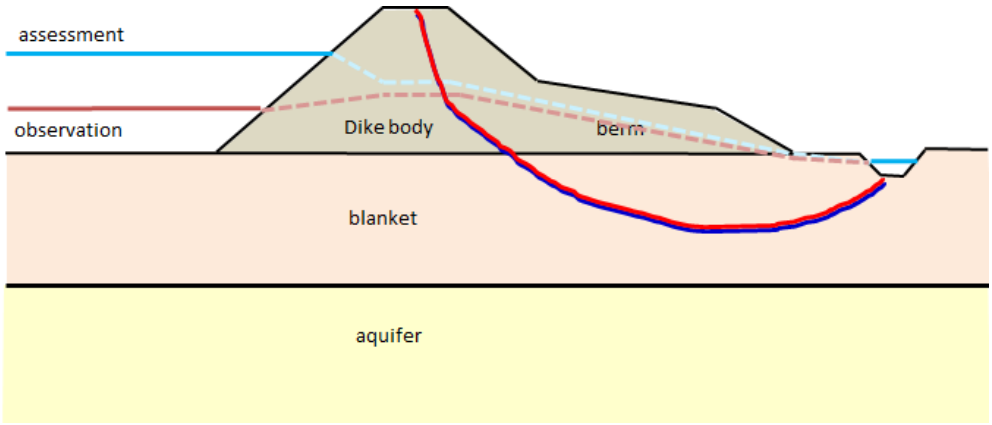


Figure 1.6: *Illustration of a situation where the critical sliding plane is the same in the assessment and the observation conditions, though the external water level differs between the both.*

Changes in time due to e.g. subsidence can be explicitly modeled and accommodated, as long as they do not affect the location of the critical slip plane significantly. On the other hand, if uplift conditions are critical for the assessment but were not observed during the survived loading, the location of the sliding plane may differ substantially between assessment and observation. Then the approximation with fragility curves is not applicable anymore.

It is emphasized that the above holds for any change between observation and assessment, not only for natural processes. The approximation with fragility curves can also be used in a reinforcement design setting, as long as the failure mode in question is not altered by the reinforcement measures. For example, a change in the outer slope angle or the replacement of the revetment will typically barely affect the inner slope stability.

1.7 Conclusions

The following main conclusions can be drawn from the examples provided in the background report (Schweckendiek and Kanning, 2016) and from the two case studies elaborated in this report:

1. The present case studies have demonstrated the applicability of reliability updating with past peformance, in particular the approach with fragility curves for slope stability, to real life conditions, namely two dike sections in the Markermeerdijken area.
2. The reliability updating method is a straightforward extension of conventional reliability analysis. In contrast to prior analysis, in reliability updating special attention needs to be paid to modeling the conditions at the time of the observation and how these differ from the assessment conditions.
3. Reliability updating with survival observation has a significant effect in terms of reducing the probability of failure, if:
   a) a significant load or load effect has been survived,
   b) the probability of failure is relatively high and dominated by epistemic (knowledge) uncertainties (typically soil properties),
   c) the structure has not changed or degraded substantially since the observation.

   With respect to (b), note that a high probability of failure also implies that there is a larger gap between the prior and the target reliability.
4. In the first case ('case green', no buildings) the probability of failure decreases significantly (orders of magnitude) through reliability updating. The results are highly sensitive to the assumptions made for the traffic load.
5. The second case ('case house') showed how aspects specific to dikes with buildings can be addressed, such as the contribution of foundation piles to the shear resistance or the effect of the weight of the building itself. The choice of which phenomena need to be included to which level of detail in the analysis is case-dependent. The effects of reliability updating for 'case house' were rather insignificant due to the high prior reliability, not due to the presence of the building.

Furthermore, the results demonstrate how probabilistic analysis avoids the conservatism necessarily introduced in a semi-probabilistic safety format, even without accounting for past performance. Dikes found to be unsafe based on a factor of safety can actually be safe in terms of the acceptable probability of failure.

Finally, the Markermeerdijken represent rather specific conditions. The dikes are constantly loaded by the lake water level, which is higher than the surface level in the hinterland. Extreme water levels on the Markermeer (roughly 1.5m higher than average) only have a very limited load effect in terms of increased pore water pressures in and under the dike. The dike stability

is more sensitive to traffic loads than to water level changes. These characteristics are not representative for the majority of the Dutch primary flood defenses.

1.8 Recommendations

We have the following main recommendations for further developing the method and enabling its use in practice:

1. Since the Markermeerdijken have specific characteristics, it will be necessary to investigate more case studies with varying characteristics which are more representative for the majority of the dikes in the primary Dutch flood protection system. Additional cases will also help to obtain better insight into how much effect can be expected in which situations. The POV-M research project of the Dutch flood protection program HWBP (*Hoogwaterbeschermingsprogramma*) includes plans to analyze three dike sections along the Hollandse IJssel, including the possibility of generating additional useful observations by test loading. Other obvious candidates are dikes in tidal estuaries, canal dikes and river dikes in the upstream area (in Dutch: bovenrivierengebied).
2. Reliability updating can also have a significant impact for other failure modes dominated by epistemic (knowledge) uncertainties (see conclusion 3). This is typically the case for geotechnical failure modes such as internal erosion (piping).
3. In the contemplated test cases the traffic load appeared to be the dominant load. For the typical Markermeerdijken conditions, as well as for many other dikes, observations of heavy prolonged precipitation and survival of the resulting increased pore water pressures conditions can be valuable information for reliability updating, too. Also load combinations such as precipitation plus traffic load can be used. Such analyses are planned to be included in the final version of this report.
4. Probabilistic modeling of the response of the phreatic surface level to loading by the external water level, possibly in combination with correlated precipitation, remains challenging and requires further research to establish best practices.
5. The approach with fragility curves is an approximation which brings about limitations. Other methods should be investigated in order to use reliability updating to its full potential, including updating of the (joint) probability distribution of the basic random variables. Even though some more advanced approaches may not be suitable for practical application, they can provide valuable insights and benchmark results for more practicable approximate methods.
6. Dike profiles with buildings and other objects are becoming increasingly important in Dutch dike assessments and reinforcements. In the current study, the relevant aspects such as the effect of the weight of a building or its pile foundation, have only been treated in a simplified fashion. For cases where these aspects matter, more robust and accurate approaches need to be applied and, if necessary, developed.
7. For dike sections where the traffic load is dominant, accurate modeling of the traffic load is essential for both assessment and the observation conditions. For example, the amount of traffic load which can be reasonably accounted for as observed and survived requires careful consideration.
8. A regular (a-priori) probabilistic analysis can be beneficial, even without considering survival observations. Probabilistic analyses also allow refining aspects such as the probabilities of external loads and load combinations, or the assessment of residual profiles or residual strength after initial sliding. Such refinements can potentially lead to significant reductions in the failure probability estimates and should be made available to practitioners.
9. The presented research efforts have focused on the assessment of an existing dike, potentially with small changes in conditions due to degradation. The approach itself also has potential for the survival information to be exploited in dike reinforcement designs. It is recommended to investigate this in further research.

In this context it should be noted that the present analyses have the characteristics of and 'advanced assessment' (in the Dutch safety assessment framework: toets op maat). The basic requirement for an 'advanced assessment' is the reliability estimate needs to be substantiated using state-of-the-art methods and models and defensible assumptions.

The remainder of this report contains detailed information and discussions of the case studies.

2 Introduction

2.1 Problem description and context

Slope stability assessments of dikes, just like most geotechnical problems, are typically dominated by the large uncertainties in soil properties. The resulting probabilities of (slope) failure are often rather large compared to the failure rates observed in the field, as experienced in the Dutch VNK2 project (Rijkswaterstaat, 2014) and illustrated in Figure 2.1.

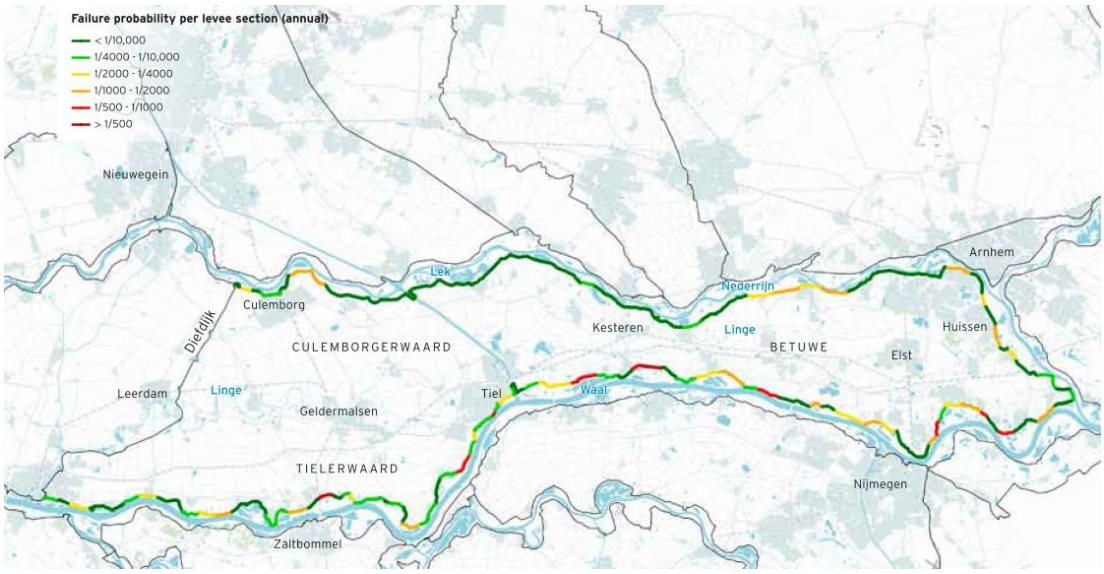


Figure 2.1: *Failure probabilities of the dike system Betuwe/Tieler- en Culemborgerwaarden according to Rijkswaterstaat (2014)*

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

Rijkswaterstaat is conducting a project to operationalize the concept of Reliability Updating with Past Performance (RUPP; in Dutch often referred to as *bewezen sterkte*) for advanced safety assessments and reinforcement designs of the primary Dutch flood defenses. Reliability updating means to update the estimate of the probability of failure using observations of past performance, here specifically the survival of observed load conditions. The focus in this first phase of the project is on the failure mode of instability of the inner slope, as many dikes were found not to meet the safety criteria for this failure mode in the statutory safety assessment of the Dutch primary flood defenses (Inspectie Verkeer en Waterstaat, 2011).

2.2 Objectives of the long-term development project

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

1 to develop and document a scientifically sound and practicable approach,

2 to confirm and illustrate the practical applicability of the approach on test cases with a level of detail and complexity which is representative for real life conditions.

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

1 **Background report** containing a scientifically sound description of the theory,

2 **Case studies** for testing and illustrating the applicability,

3 **Manual** containing a description of the method and its application for practitioners,

4 **Software** facilitating (a) probabilistic slope stability analysis and (b) use of the RUPP method by practitioners.

Notice that this test case report (2) and the accompanying background report (Schweckendiek and Kanning, 2016) (1) are primarily aimed at an expert reader in order to assess the soundness of the approach and the envisaged application, while the manual (3) will mainly address a non-expert audience.

2.3 Objectives of this report and approach

The main objective of this report is to demonstrate the practical applicability of reliability updating with observed load conditions to realistic dike stability assessments, using the approach with fragility curves as described in in the accompanying background report (Schweckendiek and Kanning, 2016).

The secondary objective is to gain insights into the potential impact (and the conditions governing the impact) of applying reliability updating to the failure mode slope instability, in particular for the Markermeerdijken. The Markermeerdijken (between Amsterdam and Hoorn) were found to be unsafe in the last national safety assessment of primary flood defenses (Inspectie Verkeer en Waterstaat, 2011). Consequently, a reinforcement program has been started, which will soon enter the design stage. The proof-of-concept study by Schweckendiek and Van der Krogt (2015) suggested that applying reliability using survived load conditions for the Markermeerdijken could have a significant impact on the scope of the reinforcement project. The current project aims to provide additional insights into how much effect can be expected for typical Markermeerdijken dike sections.

In order to address both objectives, two case studies from the Markermeerdijken area are investigated to a level of detail as commonly achieved in 'advanced safety assessments' (in Dutch: toets op maat). One of the dike sections is a regular clay dike on a peat subsoil without buildings or objects (in the remainder referred to as 'case green'); the other case is a clay dike on similar ground conditions but with a building in the cross section ('case house'), as a significant part of the Markermeerdijken has buildings in the profile.

We emphasize that the present report is not a formal assessment of the Markermeerdijken. The objective is to illustrate the relative impact of reliability updating with survival observations for various modeling choices. In order to realistically model the slope stability, the choices made by Halter *et al.* (2015) and (Zwanenburg, 2014a) at the start of the analyses were

followed as much as possible. These choices may have been subject to amendments, for example due to additional research carried out for the on-going dike reinforcement project.

Most conclusions drawn from the present analyses are case-specific and, therefore, cannot be generalized easily based on the results of only two similar cases.

2.4 Outline

After this introduction, the main assumptions and choices (i.e. starting points) made in this study with respect to slope stability modeling and reliability updating are described in chapter 3. Chapters 4 and 5 describe the case studies 'case green' and 'case house' respectively, starting with a description of the regular reliability analysis followed by the reliability updating and sensitivity analyses. The main observations from both case studies are summarized in chapter 6.

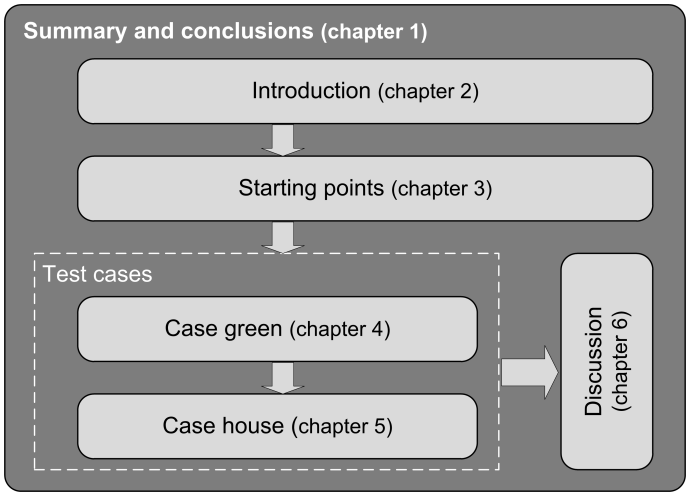


Figure 2.2: *Visual outline*

3 Starting points

This chapter discusses the main assumptions and choices (i.e. starting points) made in this study with respect to slope stability modeling and reliability updating for both case studies. For details refer to appendix B.

All modeling choices and assumptions have been made as much as possible in accordance with the Dutch safety assessment framework to be introduced in 2017 (WBI-2017). By doing so, the results are supposed to resemble the outcome of an 'advanced assessment' (in Dutch: toets op maat) closely.

Additional assumptions or deviations from the starting points shown in appendix B are stated explicitly per case in chapters 4 and 5 respectively.

3.1 Slope stability analysis and shear strength models

The modeling of dike stability in terms of the stability analysis methods and soil shear strength models has been carried out with the models as envisaged in WBI-2017, most importantly:

1. The stability of the inner slope of the dike in analyzed using two-dimensional limit equilibrium method Uplift-Van, which is capable of handling non-circular sliding planes, typically encountered in the Markermeerdijken area.
2. The follwing shear strength models are used:
   - for soil layers with expected drained behavior: Mohr-Coulomb (mainly sand layers)
   - for soil layers with expected undrained behavior: undrained shear strength based on critical state theory, more specifically the CSSM model as defined in WBI-2017 (Van Deen and Van Duinen, 2016)
   - The effects of partial saturation are neglected; soil layers are either assumed to be fully saturated or dry.

The D-Geostability software (Deltares, 2016) was used for the slope stability analysis.

3.2 Time of the assessment and time-dependent change in parameters

The time for which the assessment has to be done is typically in the future for Dutch dikes. For this report, the assessment time is 2023. This means that the anticipated assessment conditions in the year 2023 have to be modeled. This has an effect on the parameters that show time dependent behavior. In this report, the geometry is impacted as well as the water level distribution. The geometry is impacted by subsidence, which is approximately 0.01 m/year. Since the measured geometries are based on measurements in the year 2000, 0.25 m subsidence is incorporated in the assessment and 0.15 m in the observation using daily conditions. Hence, the difference in elevation between assessment and observation is 0.1 m. The pore water pressure modeling is adapted correspondingly. For the water level distribution, the distribution that is developed within the WBI-2017 project is used, this is the water level distribution for 2023 and includes time-dependent effects if applicable. Please refer to appendix B for more details about the above-mentioned points.

3.3 Model uncertainty

The model uncertainty of the applied Uplift-Van model in combination with CSSM undrained shear strength (i.e. SHANSEP) is accounted for in terms of a model uncertainty factor m_d (i.e. the calculated stability factor is multiplied by this random factor). According to Van Duinen (2015) the model factor for typical Dutch dikes follows a lognormal distribution with mean value $\mu_{m_d} = 1.005$ and standard deviation of $\sigma_{m_d} = 0.033$.

As there is still discussion about the probability distribution for the model uncertainty factor, we recommend practitioners to check and use the latest insights or consensus in future real-life probabilistic slope stability analyses.

3.4 Soil parameters and statistical characterization

As far as available, the soil parameters, including statistical parameters, were based on Halter *et al.* (2015). This includes S-ratios, m-exponents and POP-values of the CSSM model for different types of peat and clay, derived from a regional data set of direct simple shear tests (peat) and triaxial tests (clay). Notice that statistics based on a regional data set with a limited sample size per soil type leads to relatively large parameter uncertainty compared statistics based on local data such as derived in the DoV2-project (Zwanenburg, 2014a). Degradation of the shear strength parameters in time is not considered in this study, since there is no indication of significant degradation of the shear strength so far.

3.5 Pore water pressures

The pore water pressures were modeled in accordance with the assumptions made in the DoV2-project (Zwanenburg, 2014a; Halter and Effing, 2011). Rainfall is implicitly modeled with the phreatic surface. The modeling features of the so-called 'Waternet creator' are utilized in the used version of the D-Geostability software (Van Duinen, 2014). The Waternet creator enables modeling a parametric response of the pore water pressures to changes in the boundary conditions (e.g. the external water level), which enables stochastic treatment of the pore water pressures in the reliability analysis (Kanning, 2016).

3.6 External loads (traffic loads and buildings)

This section summarizes the assumptions commonly made in Dutch safety assessments for external loads such as traffic loads and the effects of the weight of buildings in the dike cross section.

Traffic load is a random phenomenon and we preferably model it as a random variable. However, since there are no well-established traffic load distributions (nor data to derive them), codified design values for traffic have been used in order to approximate the (FORM) design value to obtain a good approximation for the probability of failure. A uniform load of 13.3 kN/m^2 is assumed over a width of 2.5 m in the assessment conditions, usually on the crest of the dike. The degree of consolidation defines the (excess) pore water pressures as result of the load per soil layer. For both cases, the degree of consolidation was taken as 20% for the clay and peat layers, according to the choices made in the DoV2-project.

In standard assessments, the common conservative approach is to neglect the positive influences of a building on the slope stability. Essentially, a house is modeled as a gap in the dike body and the weight of the house is not taken into account, neither as external load acting on the moment equilibrium not the effect on the effective stresses.

Notice that the above assumptions are conservative assumptions for standard assessments.

Since the present analyses have the character of an advanced assessment, deviating assumptions can be made, if substantiated properly. For this reason, the assumptions above are considered in the so-called 'base case', variations of which will be considered in sensitivity analyses.

As discussed in Schweckendiek and Kanning (2016), conservative assumptions for the assessment conditions are by definition not conservative for the observed conditions. For this reason the 'base case' considers cautious assumptions for the observed conditions, namely no traffic load and a best guess estimate of the weight of a building, as detailed in appendix B. As for the assessment conditions, variations of which will be considered in sensitivity analyses.

3.7 Reliability analysis

The reliability analyses carried out in this study produce annual probabilities of failure, in accordance with the definitions of reliability targets in the Dutch safety assessment and design standards (OI, 2015). The reliability estimates in terms of probability of failure or reliability index are obtained using fragility curves as described in the accompanying background report (Schweckendiek and Kanning, 2016), using FORM to compute the fragility points. Appendix A contains a description of the workflow and software used[1].

3.8 Observed load conditions

For both test cases the base case considers survival of the daily conditions, followed by variations and sensitivity analyses. When referring to daily conditions in this report, technically we are referring to conditions for the mean value of the APT (arbitrary point in time value). In the specific case of the Markermeerdijken, the dike stability is relatively insensitive to the external water level, mainly due to the low permeability of the dike material itself and the relatively thick underlying low-permeability substrata. At the same time, the dike is constantly loaded by the lake and by the constantly high phreatic surface levels inside the dike. This combination makes the daily loading conditions promising for reliability updating, which is rather uncommon for other types of dikes such as river dikes. In the daily conditions, there are also relatively low uncertainties (e.g. phreatic level) compared to historic events with higher water levels. Subsequent to the daily conditions we also explored the effects of other load observations by means of sensitivity studies with higher water levels and traffic loads.

3.9 Epistemic versus aleatory uncertainties

Schweckendiek and Kanning (2016) showed the importance of making the distinction between time-invariant properties (i.e. epistemic, reducible uncertainty) and properties that are variable in time (i.e. aleatory, irreducible uncertainty). In the presented approach, we chose to assign the basic random variables to either category, while in reality the respective probability distributions may contain contributions of both epistemic and aleatory uncertainty. In other words, we need to decide per random variable whether the uncertainty is predominantly epistemic or predominantly aleatory. In reliability updating using survival information, aleatory is a safe choice in case of doubt, as the effect in the reduction of the probability of failure is less than if choosing for epistemic.

Table 3.1 contains the variables that affect the reliability of a dike for slope stability of the inner slope as considered in the WBI-2017 safety assessment framework, including a column indicating whether we assume perfect auto-correlation in time for the two case studies elaborated

[1] A manual (in Dutch: werkwijzer) to enable practitioners to carry out reliability analyses and updating for slope stability in under development as a product of the overarching development project.

in this report[2]. Most variables can be modeled as continuous stochastic variable (e.g. soil parameters) and can be implemented directly in the reliability analysis through the fragility curves. Some variables cannot be modeled practically as continuous stochastic variable (e.g. geometry), therefore have to be implemented via a (discrete) scenario, this is also indicated in Table 3.1. Notice that the volumetric weight is implemented as a deterministic variable, as is the traffic load. The sensitivity analysis for including the volumetric weight as random variable (see appendix D) shows that effects of the volumetric weight on the moment equilibrium and on the yield stress estimates roughly cancel each other in the contemplated conditions.

Table 3.1: *Random variables in slope stability of dikes with undrained analysis, including the modelling choices relevant for reliability updating as considered for the two case studies in this report*

Variable	Category	Can be modelled as continuous stochastic variable?	Correlated in time (fully)	Implementation
Su ratio, S	Soil property	yes	yes	in fragility curve
Strength increase exponent, m	Soil property	yes	yes	in fragility curve
Yield stress, σ'_y	Soil property	yes	yes	in fragility curve
Volumetric weight, γ	Soil property	yes	yes	deterministic
Friction angle sand, ϕ	Soil property	yes	yes	in fragility curve
Outside water level, h	Geohydrological and load	yes	no	'outside' fragility curve
Leakage Length outside, λ_{out}	Geohydrological	yes	yes	in fragility curve
Leakage Length inside, λ_{in}	Geohydrological	yes	yes	in fragility curve
Intrusion Length, IL	Geohydrological	yes	yes	in fragility curve
Phreatic line, PL	Geohydrological	yes	no	scenario
Polder water level, h_p (best estimate)	Geohydrological	yes	no	deterministic
Traffic load, T	Load	yes	no	deterministic (design value)
Subsidence (best estimate)	Schematisation	no	yes	deterministic
Soil layering	Schematisation	no	yes	scenario
Model uncertainty, m_d	Model	yes	yes	in fragility curve

More considerations on the choices made for the two case studies are provided in appendix B. Note that most modeling choices are case-specific and not necessarily generally applicable.

[2]The choices for assuming a variable as time-invariant or for modeling it with a continuous or discrete probability distribution can vary from case to case. For example, in some situations the intrusion length is very sensitive to the duration of the high water conditions. If the duration is random and not explicitly accounted for, the intrusion length can better be assumed as uncorrelated in time.

4 Case green: clay dike on peat

This chapter describes the test case 'case green', including geometry, assumptions and modeling choices. Both the assessment conditions as well as the observed conditions are analyzed, in first instance for a so-called base case followed by parametric studies and sensitivity analyses. The base case consists of a set of conservative assumptions for the assessment and observation situation. For a description of the reliability analyses and derivation of fragility curves, reference is made to Appendix A.

The first step is to compute the prior probability of failure based on the assessment fragility curve (section 4.1). This is the probability of failure without the effects of incorporating the survival information. Subsequently, the observation fragility curve is derived and reliability updating is performed for the base case (section 4.2) in the so-called posterior failure probability and, equivalently, the posterior reliability index.

The water level is not the only dominant load variable for case green, traffic appears to be important as well. Hence, the effect of traffic load on reliability updating in case green is investigated in section 4.3. The sensitivity of the prior and posterior probability of failure to different assumptions in modeling the phreatic surface level is investigated in (section 4.4); the sensitivity to the thicknesses of the clay and peat layers under the dike is contemplated in (section 4.5).

Finally, benchmark analyses are presented comparing the results obtained with the approximation with fragility curves and exact solutions, such as obstained with Monte Carlo Simulations (MCS) directly using the stability models.

A summary of the results and more elaborate conclusions are presented in chapter 6.

4.1 Prior analysis

This section describes the prior analysis for the base case of the assessment conditions. All these results do not yet consider the influence of an observation. The base case is close to the detailed assessment in WBI-2017, i.e. considering a set of conservative assumptions for pore pressures, traffic loads etc.

4.1.1 Assessment conditions (base case)

This section describes the assumed case-specific assessment conditions of 'case green', for the general starting points refer to appendix B.2.

Geometry and soil layering

The surface and subsoil geometry is shown in Figure 4.1. 'Case green' consists of a clayey dike body on peat and clay layers. The assessment conditions assume 0.25 m anticipated subsidence of the hinterland surface between the measurement of the geometry and the reference period for the assessment (i.e. the year 2023). For the soil parameters refer to Tables B.1 and B.3 in the appendix.

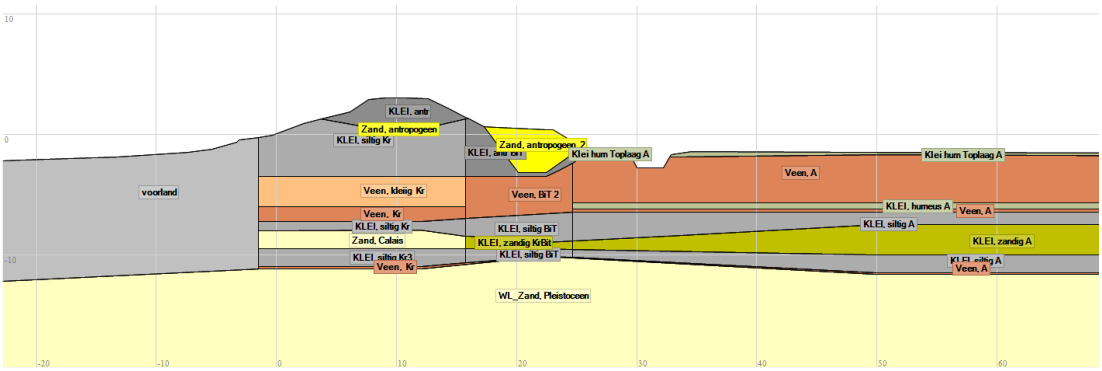


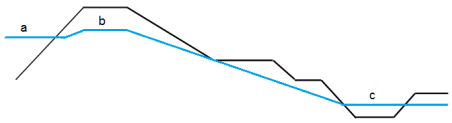
Figure 4.1: *Case green: Geometry and soil layering.*

Geohydrological boundary conditions

The phreatic line is modeled using a set of discrete points in the cross section at locations a, b and c. Table 4.1 provides the case-specific parameters. Note that the relation between the outer water level (a) and the phreatic level in the dike body (b) is modeled as a linear relation according to: $b = 1.348 + 0.419a$ [1].

During daily conditions, the stationary head in the first aquifer ('WL Zand Pleistoceen') is measured at NAP -1.84 m (Zwanenburg (2014b). For the leakage lengths values of $\lambda_{in} =$ 2000 m and $\lambda_{out} =$ 3000 m were assumed in order to match the measured response factor of 0.4 in head level (see section B.2).

Table 4.1: *Case green: schematisation of the phreatic level for two arbitrarly chosen water levels, assessment situation.*



Phreatic level in [m] above NAP, per location			
example water level h	a	b	c
Daily	-0.40	+1.18	-1.57
High water	+1.15	+1.83	-1.57

Traffic load

In the base case, the traffic load is assumed to act on the the crest with a 2.5 m wide uniform load of 13.3 kN/m^2 (red area indicated in Figure 4.2).

4.1.2 Prior analysis (base case)

Deterministic analysis

The deterministic analysis provides insight in the stability factors and critical slip planes at low and high water levels. The stability factors are shown in Table 4.2.

The safety factors with low characteristic values were determined using 5%-quantile values for resistance parameters and no material factors applied, which is in line with the latest recommendations for WBI-2017 (Kanning *et al.*, 2015). For the design values, the safety factor with characteristic values is divided by the model factor $\gamma_d = 1.06$.

[1] Here, a and b are the water level and the phreatic line level in the dike body respectively, and a linear interpolation is implemented between the two values shown in Table 4.1. The underlying assumption is that high lake water levels are practically perfectly correlated with pro-longed precipitation events affecting the phreatic surface by precipitation (for details refer to appendix B.2).

Table 4.2: *Case green: Safety factors for mean, characteristic and design values for the base case assumptions (charateristic values were taken as 5%-quantiles of the resistance parameters).*

Example water level h	SF Mean values	SF Characteristic values	SF_{char}/γ_d*
NAP-0.40m	1.13	0.73	0.69
NAP+1.15m	1.11	0.71	0.67

*) γ_d =1.06

The results in Table 4.2 essentially show that (a) the influence of the external water level on the stability is very limited (notice that the considered range roughly represents the difference between average and design values) and (b) the effect of the large difference between mean values and characteristic values for the soil strength properties. The explanation for the low influence of the external water level lies in the limited effect on the pore water pressures along the slip plane. The phreatic surface in the dike is already rather high in daily conditions, to which an increase of the external water level does not add much. Furthermore, the intrusion of groundwater into the peaty and clayey blanket does not affect the pore water pressures significantly, either.

Notice that the safety factor with mean values is clearly above 1, meaning that we do expect the dike to be stable. In discussions about survived conditions there is often a misunderstanding that safety factors below unity indicate that either the model or the assumptions made are wrong, because the dike would have failed by now, if they were correct. Yet, most of the time the safety factors referred to were obtained with characteristic or design values, in which case we can only conclude whether the dike complies with the safety requirements, not if it matches a performance observation.

The critical slip surface for mean values as shown in Figure 4.2 is relatively deep and goes through both the peat and clay layers. The critical slip surfaces for characteristic and design values are similar, also for high and low external water levels.

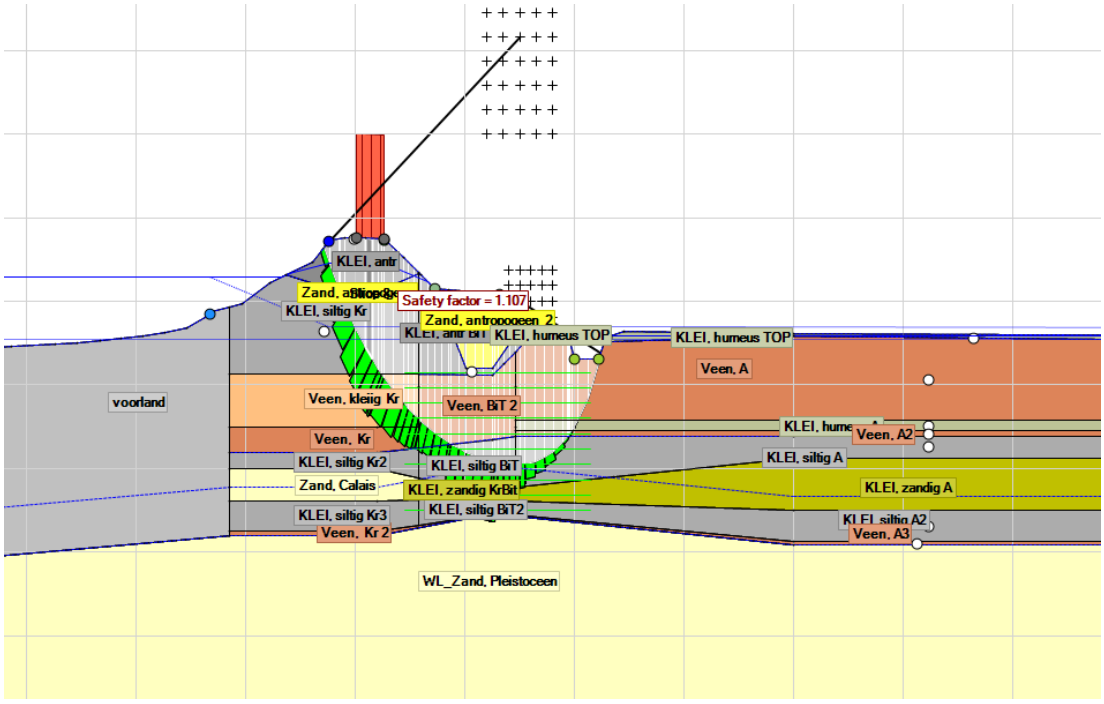


Figure 4.2: *Case green: Slip surface and safety factor with mean values and for a water level of 1.15 m above NAP.*

Fragility curve

The results of FORM analyses conditional on different water levels are reflected in the fragility points listed in table Table 4.3 and in the beta-h curve depicted in Figure 4.3.

Table 4.3: *Case green: calculated reliability index (β) and probability of failure (p_f) for different outside water levels - base case assessment.*

h[m] above NAP	$\beta \mid h[-]$	$p_f \mid h[-]$
-0.40	0.95	$1.7 \cdot 10^{-1}$
0.00	0.89	$1.9 \cdot 10^{-1}$
0.40	0.84	$2.0 \cdot 10^{-1}$
0.80	0.78	$2.2 \cdot 10^{-1}$
1.20	0.73	$2.3 \cdot 10^{-1}$

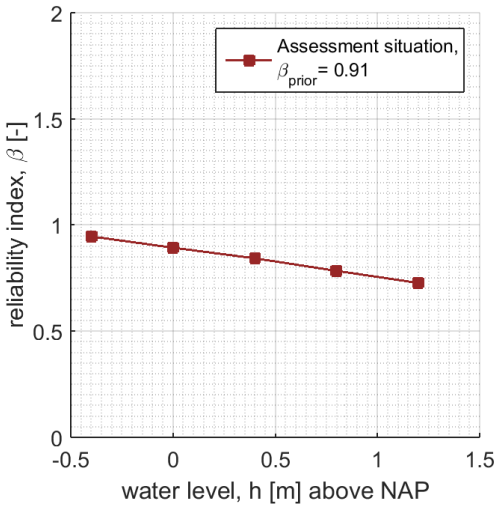


Figure 4.3: *Case green: beta-h curve showing the reliability indices conditional to a range of water levels.*

Prior reliability

Integrating beta-h over the water levels weighted by the probability (PDF) of the water level (Figure B.4) give the prior probability of failure as shown in Table 4.4.

Table 4.4: *Case green: prior reliability index (β) and probability of failure (p_f) for the base case conditions.*

Design point h[m] above NAP	$\beta[-]$	$p_f[-]$
-0.15	0.91	$1.8 \cdot 10^{-1}$

The prior annual reliability index for case green of 0.91 (corresponding to an annual probability of failure of 18%) is very high and clearly far from the acceptable values of 4.6 in the Markermeerdijken area as stated in OI (2015). The design point value of the water level is rather close to the mean lake level of 0.4 m below NAP, confirming the insignificant share of the uncertainty in the water level with respect to the total uncertainty.

The FORM influence coefficients (α^2) are presented in Table 4.5 (additionally, the design point values in real space are reported in appendix E.1). The squared influence coefficients indicate that more than 80% of the total uncertainty stems from the soil strength parameters. Note that the related uncertainty is supposed to be epistemic and reducible.

Table 4.5: *Case green: FORM influence coefficients (α^2) for the base case assessment conditions. "Pore water pressures" refers to the sum of the α^2-values of intrusion length and leakage lengths.*

Variable	α^2
S-ratio	0.50
m-exponent	0.06
Yield Stresses	0.29
Friction angle	0.00
Pore water pressures	0.00
Model uncertainty	0.15
Outside water level	0.00
Total $\Sigma(\alpha^2)$	1.00

The slip plane for the stability with design point values is presented in Figure 4.4. Its shape is virtually the same as for mean values. Notice that the safety in the design point does not necessarily equal one but the design point value of the model factor.

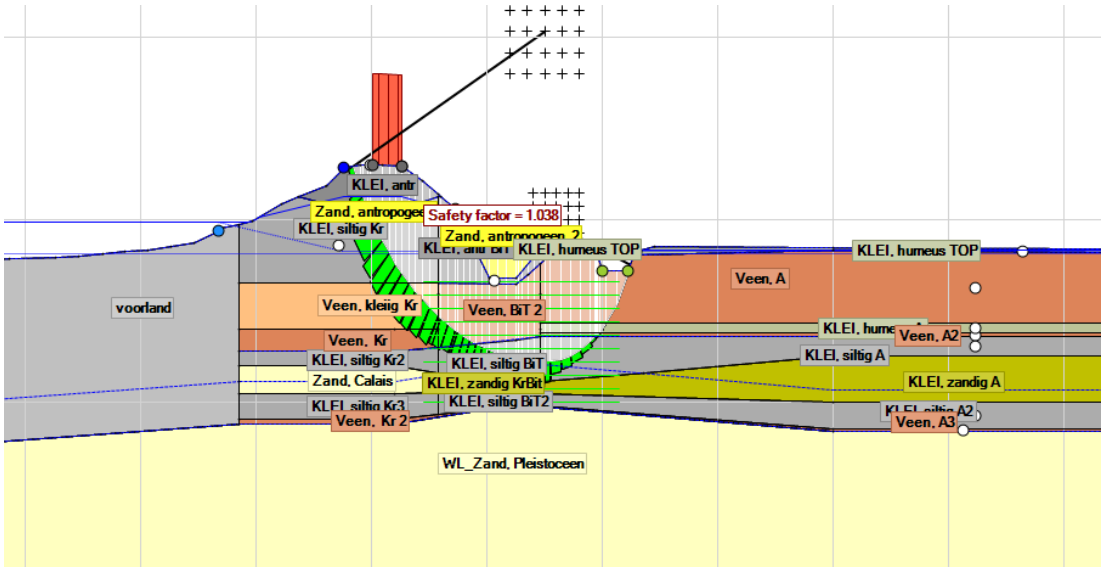


Figure 4.4: *Case green: slip surface and safety factor in the design point for the base case assessment conditions.*

4.2 Reliability updating

4.2.1 Observation conditions (base case)

This section describes the observation conditions for the base case of 'case green'. There are the following differences between the assessment and the observation conditions:

- **Geometry**: There is 0.1 m less subsidence in the observation as compared to the assessment. This is reflected in a 0.1 m thicker peat layer 'Veen A'; see section 3.2.
- **Geohydrology**: The polder level (c) is also 0.1 m higher polder level in the observation that reflects the difference in subsidence between the observation and assessment.
- **Traffic**: In the observation base case, there is no traffic load accounted for, as opposed to the 13.3 kN/m^2 traffic load in the assessment.

Neglecting the traffic load in the observation conditions is a conservative assumption, as we do not account for the survival of a traffic load or its load effect in the observation conditions. Incorporating traffic loads in the observation conditions increases the effect of reliability updating; the posterior probability of failure decreases further.

The remaining parameters are the same in observation and assessment.

4.2.2 Fragility curve for observation conditions (base case)

This section discusses the fragility curve of the observation situation. Note that this is excluding the observed water level. Reliability updating with various observed water levels is discussed in the next section.

Deterministic analysis

Table 4.6 contains the safety factors for mean, characteristic and design values for the observation conditions. Comparison with the assessment condition (Table 4.2) illustrates that all reported safety factors are 0.05 higher in the observation than in the assessment. The difference partially reflects the explicitly modeled degradation through subsidence[2] and partially the difference in assumed traffic loads.

Table 4.6: *Case green: calculated SF for mean and characteristic values - base case observation.*

Example water level h	SF Mean values	SF Characteristic values	SF_{char}/γ_d*
NAP-0.40m	1.18	0.78	0.74
NAP+1.15m	1.16	0.76	0.72

*) γ_d =1.06

Fragility curve

The reliability updating method with fragility curves requires determining fragility curves for both the assessment as well as the observation conditions. Table 4.7 contains the reliability indices and probabilities of failure conditional to the the set of external water levels used to determine the fragility curve for the observation conditions. Figure 4.5 compares the thus obtained fragility curves for assessment and observation. Both fragility curves are relatively flat, indicating limited dependence on the water level, and the reliability indices for the observation are clearly higher than for the assessment. According to the theoretical considerations

[2] Subsidence on the landside of a dike typically leads to an increasing head difference over the structure (i.e. load) in time. That is why we classify it as degradation here; the stability decreases with time.

in Schweckendiek and Kanning (2016) we do not expect a significant reliability updating effect for these base case conditions.

Table 4.7: *Case green: calculated reliability index (β) and probability of failure (p_f) for different water levels for the base case - observation.*

h [m] above NAP	$\beta \vert h[-]$	$p_f \vert h[-]$
-0.40	1.57	$5.8 \cdot 10^{-2}$
0.00	1.51	$6.5 \cdot 10^{-1}$
0.40	1.46	$7.2 \cdot 10^{-1}$
0.80	1.40	$8.1 \cdot 10^{-1}$
1.20	1.34	$9.0 \cdot 10^{-1}$

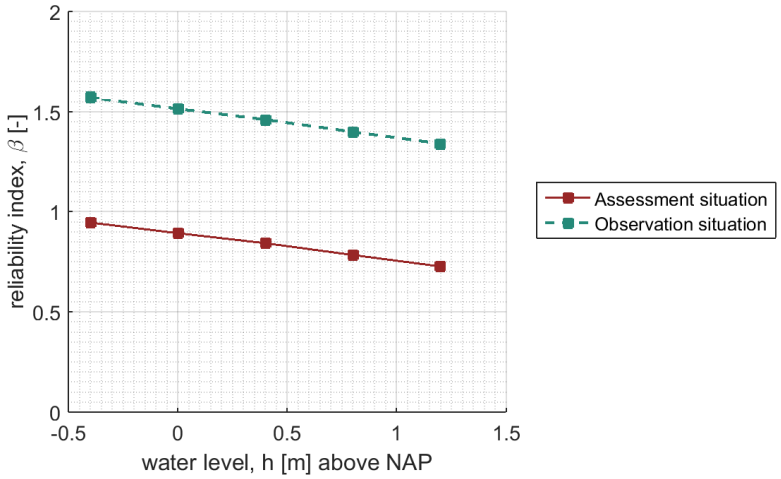


Figure 4.5: *Case green: fragility curves for the assessment and observation conditions for the base case.*

4.2.3 Correlation between assessment and observation

As stated in the background report (Schweckendiek and Kanning, 2016), we need to account for the auto-correlation in time between the assessment conditions and the observation conditions. High auto-correlation implies that most of the uncertainty is epistemic and reducible, in which case the effect of reliability updating is higher than for low correlation, in which case more aleatory, non-reducible uncertainty is involved.

The (linear) correlation coefficient ρ between the two resistance terms, h_c and $h_{c,obs}$, can be estimated by the following formula[3]:

$$\rho \approx \sum_i \alpha_i^p \alpha_i^f \rho_i^{p,f} \tag{4.1}$$

where α_i^p and α_i^f are the FORM influence coefficients of variable i for the observation (past) and for the assessment (future) conditions, respectively. In this report we use the average α_i of the fragility points in the beta-h curves. The correlation coefficient $\rho_i^{p,f}$ describes the correlation between variable i between the observation and the assessment, thus effectively the auto-correlation in time for the individual random variables. In this specific case study,

[3]Note that this formula works if we assume the basic random variable to be uncorrelated to eachother. For correlated variables an extended version of the formulation needs to be used.

we assume each basic random variable to be epistemic (i.e. time-invariant, see Table 3.1, implying $\rho_i^{p,f} = 1$).

Table 4.8 contains the FORM influence coefficients for case green.

Table 4.8: *Case green: FORM influence coefficients (α) for the base case assessment and also observation situations.*

	α_i^f				α_i^p					
Variable i	h=-0.4	h=0.4	h=1.2	average	h=-0.4	h=0.4	h=1.2	average	$\rho_i^{p,f}$	$\alpha_i^p \alpha_i^f \rho_i^{p,f}$
S-ratio	0.712	0.707	0.702	0.707	0.705	0.701	0.696	0.701	1.0	0.495
m-exponent	0.235	0.247	0.256	0.246	0.217	0.229	0.239	0.228	1.0	0.056
Yield Stresses	0.536	0.538	0.542	0.539	0.538	0.541	0.544	0.541	1.0	0.291
Friction angle	0.003	0.003	0.002	0.003	0.003	0.003	0.003	0.003	1.0	0.000
Pore water pressures	0.017	0.010	0.009	0.012	0.029	0.022	0.020	0.024	1.0	0.000
Model uncertainty	0.388	0.386	0.383	0.386	0.407	0.405	0.402	0.405	1.0	0.156
									$\sum_i =$	0.999

Notice that the correlation does change slightly with the water level between the fragility points, yet the correlation coefficient of 0.999 implies that we have virtually perfect auto-correlation in time for case green.

4.2.4 Reliability updating (base case)

Using the fragility curves for the assessment and observation, the posterior probability of failure is calculated for various water levels (h^*) considered as observed (see Schweckendiek and Kanning (2016) for the respective method). It should be noted that the observed water level -0.2 m above NAP is the daily summer level, +0.2 m above NAP is approximately the high water level of 1998 and +1.0 m above NAP is a fictitious extreme observation purely to illustrate the sensitivity of the effect to the observed water level. The resulting posterior reliability estimates in Table 4.9 indicate that the reliability updating effect for the base case with conservative assumptions is indeed not very significant. The prior reliability was 0.91 and the posterior reliability index for daily conditions is 1.14. Also increasing the observed water levels does not make a large difference for the effect on the reliability index, which is not surprising given the relative insensitivity of the stability to the water level.

Table 4.9: *Case green: calculated posterior reliability index (β) and probability of failure (p_f) - base case.*

h^* [m] above NAP	$\beta \vert h^*[-]$	$p_f \vert h^*[-]$
-0.2	1.14	$1.3 \cdot 10^{-1}$
$+0.2$	1.17	$1.2 \cdot 10^{-1}$
$+1.0$	1.26	$1.0 \cdot 10^{-1}$

The fact that we see a limited influence of reliability updating, even though the case is dominated by epistemic uncertainties in the soil properties leading to a high probability of failure, is mainly due to the significant difference between assessment and observation situation. The difference is caused mainly by the radical choice to (deterministically) account for the presence of a traffic load in the assessment and none at all in the observation conditions for the base case. Therefore, the influence of the assumptions regarding the traffic load are further investigated in the next section.

4.3 Sensitivity to the traffic load

Figure 4.7 depicts the influence of traffic load and outside water level on the stability factor and the reliability index. Both metrics are clearly more sensitive to changes in traffic load than changes in outside water level for the contemplated relevant ranges. This illustrates the actually dominant load variable in this case is the traffic load, not the water level. The implication in a reliability updating context is that (a) the (probabilistic) modeling of the traffic load is crucial and (b) that observations of survived traffic loads may be more influential than observations of observed water levels.

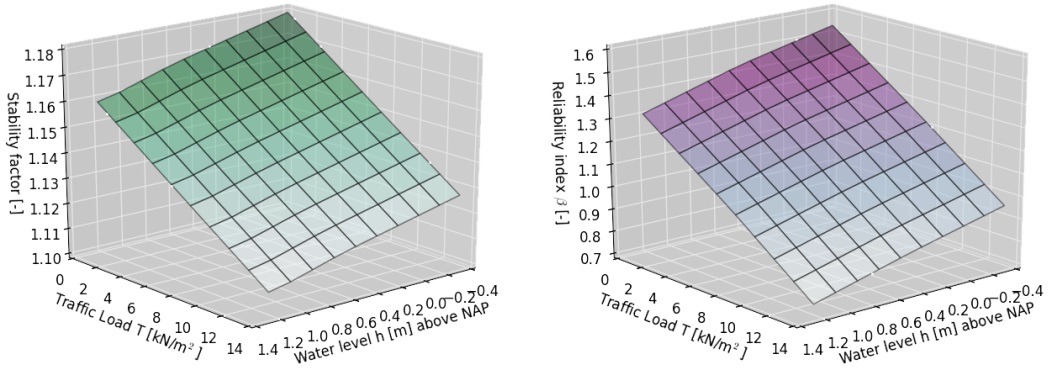


Figure 4.6: *Case green: Safety factor (based on mean values) and prior reliability index as function of traffic load and water level (h) and traffic load (T).*

4.3.1 Beta-h curves for various traffic loads

In order to further investigate the sensitivity of the posterior reliability estimate to the traffic load, we consider three different values for the traffic load for both the assessment and observation conditions:

a. no traffic load (0 kN/m^2)

b. design value (13.3 kN/m^2)

c. intermediate value (7 kN/m^2, arbitrarily chosen)

Table 4.10 and Figure 4.7 contain the fragility curves for assessment and observations conditions for the traffic load values stated above. For the same traffic load in assessment and observation, the curves are very close to each other. The remaining difference is due to the assumed subsidence between the moment of observation and the reference period for the assessment. On the other hand, the differences for different traffic loads are large. The base case, in fact, combines the curve at the very bottom for the assessment with the curve at the very top for the observation. This makes the limited effect of reliability updating plausible from a theoretical standpoint.

Table 4.10: *Case green: fragility curves for water levels (h) and traffic load (T in kN/m^2)*

	Assessment, $\beta\|h$			Observation, $\beta\|h$		
h [m] above NAP	T = 0	T = 7	T = 13.3	T = 0	T = 7	T = 13.3
-0.4	1.53	1.23	0.95	1.57	1.27	0.99
0.4	1.43	1.13	0.84	1.46	1.16	0.89
1.2	1.30	1.01	0.73	1.34	1.05	0.77

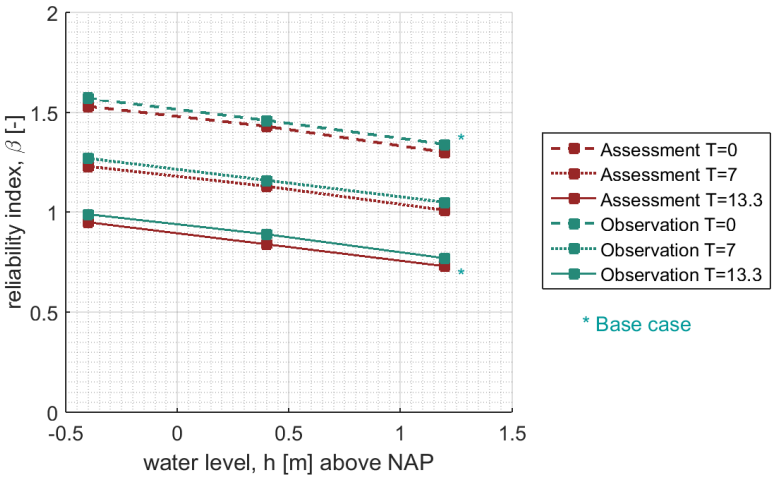


Figure 4.7: *Case green: fragility curves for water levels (h) and traffic load (T in kN/m^2)*

The subsequent sections discuss the effect for the following assumptions regarding the traffic load in the assessment conditions, while varying the observed traffic load:

1 no traffic load,

2 design value traffic load,

3 probability distribution for the traffic load ('PMF', defined in section 4.3.4).

Notice that for these assumptions also the prior reliability estimate changes to the values as given in Table 4.11. The value for the traffic load with probability distribution (PMF) is virtually the same as for no traffic load as the traffic load is considered to occur only with low probability together with water levels relevant for inundation.

Table 4.11: *Case green: prior reliability index β for different assumptions regarding the traffic load (T).*

traffic load T [kN/m^2]	prior reliability index $\beta[-]$
13.3	0.91
0	1.50
PMF	1.50

4.3.2 Assumption 1: No traffic load in assessment

Figure 4.8 shows the posterior reliability considering combined observations of both the water level (h^*) and the traffic load (T*). The observed traffic load is shown on the horizontal axis and the posterior reliability on the vertical. The prior reliability index of 1.50 updates to 2.51 for observation of survival of the daily water level and without considering an observed traffic load. Furthermore, the posterior reliability index is highly sensitive to the observed traffic load.

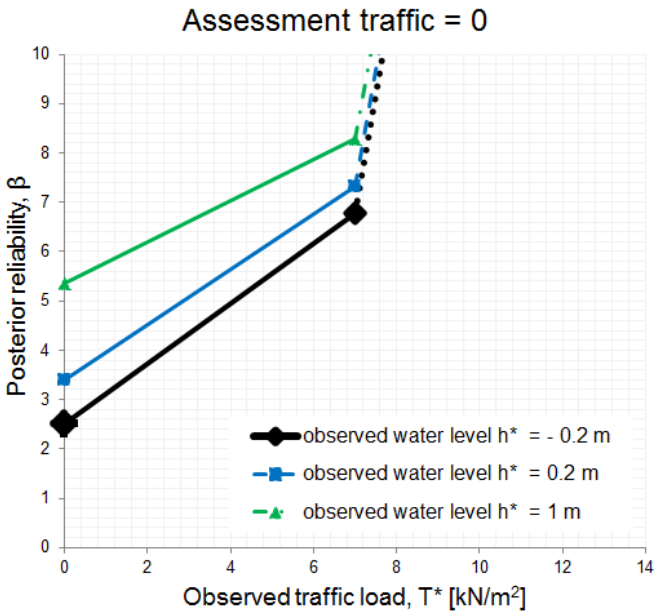


Figure 4.8: *Case green: posterior reliability for combinations of observed water level (h^*) and observed traffic load (T*) with a traffic load of T = 0 kN/m^2 for the assessment conditions. The posterior reliability is computed for 3 values of the observed traffic loads: 0, 7 and 13.3 kN/m^2.*

4.3.3 Assumption 2: Design value of the traffic load

This section presents the same results as the previous section with the only difference that the design value of 13.3 kN/m^2 is considered for the traffic load in the assessment conditions. Figure 4.9 shows that, as expected, the increase in reliability is much less pronounced than for assuming no traffic load in the assessment. The reliability updating effect increases with observed traffic load, but much less compared to the precious section. For example, only an observed traffic load that equals the assessment situation (i.e. T* = 13.3 kN/m^2) implies substantial increase in reliability index (from 1.14 to 2.17). Supposing the traffic load observation was made in combination with a higher water level (green and blue lines) we see significantly more effect.

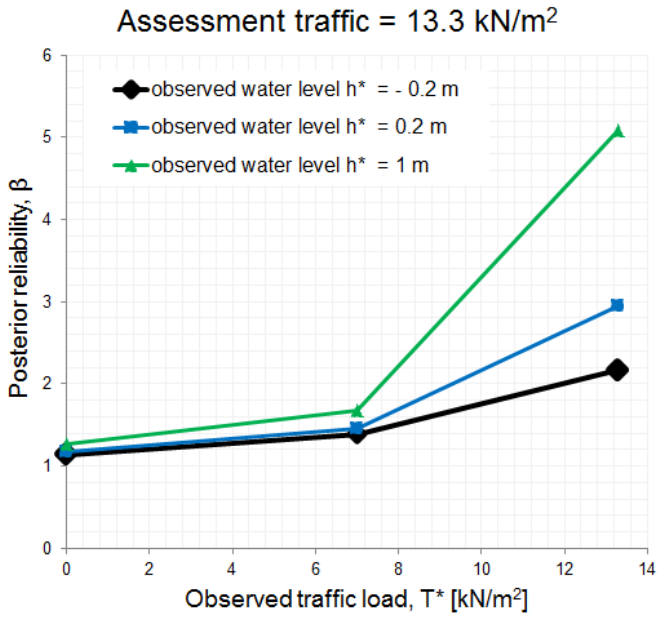


Figure 4.9: *Case green: posterior reliability for combinations of observed water level (h^*) and observed traffic load (T*) with a traffic load of T = 13.3 kN/m^2 for the assessment conditions. The posterior reliability is computed for 3 values of the observed traffic loads: 0, 7 and 13.3 kN/m^2.*

4.3.4 Assumption 3: Probability distribution for the traffic load

The previous assumptions considered assessment conditions with and without traffic. Ideally, we would use a probability distribution for the occurrence of traffic loads on an annual basis. Such a distribution is not available, so in this sensitivity study we employ a probability mass function (PMF), assigning probabilities to the values 0, 7 and 13.3 kN/m^2 as defined in Table 4.12.

Table 4.12: *Case green: probability mass function (PMF) of the traffic load in the assessment conditions for sensitivity analyses purposes.*

traffic load T [kN/m^2]	probability
0	0.989
7	0.010
13.3	0.001

The reason for assigning rather low probability values to heavy traffic in 'case green' is that there is no road on the crest, only a bicycle lane. The actual road is on the berm, where the influence on the stability is much less (see additional sensitivity analysis in in appendix, section E.3). That means that heavy traffic on the berm would only be expected in emergency conditions such as for flood fighting purposes. Even for flood fighting, the likely access road to critical locations would be the road on the berm.

As stated earlier, the prior probability based using the probability distribution for the traffic load is very close to the case without traffic (assumption 1). Also the posterior reliability estimates depicted in Figure 4.10 are very close to assumption 1, without traffic.

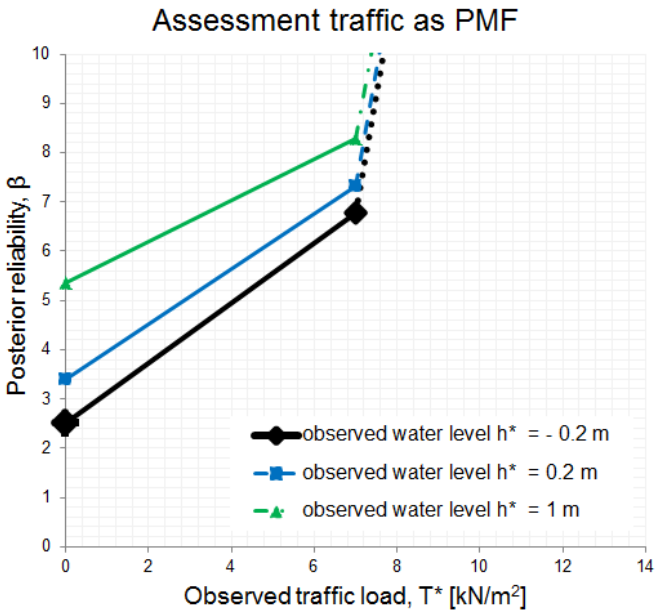


Figure 4.10: *Case green: posterior reliability for combinations of observed water level (h^*) and observed traffic load (T*) with a probability distribution for the traffic load for the assessment conditions as defined in Table 4.12. The posterior reliability is computed for 3 values of the observed traffic loads: 0, 7 and 13.3 kN/m^2.*

Note that in Figures 4.8, 4.9 and 4.10 the analyses were carried out for the values 0, 7 and 13.3 kN/m^2 of the observed traffic load. The lines in the figures are linear interpolations between the computed points. For very low probabilities of failure, the computed reliability indices become rather meaningless. Hence, especially the interpolation between the known points in the high reliability region needs to be treated with care.

4.3.5 Traffic location

The analyses presented so far consider the traffic load at the crest of the dike, for both the assessment and the observation situation. Computations were repeated considering the traffic load at the berm of the dike (where there is actually a road). The same qualitative conclusions can be drawn as for the previously presented assumptions. However, since the traffic load at the berm has less impact on the stability than at the crest, the prior reliability estimates are higher and relative reliability updating effect is lower. For details refer to appendix, section E.3.

4.4 Sensitivity to phreatic surface level

In the present study, the level of the phreatic surface is considered as dependent on the outside water level, using a deterministic (relative) response to the lake water level (appendix B). In this section we contemplate the sensitivity of the prior and posterior probability of failure to two modeling assumptions regarding the phreatic surface level. First, this subsection examines the effect of assuming the relative response of the phreatic surface to external loading to be uncertain (random) as opposed to deterministic; secondly, we investigate the effect of a very high (fictitious) observed phreatic level as an indication of the potential impact of monitoring pore pressures during extreme precipitation events.

4.4.1 Uncertain phreatic response

As opposed to the analyses presented hithero, in this section we consider the phreatic surface response so far as the "best guess" (expected values), adding to that a random error with a normal distribution with a standard deviation 0.30 m (i.e. $error \sim N(0, 0.3)$). Note that the realization of the random error is applied to all characteristic points in the phreatic line model evenly. In other words, the error term is assumed to be a systematic estimation error and fully correlated in the cross section.

The sensitivity analysis is made for the assumption of no traffic load in both the assessment and the observation conditions. The resulting fragility curves (Figure 4.11) show that including the random error (green lines) slightly increases the conditional reliability indices for both assessment and observation. Apparently, including the possibility of a slightly more favorable response of the phreatic surface has a greater net effect than the possibility of slightly more unfavorable response.

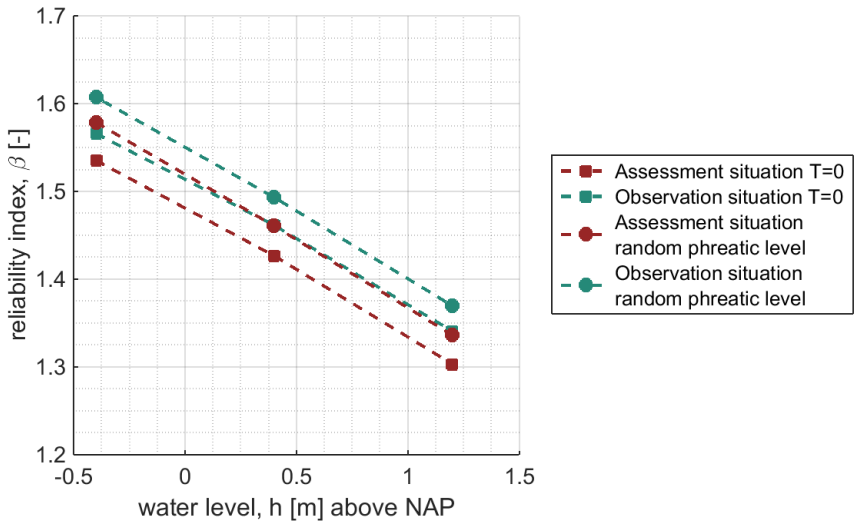


Figure 4.11: *Case green: fragility curves for the assessment and observation conditions for the phreatic level sensitivity analysis assuming a random error in the response of the phreatic surface to the lake water level.*

Consequently, the resulting prior reliability index (Table 4.13) is also slightly higher than without random error. Overall, the effect of taking a random error with the assumed magnitude ($\sigma = 0.3$) is insignificant in the prior analysis, which is confirmed by the design point value of the random error of 0.05 m and its influence coefficient of $\alpha^2 = 0.01$.

Table 4.13: *Case green: prior reliability index (β) and probability of failure ($P(F)$) for the phreatic level sensitivity analysis assuming a random error in the response of the phreatic surface to the lake water level.*

	$\beta[-]$	$P(F)[-]$
Assessment T=0kPa	1.50	$6.7 \cdot 10^{-2}$
Assessment T=0kPa random err	1.54	$6.2 \cdot 10^{-2}$

The posterior reliability in this sensitivity study is calculated for an observed water level of -0.20m w.r.t. NAP. Note that, as we consider the error in the phreatic line to be random (aleatory), the correlation between assessment and observation resistance as represented by the fragility curves decreases slightly when including the error (from $\rho = 0.995$ to $\rho = 0.9898$). However, this decrease in correlation does not seem to affect the reliability updating effect with the posterior reliability including the random error still being slightly higher that without (Table 4.14), apparently because the prior reliability is also higher.

Table 4.14: *Case green: posterior reliability index (β) and probability of failure (p_f) for the phreatic level sensitivity analysis assuming a random error in the response of the phreatic surface to the lake water level.*

	$\beta[-]$	$p_f[-]$
Assessment T=0kPa | Observation T=0kPa	2.34	$9.6 \cdot 10^{-3}$
Assessment T=0kPa random err | Observation T=0kPa random err	2.52	$5.9 \cdot 10^{-3}$

4.4.2 Effect of high observed phreatic level

The effect of a higher observed and survived phreatic level is investigated in this section by assuming observed phreatic levels of +1.8m NAP and +2.0m NAP (below the crest), independent of the outside water level (e.g. as if caused by heavy precipitation without increased lake level). The corresponding fragility curves for assessment and observation are depicted in Figure 4.12.

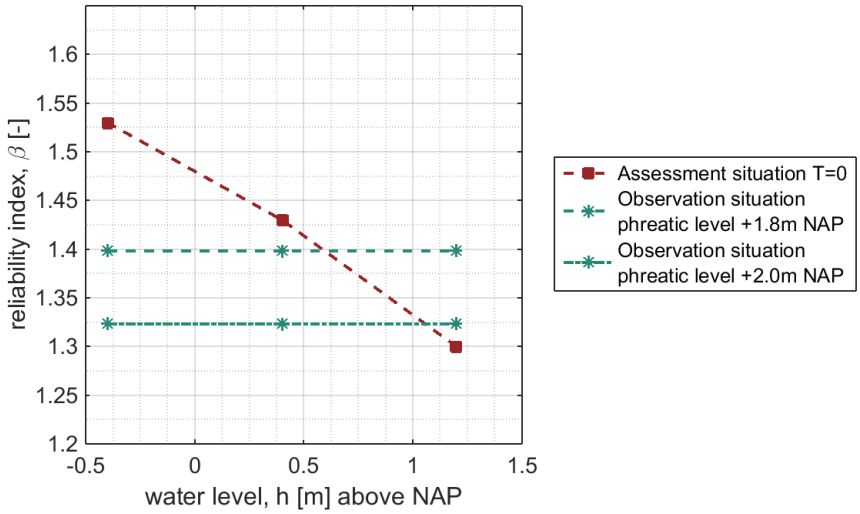


Figure 4.12: *Case green: fragility curves for the assessment and observation conditions with high observed phreatic level.*

The resulting posterior posterior reliability indices as summarized in Table Table 4.15 are

Table 4.15: *Case green: prior reliability index (β) and probability of failure (p_f) for the phreatic level sensitivity analysis.*

	$\beta[-]$	$p_f[-]$
Assessment T=0kPa | Observation T=0kPa phreatic level +1.8m NAP	3.21	$6.6 \cdot 10^{-4}$
Assessment T=0kPa | Observation T=0kPa phreatic level +2.0m NAP	4.74	$1.2 \cdot 10^{-6}$

significantly higher than the reference value of 2.3 obtained for reliability updating with the daily conditions (and zero traffic load). The reason for the high obtained reliability indices is that observing such high phreatic surface levels in the dike comes close to the worst credible loading condition in terms of phreatic surface. Note that the investigated levels are fictitious and merely give an indication of the effect if such phreatic levels had been monitored in the past, which to our knowledge they have not, or would be in the future.

4.5 Sensitivity to stratification (thicker clay layer)

Recent geological research into the Markermeerdijken subsoil conditions (Vos and De Vries, 2016) confirm the stratification as assumed for the base case in terms of the soil types and layer thicknesses for the specific area where case green is located. At the same time, the report confirms the presence of thicker clay layers in the Northern part of the Markermeerdijken. According to Figure 4.13, the highest level of the silty clay layer (in Dutch: Klei, siltig) layer is around -5 m NAP. In order to investigate the effects of the presence of such a thick clay layer on the prior and posterior reliability, the stratification has been amended as illustrated in Figure 4.14 by increasing the clay layer thickness accordingly. The the pore pressures in this variation are assumed hydrostatic from the phreatic level until the bottom of the peat layer; the remaining modeling assumptions (subsidence, parameters, etc.) are identical to the base case.

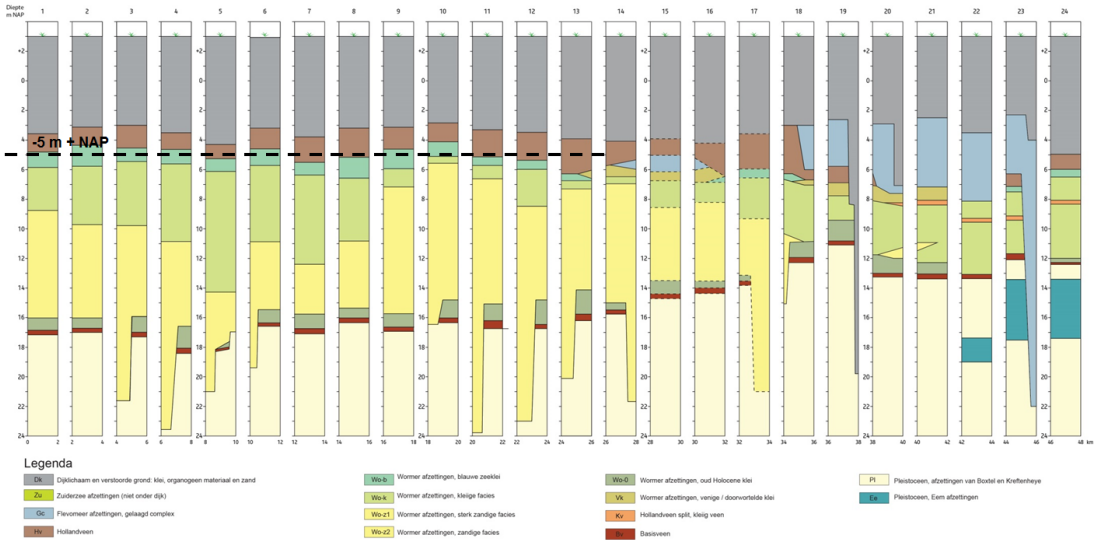


Figure 4.13: *Schematic stratigraphic profile columns (2 km) along the Markermeerdijk Hoorn-Amsterdam below the dike crest (Vos and De Vries, 2016).*

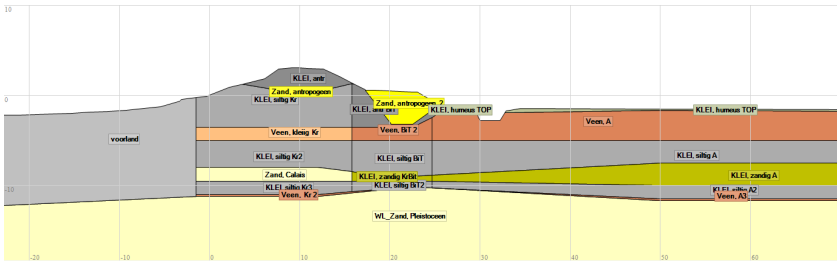


Figure 4.14: *Case green: adjusted soil schematisation by increased thickness of the Klei, Siltig layer.*

The conditional reliability indices for this adjusted schematisation are shown in Table 4.16. The corresponding fragility curves in Figure 4.15 show that the adjusted stratification leads to significantly lower reliability for both assessment and observation, which is plausible as the clay in the cross section has a lower shear strength than the peat which it replaced.

Table 4.16: *Case green: calculated reliability index (β) and probability of failure ($P(F)$) for different water levels for the adjusted soil schematisation - assessment.*

	assessment		observation	
h [m] above NAP	$\beta(h)$	$P(F\|h)$	$\beta(h)$	$P(F\|h)$
-0.40	0.25	$4.0 \cdot 10^{-1}$	0.85	$2.0 \cdot 10^{-1}$
0.00	0.12	$4.5 \cdot 10^{-1}$	0.74	$2.3 \cdot 10^{-1}$
0.40	0.01	$5.0 \cdot 10^{-1}$	0.61	$2.7 \cdot 10^{-1}$

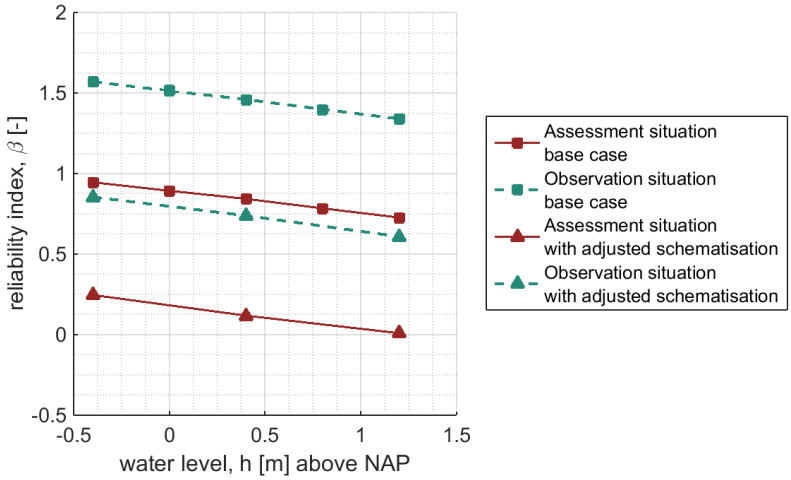


Figure 4.15: *Case green: fragility curves for the assessment and observation conditions for base case and the adusted soil schematisation.*

Taking into account the probability distribution of the lake water level results the prior reliability as summarized in Table 4.17, which is also lower than the reliability for the base case without traffic load, which was $\beta = 0.91$ (see Table 4.4).

Table 4.17: *Case green: prior reliability index (β) and probability of failure ($P(F)$) for the sensitivity analysis with thicker clay layer.*

Design point h[m] above NAP	$\beta[-]$	$p_f[-]$
-0.16	0.19	$4.2 \cdot 10^{-1}$

The posterior reliability is calculated for the observed average summer water level (($h^* = -0.2m$+NAP), using the fragility curves for the assessment and observation. The posterior reliability as summarized in Table 4.18 shows that the reliability updating effect is comparable to the base case, yet with a lower posterior reliability index (base case was $\beta = 1.14$, see Table Table 4.9) as also the prior reliability was lower. Hence, the contemplated variation does not change the conclusions drawn from the analysis hitherto regarding relative impact and applicability of reliability analysis and updating.

Table 4.18: *Case green: calculated posterior reliability index (β) and probability of failure ($P(F|\varepsilon)$) for the sensitivity analysis with thicker clay layer.*

h^* [m] above NAP	β	$P(F\|\varepsilon)$
-0.2	0.60	$2.7 \cdot 10^{-1}$

4.6 Prior and posterior analysis exact compared with fragility curves

The tables in this section present the prior and posterior results for both the exact Monte Carlo Simulation (MCS) directly carried out with the stability model and the approximation with fragility curves (FC) for the case green. Furthermore, for the prior analysis also results for Directional sampling (DS, also directly with the stability model) are presented. Also information on computation time is provided.

Prior reliability

The prior results (β_{prior}) are shown in the following tables: Table 4.19 for an assessment traffic load of T=13.3 kN/m^2 (base case) and Table 4.20 for T=0 kN/m^2 (i.e. assumption 1).

Table 4.19: *Case green: Comparison of the prior reliability estimate for different reliability methods, with a traffic load of T=13.3 kN/m^2.*

Method	reliability index β	number of computations	computation time
FC	0.91	672	< 30 min
MCS	1.01	2179	$\sim$ 1 h
DS	1.02	4658	$\sim$ 4 h

Table 4.20: *Case green: Comparison of the prior reliability estimate for different reliability methods, with a traffic load of T=0 kN/m^2.*

Method	reliability index β	number of computations	computation time
FC	1.50	756	< 30 min
MCS	1.51	5669	$\sim$ 2 h
DS	1.48	7448	$\sim$ 8 h

The results of the prior reliability index are in good agreement for the three methods.

Note that directional sampling is usually computationally much more efficient than MCS for problems with low probabilities of failure (say reliability indices of 3 and higher). The fact that DS does not outperform MCS here is due to the relatively high probability of failure and the relatively large number of random variables.

Posterior reliability

Tables 4.21 and 4.22 present the posterior results for different observed and survived water

levels. Table 4.21 considers different traffic load conditions for the assessment and observation (base case), while Table 4.22 considers the same traffic load conditions for the assessment and observation (i.e. both T=0 kN/m^2).

Table 4.21: *Case green: Comparison of the posterior reliability estimates with FC and MCS for different observed water levels (h^*) for the base case, i.e. assessment with T=13.3 kN/m^2 and observation with T=0 kN/m^2; $\beta|h^*$ is the posterior reliability index for survived water level h^*, n the number of computations and t the computation time*

	h^*= NAP -0.2			h^* = NAP +0.2			h^* = NAP +1.0		
Method	$\beta\|h^*$	n	t	$\beta\|h^*$	n	t	$\beta\|h^*$	n	t
FC	1.14	-	< 5 min	1.17	-	< 5 min	1.26	-	< 5 min
MCS	1.38	4515	$\sim$ 1.5 h	1.40	4739	$\sim$ 1.5 h	1.47	5475	$\sim$ 2 h

Table 4.22: *Case green: Comparison of the posterior reliability estimates with FC and MCS for different observed water levels (h^*) and no traffic load in assessment and observation (T=0 kN/m^2); $\beta|h^*$ is the posterior reliability index for survived water level h^*, n the number of computations and t the computation time.*

	h^*= NAP -0.2			h^* = NAP +0.2			h^* = NAP +1.0		
Method	$\beta\|h^*$	n	t	$\beta\|h^*$	n	t	$\beta\|h^*$	n	t
FC	2.51	-	< 5 min	3.39	-	< 5 min	5.36	-	< 5 min
MCS	2.77	146868	$\sim$ 2 d	3.27	162832	$\sim$ 4 d	- $^+$	-	-

$^+$) computations did not converge in the available time (would take months on a single PC)

Also the posterior results of the two methods, the approximation with fragility curves (FC) and the exact solution with MCS directly using the stability models are in good agreement. Note that there are several error sources in the approximation with fragility curves which may lead to differences: the interpolation of the beta-h, approximating the CDF of the resistance with the beta-h curve and the computation of the fragility curves (here) with FORM.

5 Case house: clay dike on peat with house

This chapter describes the test case 'case house' in detail. Case house is a clay dike on peat that includes a house in the cross section. The geometry, subsoil modeling choices, parameters and other choices are described. Analyses are made for both the assessment conditions as well as the observed conditions. For both, a base case is considered for which a deterministic analysis and fragility curve(s) are calculated. For a description of the derivation of fragility curves, reference is made to Appendix A. Similar to case green, the base case uses conservative assumptions for the observation and the assessment. For traffic, this means no traffic in the observation and traffic in the assessment. Since the house in case house has a positive effect on slope stability, the conservative choice used for the base case without house in the assessment and with house in the observation.

The first step is to compute the prior probability of failure based on the assessment fragility curve (section 5.1). This is the probability of failure without the effects of updating. Subsequently, the observation fragility curve is derived and reliability updating is performed for the base case (section 5.2). This results in a posterior failure probability that includes a water level observation.

Similar to case green in the previous chapter, also the traffic load influence is considered (section 5.3). The effect of the house is considered as well in this section as this is the main difference with case green. Subsequently, the influence of typically neglected resistance variables is analyses in section 5.4.

Finally, a summary of the results and more elaborate conclusions are presented in chapter 6.

5.1 Prior analysis

5.1.1 Assessment conditions (base case)

This section describes the starting points for the modeling choices of the assessment situation of this test case 'clay dike on peat with a house', *i.e.* 'case house'. For the general starting points and more details on the starting points, refer to section B.2, the case-specific situation is described below.

Geometry and soil layering

See starting points in section B.2. The 2D soil layering is shown in Figure 5.1 and is based on local CPTs (see Figure B.2). As can be seen in this figure, the house in this cross-section is not shown and only the gap the house leave is shown. The foundation of the house is at a depth of -0.76 m above NAP.

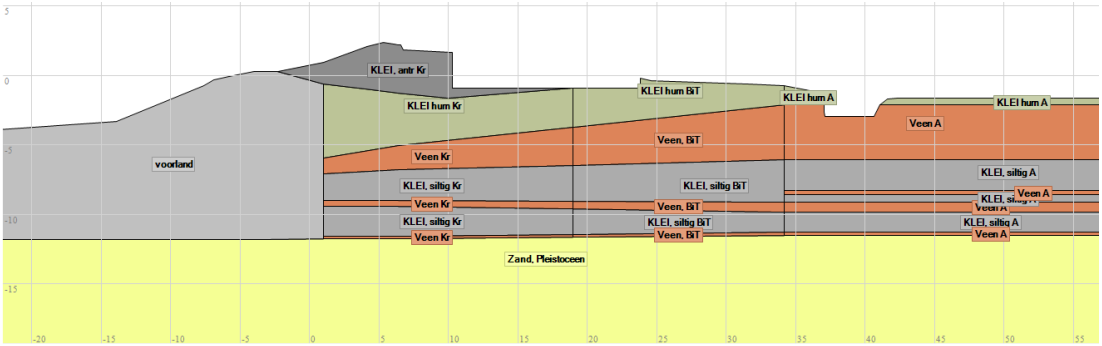


Figure 5.1: *Case house: Geometry and soil layering.*

To assess the slope stability of a dike with a building, it is common practice to assess a cross-section without the building. If the stability is guaranteed for this cross section, the situation with the building will be too, as long as the building does not contribute actively to slope instability. In case house, the building is mostly present at the resisting part of the slip plane of slope stability, hence, the weight of the house is contributing to stability. In this report, only deep slip circles are considered. Local slip circles that would occur due to the gap of the house are not included.

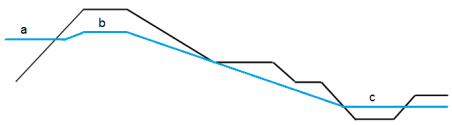
Geohydrological

The phreatic line is modeled at locations (a), (b) and (c), as shown in Table 5.1. Since the base case has a lowered surface level due to subsidence (similar as case green), the polder water level is also lowered to the same extend. This is taken into account in Table 5.1. The difference, between the outer water level (a) and the phreatic level in the dike body (b), varies and is considered for different water levels using: $b = 1.013 + 0.406a$[1].

During daily conditions, the stationary head level in the first aquifer layer is measured at NAP -2.08 m (PL2), Zwanenburg (2014b). The leakage lengths are using the bathymetric map of the IJsselmeer area and the measured response factor in head level (mentioned in section B.2).

The values of the leakage lengths are taken as $\lambda_{in} = 3200$ m and $\lambda_{out} = 4800$ m. For more details about the geohydrological modeling, refer to section B.2.

Table 5.1: *Case house: schematization of the phreatic level for two arbitrary water levels, assessment conditions*



Phreatic level in [m] above NAP, per location			
example water level h	a	b	c
NAP-0.40m	-0.40	+0.85	-1.90
NAP+1.15m	+1.15	+1.48	-1.90

[1] Here, a and b are the water level and the phreatic line level in the dike body respectively, and a linear interpolation is implemented between the two values shown in Table 4.1. The underlying assumption is that high lake water levels are practically perfectly correlated with pro-longed precipitation events affecting the phreatic surface by precipitation (for details refer to section B.2).

5.1.2 Prior analysis (base case)

Deterministic analysis

In order to review the current schematisation, two analyses are made. The check has the purpose to verify that computations are running smoothly and to identify and verify the behaviour of the dike (i.e. stability factor) for low and high water levels. The first analysis shows the calculated stability factor (SF) with mean values of all the uncertain parameters. The second shows the calculated SF with characteristic low values (i.e. 5%-percentile values and no material factors applied [2]) for the strength parameters. The SF is calculated for a high and low outside water level. The stability factors are shown in Table 5.2 and the slip planes is shown in Figure 5.2 and Figure 5.3.

It is important to mention that the current study was restricted to large slip planes and did not contemplate local instabilities near the house. The considered slip plane has a large length which means that relatively much shear strength can be mobilized.

The computed safety factors are high for mean values. There is a limited difference in SF between the considered water levels. Also, the difference in SF between SF based on mean and characteristic values is large, indicating a large contribution of shear strength properties.

Table 5.2: *Case house: calculated stability factor for mean and characteristic values - base case assessment.*

Example water level h	SF Mean values	SF Characteristic values	SF_{char}/γ_d*
NAP-0.40m	1.83	1.08	1.02
NAP+1.15m	1.78	1.06	1.00

*) γ_d =1.06

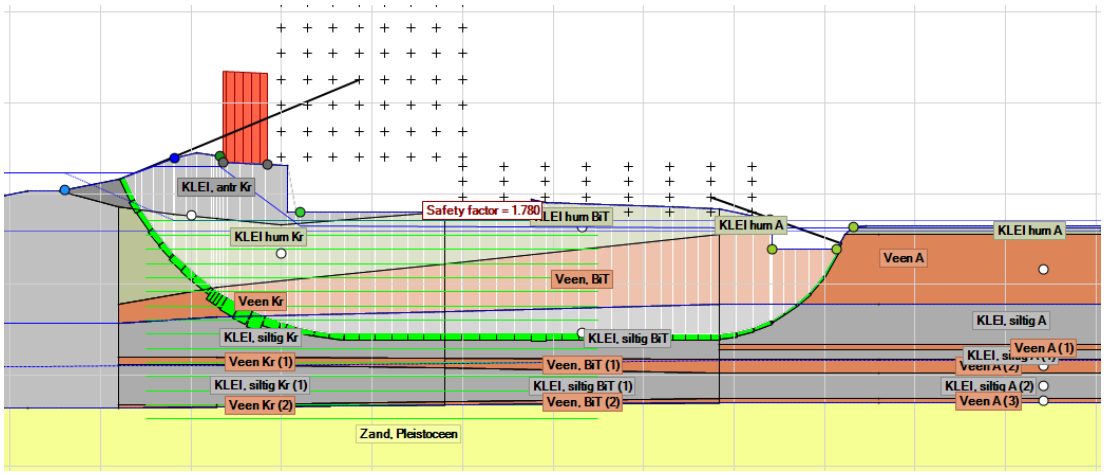


Figure 5.2: *Case house: Stability factor with mean values for an example water level.*

[2] This is in line with the lastest recommendations for WBI2017 - Kanning *et al.* (2015)

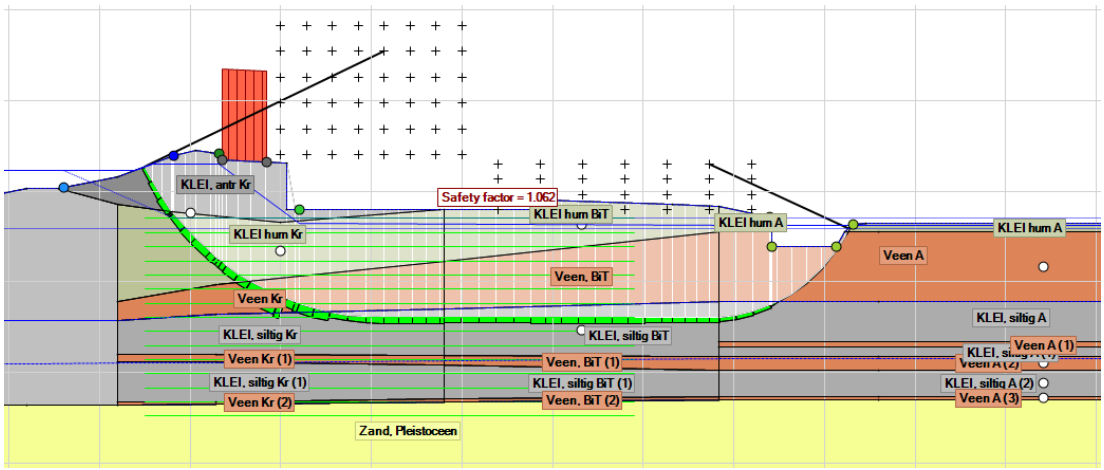


Figure 5.3: *Case house: Stability factor with characteristic values for an example water level.*

Fragility curve

Using FORM, probabilistic calculations are made for fixed water levels. A relevant range of NAP-0.40m until NAP+1.20m is chosen. The reliability indices and failure probabilities conditional to the respective outside water level are shown in Table 5.3. The conditional failure probability as a function of water level is the fragility curve.

Table 5.3: *Case house: calculated reliability index (β) and probability of failure (p_f) for different outside water levels - base case assessment.*

h[m] above NAP	$\beta \vert h[-]$	$p_f \vert h[-]$
-0.40	5.65	$8.0 \cdot 10^{-9}$
$+0.40$	5.57	$1.3 \cdot 10^{-8}$
$+1.20$	5.44	$2.7 \cdot 10^{-8}$

Prior analysis

Using the fragility curve above and the (Gumbel) distribution of the outside water level (Figure B.4), the prior probability of failure is calculated using integration (Schweckendiek and Kanning, 2016). The result is shown in Table 5.4. The prior reliability index is much higher than case green (also the SF is much higher for case house than case green). This may be explained due to the longer berm due to the house of the house and the corresponding larger slip circle.

The resulting influence coefficients (α^2) are shown in Table 5.5 and the values of the variables in the design point in appendix F.1. The α^2 values show the contribution of each uncertain variable to the failure probability. For case house, the dominant variables to the failure probability are shear strength related (S-ratio, Yield stress) and model uncertainty. The contribution of water level uncertainty is negligible. This is also confirmed by the design point of the water level that is close to the mean. The slip plane at the design point is presented in Figure 5.4.

Table 5.4: *Case house: calculated prior reliability index (β) and probability of failure (p_f) - base case assessment.*

Design point h[NAP .. m]	β	p_f
-0.15	5.62	$9.5 \cdot 10^{-9}$

Table 5.5: *Case house: calculated influence coeficients (α^2) in the design point - base case assessment.*

Variable	α^2
S-ratio	0.59
m-exponent	0.02
Yield Stresses	0.24
Friction angle	0.00
Pore water pressures	0.00
Model uncertainty	0.14
Outside water level	0.00
Total	1.00

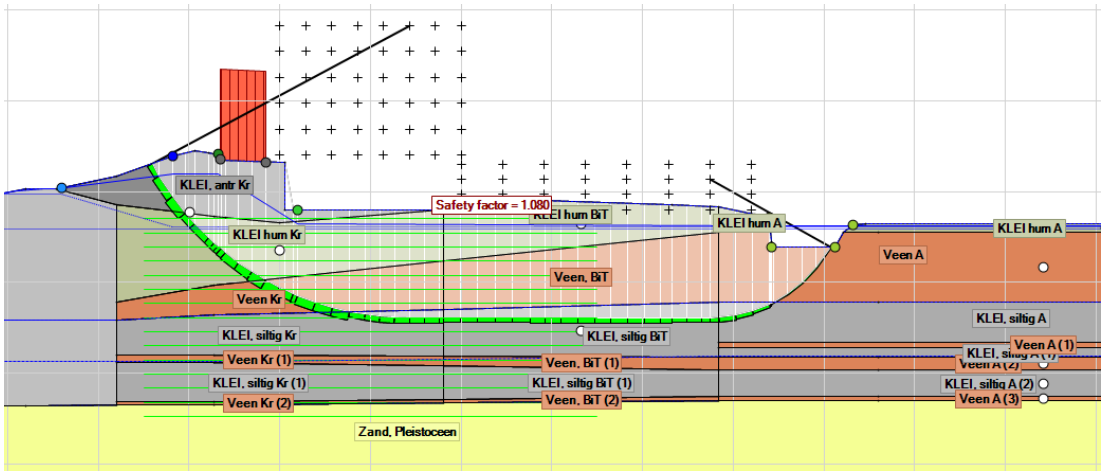


Figure 5.4: *Case house: Stability factor in the design point equals the value of the model factor in the design point, since $g = SF \cdot m_d - 1 = 0$.*

5.2 Reliability updating

5.2.1 Observation conditions (base case)

This section describes the assumptions for the modeling of the observation situation of the base case of 'case house'; see also section B.3. There are the following differences between the assessment and the observations:

- **Geometry**: There is 0.1 m less subsidence in the observation as compared to the assessment. This is reflected in a 0.1 m thicker peat layer 'Veen A'.
- **Geohydrology**: The only difference is a 0.1 m higher polder level (c) in the observation that reflects the difference in subsidence between the observation and assessment. See also section B.4.4.
- **Traffic**: In the base case, there is no traffic assumed in the observation; opposed to the 13.3 kN/m^2 traffic load in the assessment.
- **House load** The house load is modeled as present during the observation, see appendix F.3. The load is modeled as a permanent uniform load of H = 17 kN/m^2. The house load is not incorporated in the base case in the assessment since this is a conservative choice.

The remaining parameters are the same for observation and assessment conditions.

5.2.2 Fragility curve for observation conditions (base case)

This section discusses the fragility curve of the observation conditions. Note that this is excluding the observed water level. Reliability updating with various observed water levels is discussed in the next section.

Deterministic analysis

In order to analyze the difference, in terms of stability factor (SF), between the assessment and the observation conditions, four computations are made. Two with mean values and two with characteristic low values (*i.e.* 5%-percentile) - see Table 5.6. Also the design value of the SF is computed, which is the SF given characteristic values divided by the model factor (γ_d). There is around 0.2 difference in SF between the observation and assessment conditions (compare Tables 5.2 and 5.6).

Table 5.6: *Case house: calculated stability factor for mean and characteristic values - base case observation.*

Example water level h	SF Mean values	SF Characteristic values	SF_{char}/γ_d*
NAP-0.40m	2.01	1.24	1.17
NAP+1.15m	1.95	1.22	1.15

*) γ_d =1.06

Fragility curves base case

Using FORM, probabilistic calculations are made for fixed water levels. A relevant range of -0.40 m until +1.20 m above NAP is chosen. The reliability indices conditional to the respective outside water level are shown in Table 5.7. The beta-h curve (equivalent to the fragility curve that is the conditional failure probability as function of water level curve) is shown in Figure 5.5. Similar to case green, the observation curve and the assessment curve are separated (due to i.e. traffic load) and show limited sensitivity to the outside water level.

Table 5.7: *Case house: calculated reliability index (β) and probability of failure (p_f) for different outside water levels - base case observation.*

h [m] above NAP	$\beta \mid h$ $[-]$	$p_f \mid h$ $[-]$
-0.40	7.07	$7.7 \cdot 10^{-13}$
$+0.40$	6.88	$3.0 \cdot 10^{-12}$
$+1.20$	6.62	$1.8 \cdot 10^{-11}$

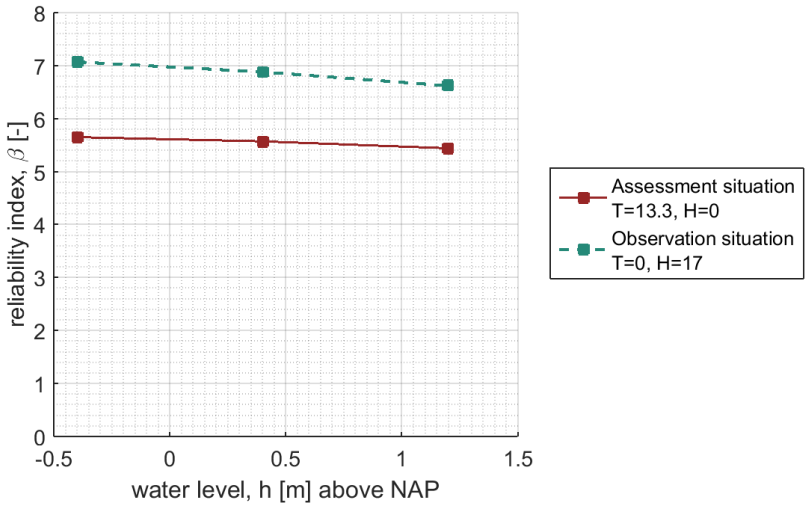


Figure 5.5: *Case house: fragility curves for base case assessment and observation.*

5.2.3 Correlation between assessment and observation

As refereed for the case green, using the information derived from the assessment and observation fragility curves and the auto-correlation in time of the individual basic random variables, we can estimate the correlation between the assessment and observation conditions, i.e. ρ - see eq.(4.1).

Table 5.8 presents the FORM influence coefficients for the case house, in the assessment and the observations conditions are presented. The estimation of the correlation between assessment and observation (ρ) is then presented in Figure 5.6.

Table 5.8: *Case house: FORM influence coefficients (α) for the base case assessment and also observation situations.*

	α_i^f				α_i^p					
Variable i	h=-0.4	h=0.4	h=1.2	average	h=-0.4	h=0.4	h=1.2	average	$\rho_i^{p,f}$	$\alpha_i^p \alpha_i^f \rho_i^{p,f}$
S-ratio	0.768	0.765	0.760	0.764	0.773	0.772	0.783	0.776	1.0	0.593
m-exponent	0.158	0.165	0.174	0.166	0.145	0.151	0.151	0.149	1.0	0.025
Yield Stresses	0.496	0.499	0.501	0.498	0.487	0.485	0.475	0.482	1.0	0.240
Friction angle	0	0	0	0	0	0	0	0	1.0	0
Pore water pressures	0.021	0.026	0.049	0.032	0.037	0.043	0.020	0.033	1.0	0.001
Model uncertainty	0.373	0.372	0.374	0.373	0.378	0.380	0.372	0.377	1.0	0.140
									$\sum_i =$	0.999

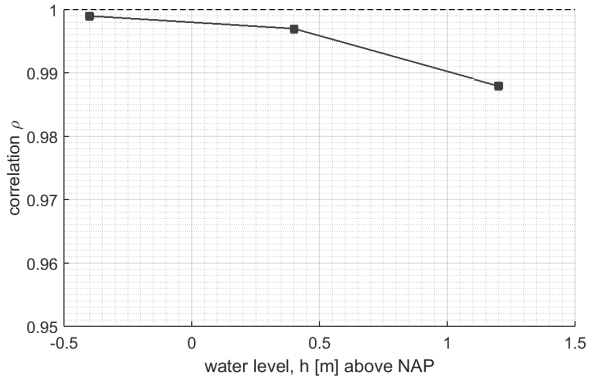


Figure 5.6: *Case house: correlation coefficient achieved based on FORM influence coefficients and individual correlation, as in eq.(4.1).*

The correlation changes very marginally with water level, and for case house, the consideration of full correlation between the assessment and observation fairly approximates the coefficient of correlation found and shown in Figure 5.6. The complete table of the random variables can be seen in appendix, section F.2.

5.2.4 Reliability updating (base case)

Using the fragility curves for the assessment and observation, the posterior probability of failure is calculated (Schweckendiek and Kanning, 2016). This is done for the same three water levels that are considered observed as case green. The result is shown in Table 5.9 for different considered observed water levels (h^*).

Table 5.9: *Case house: calculated posterior reliability index (β) and probability of failure (p_f) for the base case for three considered observed water levels (h^*).*

h^* [m] above NAP	$\beta \vert h^* \quad [-]$	$p_f \vert h^* \quad [-]$
-0.2	5.62	$9.38 \cdot 10^{-9}$
$+0.2$	5.62	$9.38 \cdot 10^{-9}$
$+1.0$	5.62	$9.39 \cdot 10^{-9}$

The decrease in probability of failure is insignificant, independent of the considered survived water level. The main reasons are the low prior failure probability and the large difference between the assessment and observed conditions. The limited updating effect for cases with high prior reliability (low failure probability) is not surprising. Survival of such a structure does not provide significant new information, as we expected the dike to be stable anyway with high probability.

5.3 Sensitivity to traffic load and building influence

5.3.1 Influence traffic load and building on prior reliability

The main difference between the assessment and the observation are (1) the presence of traffic load in the assessment condition, while this was not assumed present in the base case during the observation; and (2) the presence of the house load in the observation and the absence of a house load in the assessment. Therefore, both the sensitivity of the reliability index to the outside water level, the traffic load and the house load is investigated, see Table 5.10. The reliability index β as a function of the water level and traffic and house load is shown in Figure 5.7. Similar to case green, the traffic load has more influence on the reliability than the outside water level. For the house load, it can be seen that the house load has a positive contribution to the stability as an increased house load corresponds to a higher reliability. Also, β is more sensitive to changes in the house load than to changes in the outside water leve.

Table 5.10: *Case house: reliability index condtional to water level (h), traffic load (T) and house load (H) - assessment siutation*

	Assessment H = 0, $\beta\|h$			Assessment T = 13.3, $\beta\|h$		
h [m] above NAP	T = 0	T = 7	T = 13.3	H = 0	H = 17	H = 34
-0.4	6.29	5.96	5.65	5.65	6.42	6.35
0.4	6.17	5.88	5.57	5.57	6.31	6.30
1.2	5.84	5.61	5.44	5.44	6.09	6.24

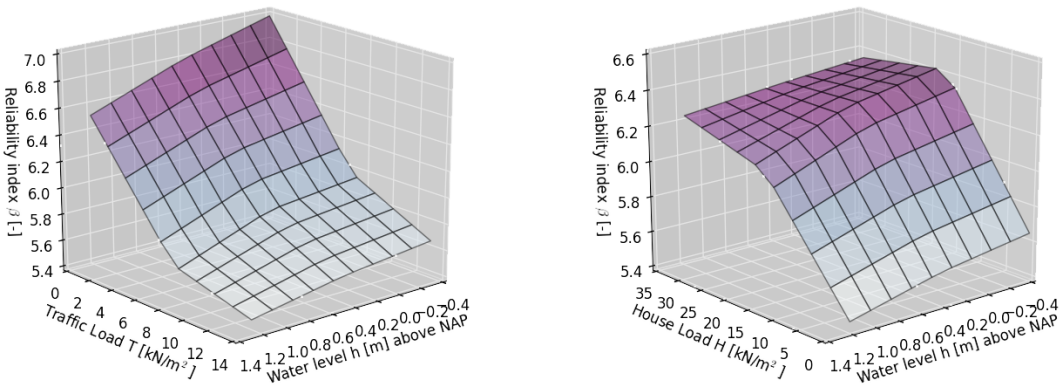


Figure 5.7: *Case house: reliability index conditional to traffic load and water level (left) and reliability index conditional to house load and water level (right) - assessment condition*

In general, if the assessment and observation conditions are closer together in terms of probabilities of failure conditional to the water level, it is expected that an observation will have a larger effect on the posterior reliability (Schweckendiek and Kanning, 2016). This is further investigated in the next sections by computing the updating effect in case traffic and house loads are considered the same in assessment and observation.

The starting points in the base case concerning the house weight are conservative (section B.3) as we obtain a lower bound for the posterior reliability index. Hence, in Assumption 1 the house load is considered in both the assessment and observation conditions while the traffic load is taken only in the assessment condition, as in the base case (see results in section 5.3.2). Meanwhile Assumption 2 considers both the house load and the traffic load in both the assessment and observation condition (see results in section 5.3.3).

5.3.2 Assumption 1: House load in the assessment

Assumption 1 considers conditions in which the house load is not only present in the observation, but also in the assessment, while in the base case it was only considered in the observation. The traffic is here assumed only present in the assessment, and no observed traffic load is accounted for. The presence of the weight of the building in both assessment and observation conditions seems the more sensible assumption as opposed to the conservative assumption in the based case. For example, even if a house collapses during a storm, the weight would still be there.

Assumption 1 is presented below, where changed values in comparison with base case are indicated **bold**:

- Assessment with traffic load (T = 13.3 kN/m^2) and house load (**H = 17 kN/m^2**)
- Observation without traffic load (T = 0 kN/m^2) and house load (H = 17 kN/m^2)

The fragility curves for assessment and observation conditions are shown in Figure 5.8. The reliability of the assessment conditions is higher than for the base case, because the house weight has a positive effect on the stability (see previous section). Consequently, the beta-h curve has moved upwards, and the prior probability of failure decreased. The prior results are shown in Table 5.11; the mean value of the water level in the design point emphasizes the limited influence of the water level also for this case.

The assessment and observation are closer to each other than in the base case (Figure 5.5). Table 5.13 illustrates that there is slightly more effect of reliability updating compared to the base case, yet it is still limited due to the low prior failure probability and the remaining difference between assessment and observation.

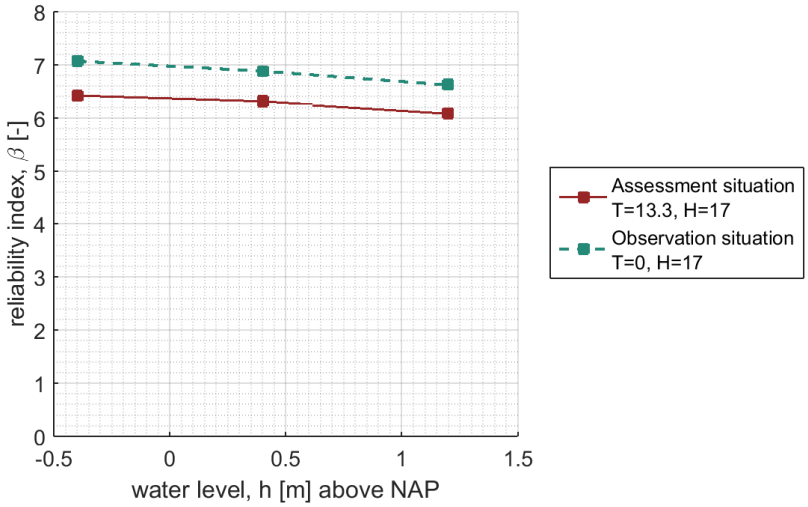


Figure 5.8: *Case house - Assumption 1: fragility curves for assessment condition and observation*

Table 5.11: *Case house - Assumption 1: calculated prior reliability index (β) and probability of failure (p_f) base case assessment.*

Design point h [m] above NAP	β	p_f
-0.15	6.38	$8.7 \cdot 10^{-9}$

Table 5.12: *Case house - Assumption 1: calculated posterior reliability index (β) and probability of failure (p_f) conditional to 3 considered observed water levels (h^*).*

h^*[m] above NAP	$\beta \mid h^* \quad [-]$	$p_f \mid h^* \quad [-]$
-0.2	6.38	$8.6 \cdot 10^{-9}$
$+0.2$	6.39	$8.5 \cdot 10^{-9}$
$+1.0$	6.40	$7.6 \cdot 10^{-9}$

5.3.3 Assumption 2: House and traffic load in both observation and assessment

Assumption 2 includes the situation where both the traffic and house load are the same in the observation and assessment. Hence, the only difference between the two is subsidence resulting in geometrical differences. Assumption 2 is presented below, the changed values in comparison with base case are indicated in **bold**:

- Assessment with traffic load (T = 13 kN/m^2) and house load (**H = 17 kN/m^2**)
- Observation case with traffic load (**T = 13 kN/m^2**) and house load (H = 17 kN/m^2)

The two fragility curves are shown in Figure 5.9. The assessment situation is the same as in Assumption 1, so the prior fragility curve is the same as in Figure 5.8. Because the house weight increases the stability, the fragility curve moved downwards, bringing the two fragility curves closer together (Figure 5.9). Now the observation and assessment conditions are very similar and the only remaining difference between the beta-h curves is caused by differences in subsidence. This is reflected in the more significant reliability updating effect as shown in Table 5.13.

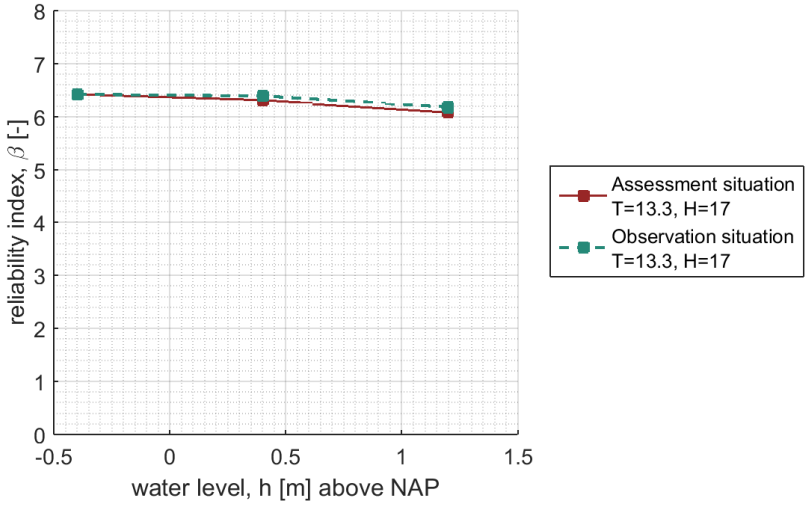


Figure 5.9: *Case house - Assumption 2: fragility curves for assessment conditions and observation*

Table 5.13: *Case house - Assumption 2: calculated posterior reliability index (β) and probability of failure (p_f) conditional to 3 considered observed water levels (h^*).*

h^* [m] above NAP	$\beta \mid h^* \quad [-]$	$p_f \mid h^* \quad [-]$
-0.2	6.64	$1.53 \cdot 10^{-9}$
$+0.2$	6.73	$8.30 \cdot 10^{-10}$
$+1.0$	7.86	$1.90 \cdot 10^{-13}$

5.4 Typically neglected resistance contributions

There are several phenomena that contribute to the stability which, however, are typically neglected in standard assessments. These could be, for example, effects of the three-dimensional shape of the failure plane beyond what is covered by the model uncertainty (m_d), such as access ramps or other objects in the dike geometry, or the stabilizing effect of pile foundations comparable to soil nailing. In standard assessments these and other aspects are often neglected as conservative assumption, as often no appropriate and practicable models are available. However, as discussed in Schweckendiek and Kanning (2016), for the observation conditions neglecting these aspects is not conservative and may lead to over-estimation of the posterior reliability.

To explore these effects, an analysis is made to quantify the effect of typically neglected resistance. As an example, the effect of unaccounted 3D effects is considered. Note that regular 3D effects are supposed to be covered in the model uncertainty. If we assume there is 10 percent unaccounted resistance, this means e.g. that the 2D stability factor has to comply to a minimum stability factor of 0.9. The limit state function in this case is then $g = SF \cdot m_d - 0.9$. For probabilistic calculations, this generally means that the reliability increases. The effects of the unaccounted 3D resistance on the beta-h curves is shown in Figure 5.10. The base case (see section 5.1) observation and assessment are considered as the starting point and adapted for the unaccounted 3D resistance. Due to the additional 3D effect, the prior reliability increases, see Table 5.14. The reliability updating is shown in Table 5.15. In both cases, there is such a difference between observation and assessment, that no reliability updating effect is visible.

This section shows that, in principle, it is possible to incorporate additional resistance contributions. Based on the simplified assumptions for sensitivity analysis here, we have shown that both the prior and the posterior reliability increase. Note however, that this sensitivity analysis is merely meant to obtain an impression of relative effects. Real-life applications will require more tailored modeling of the local conditions.

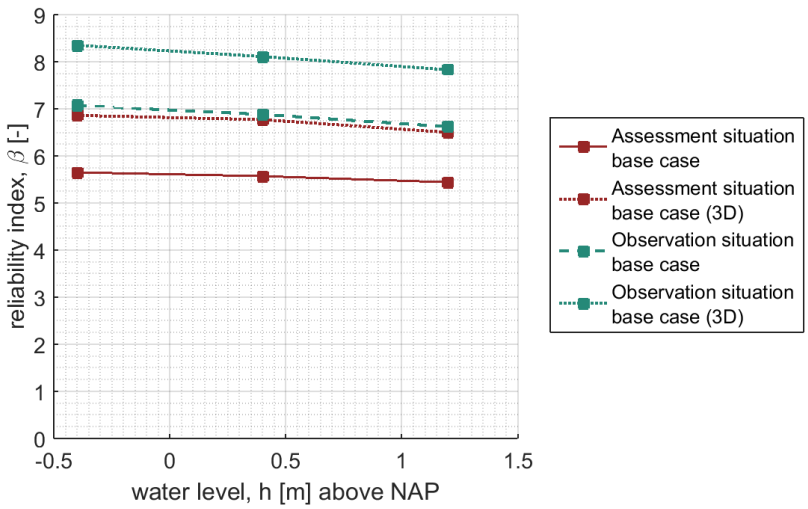


Figure 5.10: *Case house: incorporation of unaccounted 3D effect in fragility curves for assessment and observation conditions*

Table 5.14: *Case house: incorporation of unaccounted 3D effect in calculated prior reliability index (β) and probability of failure (p_f).*

Design point h[m] above NAP	β	p_f
-0.15	6.83	$4.3 \cdot 10^{-10}$

Table 5.15: *Case house: incorporation of unaccounted 3D effect in calculated posterior reliability index (β) and probability of failure (p_f) for base case assessment (3D) and observation (3D)*

h^* [m] above NAP	$\beta \vert h^* \quad [-]$	$p_f \vert h^* \quad [-]$
-0.2	6.83	$4.3 \cdot 10^{-10}$
$+0.2$	6.83	$4.3 \cdot 10^{-10}$
$+1.0$	6.83	$4.3 \cdot 10^{-10}$

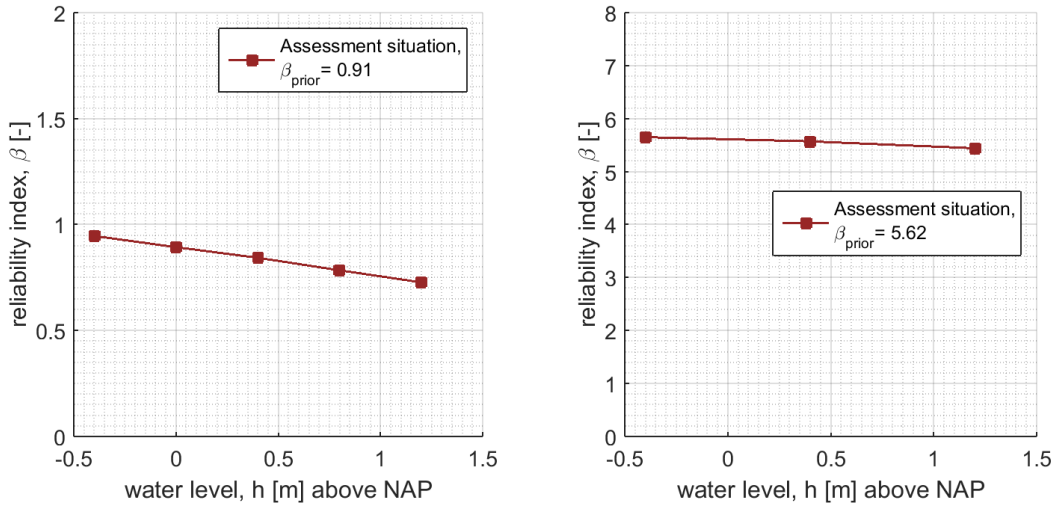
Note that the present sensitivity analysis is case-specific and cannot be easily generalized. Also, the modeling assumptions made merely serve the purpose to obtain qualitative insights into the effects in three-dimensional situations. Application to real cases will require more tailored treatment of the local conditions.

6 Discussion

This chapter summarizes the main observations for the two test cases elaborated in chapter 4 and 5 and provides additional considerations which are not necessarily directly related to the contemplated cases. For the overall conclusions and recommendations refer to chapter 1.

6.1 Results of prior analyses

Figure 6.1 summarizes the prior reliability results of 'case green' and 'case house'.



***(a)** 'case green'* ***(b)** 'case house'*

***Figure 6.1:** Fragility curves for base case assessment situation.*

The main observations are:

- 'Case green' has a relatively low prior reliability index, the probability of failure being dominated by uncertainties in the shear strength properties of the soft soil layers.
- 'Case house', on the other hand, has a relatively high prior reliability index, though the soil profile is similar to 'case green'. It must be noted that the current study was restricted to large slip planes and did not contemplate local instabilities near the house. The considered slip plane has a relatively large length which means that relatively high shear strength can be mobilized.
- The main differences between the two cases are the length of the critical slip plane (mainly due to the difference in berm geometry) and the level of the phreatic surface. The large differences between apparently similar profiles confirm the sensitivity of the stability to rather minor variations in the profile and pore water pressures at very low stress levels.
- In both cases, the stability factor and reliability index shows very low sensitivity to the water level. This is also reflected by the relatively flat fragility curves. This implies that the remaining uncertainties are dominant (i.e. resistance of the dike).

6.2 Context of the reliability index

The results of the probabilistic analyses can be placed in the context of the semi-probabilistic safety assessment as envisaged in WBI2017 (Kanning *et al.*, 2015). In order to derive acceptable factors of safety, a relation between the safety factor (with characteristic values) and reliability index was 'calibrated'. Figure 6.2 illustrates that for roughly the same safety factor a wide range of reliability indices can be expected. Hence, a tailored probabilistic analysis can avoid the conservatism necessarily introduced by a semi-probabilistic assessment criterion (which needs to be applicable for a wide range of conditions). 'Case green' falls quite close by to the relation derived for WBI2017, though it lies outside the range of relevant reliability index for Dutch dikes (grey outlined box). On the other hand, the safety factor computed for 'case house' of roughly 1.0 would imply a reliability index of at least roughly 3.5 according to the calibrated relation (red line). Yet the actual reliability index obtained by probabilistic analysis was 5.6, which is at the high end of the scatter. However, both cases appear to have a relation between safety factor and reliability index that is in line with other, not-related cases.

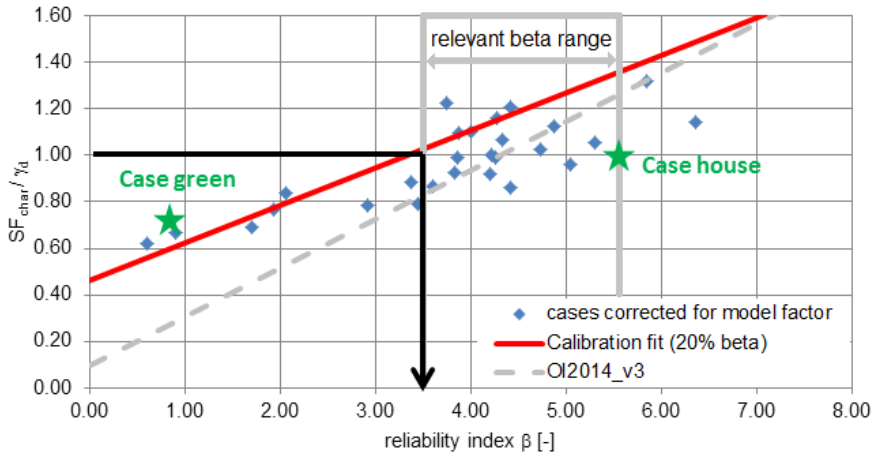


Figure 6.2: *Prior reliability index and stability factor SF_{char}/γ_d of the two test cases compared with WBI-2017 preliminary semi-probabilistic safety assessment (Kanning* et al.*, 2015).*

The implication is that dikes found to be unsafe based on a factor of safety can actually be safe in terms of the acceptable probability of failure.

6.3 Reliability updating

The reliability updating analyses of both the test cases and of the examples in Schweckendiek and Kanning (2016) have highlighted the following aspects:

6.3.1 Expected effect of reliability updating

Reliability updating with survival observation has a significant effect in terms of reducing the probability of failure, if:

a) a significant load or load effect has been survived,

b) the probability of failure is relatively high and dominated by epistemic (knowledge) uncertainties (typically soil properties),

c) the structure has not changed or degraded substantially since the observation.

For item c) we observe that, especially for rather flat fragility curves, relatively small differences between the fragility curves can lead to the reliability updating effect being small.

6.3.2 Traffic load

The traffic load has a large influence in both test cases. Assuming different loads for the assessment and the observation situations leads to a difference between the fragility curves. In combination with a flat fragility curve, this results in a relatively limited effect of reliability updating. The higher the observed traffic load with respect to the one assumed for the assessment conditions, the greater the effect on the posterior reliability.

The effect of traffic loads is commonly treated as deterministic in standard design and assessment analyses. Yet the real effects, especially in terms of the stress changes in the dike body, are rather uncertain. For an accurate (i.e. best estimate) reliability updating analysis it is desirable to treat the traffic load modeling alike in the assessment and the observation conditions. That does not mean the same value needs to be assumed, but that the modeling of the load effects should be the same and preferably the uncertainties treated explicitly. When it is not possible to model the traffic load and the related uncertainties accurately, the assumptions in assessment and observation conditions may differ to produce cautious estimates of the posterior reliability.

Arguably, at least some traffic load has occurred in the observation conditions, especially when we consider the daily conditions for the observation and not rare, extreme events. Considering the sensitivity of the posterior reliability to this parameter, it seems worthwhile to further investigate this aspect.

6.3.3 Aspects relevant with buildings

The load of a house by its dead weight (with shallow footing) was a consideration in 'case house'. In the base case for the assessment the weight was not considered as a conservative starting point, since the load effect increases the stability of the dike. Similarly, the weight of the house was considered in the observation conditions as conservative assumption. As for the traffic load, it is questionable whether this sort of conservatism is adequate for the probabilistic analyses carried out in this study, as reliability analysis ought to be as unbiased as possible and not "conservative". This is especially true for the difference in treatment between assessment and observation, as the weight of the house will not likely disappear.

There are various aspects of foundations that may influence a RUPP analysis. For instance, pile foundations of houses can increase the stability similar to soil nailing. The effect of piles on the stability was considered in a sensitivity analysis using simplifications for the increase of shear strength in the pile zone. The increase in shear strength appears rather insignificant for the contemplated case.

Note that these conclusions are case-specific and cannot be generalized. In real life cases with buildings or three-dimensional conditions, probably more sophisticated modeling tailored to the local conditions will be necessary.

6.4 Comparison TRAS method

The technical report "Technisch Rapport Actuele Sterkte (TRAS)" (ENW, 2009) describes an alternative method to calculate the posterior probability of failure, in the remainder called "TRAS method". The method is based on the same theory of Bayesian updating as underlying the present elaboration of the case studies and described in the background report (Schweckendiek and Kanning, 2016), yet it applies an additional (first-order) approximation in combining the assessment and the observation limit state. We have applied the TRAS approach (described in section 6.3 of ENW (2009)) to *case green* and compared the results with the proposed approximation with fragility curves, as well as with the results obtained by directly applying Monte Carlo simulation (MCS). The essential conclusions are presented below, for details refer to appendix G.

The general observation is that the TRAS method can severely over-estimate the posterior reliability, if we use the MCS results as reference, while the approach with fragility curves (FC) appears to give reasonably accurate results for the considered cases. The over-estimation is due to the fact that the TRAS approach does not include the possibility to accommodate (deterministic) differences between assessment conditions and the observation conditions. For the presented case, the differences are caused by the geometry differences due to subsidence and the consequentially lowered phreatic level on the polder side. The TRAS method actually requires the dike to be the same in assessment and observation conditions. The present case study of case green has shown that apparently minor differences, which cannot be accounted for with the TRAS method in a straightforward way, can have a significant impact the posterior reliability.

This calls into question the general validity of the TRAS method, or at least a clarification of the range of conditions where the approach is sufficiently accurate or valid. For conditions with little influence of the water level leading to a high correlation between assessment and observation (in the absence of other aleatory factors) and differences between assessment and observation, this does not seem to be the case. We recommend to at least produce a warning for practitioners that the TRAS method is not applicable in such cases, and ideally to give guidance for when and how it can be applied successfully.

6.5 Recommendations for the test cases

The following items are suggested as potential follow-up analyses for the test cases:

1. **Rainfall observations**: Observations of prolonged rainfall events have not yet been considered explicitly as survived load conditions. This may be interesting especially in combination with other loads such as the water level or traffic loads.
2. **Geometry and subsoil**: The two considered test cases have very similar geometry profiles and subsoil conditions. Though the relative effects are not expected to be different from a theoretical point of view, analyzing other profiles with differing ground conditions may provide additional insights.
3. **Slip planes relevant for inundation**: The relatively large freeboard in the two case studies suggests that occurrence of the critical slip planes will not necessarily lead to inundation, especially not for low water levels, as the residual profile often will be sufficient to retain the water. Including such residual profiles and the conditional probabilities of inundation given failure of the critical sliding plane can substantially increase the reliability estimates, both prior and posterior.

For the general conclusions and recommendations refer to chapter 1.

References

AHN, 2015. *Handleiding "AHN downloaden van PDOK"*. Actueel Hoogtebestand Nederland.

Calle, E., 2005. “Bewezen Sterkte bij Dijken.” *Geotechniek* 1.

Calle, E. and W. Kanning, 2013. *WTI: Effecten ruimtelijke variabiliteit op geotechnische sterkte van waterkeringen*. Report 1207805-004-ZWS-005, Deltares.

Deltares, 2016. *D-Geo Stability, User Manual. Version 16.1.1.*

ENW, 2009. *Technisch rapport actuele sterkte van dijken*. Tech. rep.

Halter, W., 2008. *Geotechnische begeleiding bij planstudie dijkversterking Markermeerdijk Edam - Amsterdam. Geohydrologische beschouwing.* Tech. rep., Fugro Ingenieursbureau B.V.

Halter, W. and B. Effing, 2011. *Geotechnische begeleiding bij planstudie dijkversterking Markermeerdijk, Hoorn - Edam - Amsterdam, Ontwerpuitgangspunten*. Report 1211-0036-000, Fugro Geoservices B.V.

Halter, W., M. Hinborch, A. Aparicio Sáez and C. Briele, 2015. *Versterking Markermeerdijk Hoorn - Amsterdam. Achtergrondrapport Grondparameters Dijken op Veen*. Report 1214-0007-040, Fugro Geoservices B.V. report 1214-0007-040.

Hasofer, A. and N. Lind, 1974. “An exact and invariant first-order reliability format.” *Journal of Engineering Mechanics Division, ASCE* 100(1): 111-121.

Inspectie Verkeer en Waterstaat, 2011. *Derde toets primaire waterkeringen*. Tech. rep., Inspectie Verkeer en Waterstaat, Ministerie Infrastructuur en Milieu.

Kanning, W., 2016. “Pore water pressure uncertainties for slope stability.” Deltares memo 1230090-034-GEO-0008.

Kanning, W., M. Huber, M. Van der Krogt, T. Schweckendiek and A. Teixeira, 2015. *Derivation of the semi-probabilistic safety assessment rule for inner slope stability.* Tech. rep., Deltares report 1220080-003-ZWS-0019.

Kremer, R., M. Van der Meer, J. Niemeijer, B. Koehorst and E. Calle, 2001. *Technisch Rapport Waterkerende Grondconstructies; Geotechnische aspecten van dijken, dammen en boezemkaden*. Tech. rep., Rijkswaterstaat, DWW.

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

OI, 2015. *Handreiking ontwerpen met overstromingskansen, Veiligheidsfactoren en belastingen bij nieuwe overstromingsnormen*. Report Concept vs. 2.5, Rijkswaterstaat Water, Verkeer en Leefomgeving.

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

Rozing, A., 2015. *Onzekerheden Waterspanningen in WTI 2017*. Memo 1220083-004-GEO-0003, Deltares.

Schweckendiek, T. and W. Kanning, 2016. *Reliability updating for slope stability of dikes. Approach with fragility curves (background report)*. Tech. rep., Deltares report 1230090-033-GEO-0001.

Schweckendiek, T. and M. Van der Krogt, 2015. *Verkenning Bewezen Sterkte Markermeerdijken Onderdeel vervolgonderzoek pompen+geotechniek*. Report 1221189-000-GEO-0004, Deltares.

Spits, L., 2012. *Dijken op veen - Grondwater Respons Onderzoek peilbuisraaien in de gemeenten Waterland en Zeevang. Deltares report 1203768-015.* Tech. rep., Deltares.

TAW, 2001. *Technisch Rapport Waterkerende Grondconstructies Geotechnische aspecten van dijken, dammen en boezemkaden*. Report, Technische Adviescommissie voor de Waterkeringen.

TNO. "REGIS-II Hydrogeologie." Http://www2.dinoloket.nl/nl/about/modellen/regis.html.

Van Deen, J. and A. Van Duinen, 2016. *Schematiseringshandleiding Macrostabiliteit*. Tech. rep., Deltares report 1220083-008-GEO-000.

Van Duinen, A., 2014. "Schematisering waterspanningen in WTI 2017 (Ringtoets)."

Van Duinen, A., 2015. *Modelonzekerheidsfactoren Spencer-Van der Meij model en ongedraineerde schuifsterkte. Programma WTI 2017, cluster Stabiliteit*. Report 1207808-001, Deltares report 1207808-001.

van Zee, L., 2014. *Evaluatie predictie kruindaling en controle dwarsprofielen Markermeerdijk*. Report 1214-0007-020, Fugro Geoservices B.V.

Vos, P. and S. De Vries, 2016. *Geologisch onderzoek naar de heterogeniteit van de ondergrond onder de Markermeerdijk tussen Hoorn en Amsterdam.* Tech. rep., Deltares, project 1230206-000.

VTV, 2007. *Voorschrift Toetsen op Veiligheid Primaire Waterkeringen*. Report ISBN: 978-90-369-5762-5, Ministerie van Verkeer en Waterstaat.

WVL, R., 2016. *Plan van aanpak Operationaliseren bewezen sterkte.* Tech. rep., Rijkswaterstaat draft report v0.4.

Zwanenburg, C., 2014a. *Dijken op Veen II. DoV werkwijze voor bepaling macrostabiliteit Markermeerdijken*. Report 1208254-032-GEO-0001, Deltares.

Zwanenburg, C., 2014b. *Dijken op Veen II: rekenvoorbeelden ter illustratie DoV werkwijze*. Report 1208254-043-GEO-0001-gbh, Deltares.

APPENDIX

A Slope reliability using beta-h curves and FORM

A.1 Introduction

This appendix describes how reliability analyses for slope stability of dikes are carried out, using FORM (Hasofer and Lind, 1974) and fragility curves (beta-h curves). The analyses are performed in the Probabilistic Toolkit (PTK) which links the macrostability kernel with the probabilistic libraries in C# programming language. Below follows a description of the workflow.

A.2 Workflow

Standard reliability approaches like FORM are efficient and fast means for the calculation of the reliability of complex systems. However, FORM can be sensitive in case of strong discontinuities or singularities of the limit state function. In case of slope stability problems, the limit state function itself is rather simple, however the stability model is sensitive to discontinuities like nonlinear material behavior, pore pressure distributions and other nonlinearities, caused by changing water levels. This formulates the need for a robust and efficient probabilistic calculation method, which can be used to calculate the reliability of a slope stability problem.

As such, it is chosen that the water level h is not considered directly as a random variable; instead, a conditional probability of failure $P_f|h_i$ is calculated, which is used to construct the the fragility curve (beta-h), presented in at a later stage of the workflow as presented in Figure A.1.

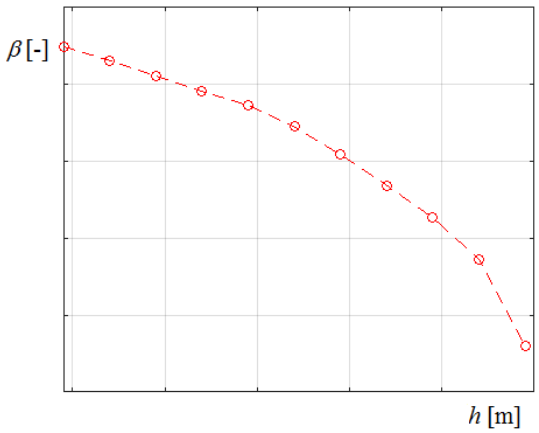


Figure A.1: *Fragility curve using the conditional reliability index $\beta|h$*

The workflow for the calculation of the reliability index and of the influence factors with the PTK, comprises the following steps:

1 At first, the distributions of random variables are defined with the corresponding mean values, standard deviations and - if needed - the correlation matrix;

2 Reliability analyses are performed using FORM: PTK prepares the parameter set and put these in an input file. PTK sends the files to the macrostability kernel. The limit state equation g as in eq.(A.1) is evaluated using the calculated stability factor.

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

Herein is SF the stability factor that is retrieved from the macrostability kernel calculation of the critical slip surface (Uplift-Van). The m_d is the model uncertainty[1] For all water levels h_i the reliability conditional on this waterlevel ($\beta|h_i$) is calculated using the stochastic soil properties. The convergence criteria of the FORM analysis and the maximum number of iterations is steered through the PTK settings;

3 Steps 1 to 2 are repeated for different water levels between a lowest water level h_{min} and a maximum water level h_{max} that depend on the case study.

At this point, the following is known for different water levels:

- conditional probability of failure $P_f|h_i$,
- the corresponding reliability index $\beta|h_i$.
- and the influence coefficients $\alpha_j|h_i$.

These results form the actual fragility curve and the estimation of the overall reliability, conditional a water level. As far as the construction of the fragility curve is concerned, a linear interpolation is applied among the fragility points in Figure A.1. The amount of the water levels that are chosen, should be such, so as the inaccuracy of the interpolation does not become an important factor (Schweckendiek and Kanning, 2016).

Regarding the evaluation of the overall reliability, an additional prerequisite is the occurrence probability, $f(h)$, of the water level, h that is under consideration. Therefore, the probability distribution of the water level should be input. The evaluation of the overall probability of failure, P_f, can be done among others by numerical integration, which is shown in eq.(A.2).

$$P_f = \int_{-\infty}^{+\infty} P_f|h_i \cdot f(h_i)dh \approx \sum_{-\infty}^{+\infty} P_f|h_i \cdot f(h_i) \cdot \Delta h \tag{A.2}$$

where Δh is the discrete interval chosen for the integration over the water level.

A.3 Output

The used software PTK, produces the further output:

- Results conditional to (a selection of) specific water levels:
  probability of failure $P_f|h_i$
  reliability index $\beta|h_i$
  vector of influence coefficients $\alpha_j|h_i$
  design point $X_j|h_i$
- Results independent of the water level (including integration over water level domain):
  probability of failure P_f
  reliability index β
  vector of influence coefficients α_j
  design point X_j
- Plot of the fragility curve (reliability vs. water level)

[1] The definition of m in van Duinen (2015) is the inverse of the definition used in this report and the values of Van Duinen (2015) are adapted correspondingly.

B Starting points

B.1 Time-invariant and time-variant variables

Schweckendiek and Kanning (2016) showed the importance of making the distinction between time-invariant properties (i.e. epistemic, reducible uncertainty) and properties that are variable in time (i.e. aleatory, irreducible uncertainty). In the presented approach, we chose to assign the basic random variables to either category, while in reality the respective probability distributions may contain contributions of both epistemic and aleatory uncertainty. In other words, we need to decide per random variable whether the uncertainty is predominantly epistemic or predominantly aleatory. Below we provide considerations for assigning variables to either category as indicated in Table 3.1.

Schematisation

The variables/parameters categorized as schematisation are considered in scenarios. To a scenario a certain probability of occurrence is given, which is related to the associated uncertainty. The scenario can be time-variant or time-invariant. 'Soil layering' is constant in time (on this scale), so fully correlated in time. Subsidence can take different values for the past and future (time-variant) and will be therefore different in assessment and observation situations.

Soil properties

The soil properties (Su ratio, strength increase exponent, yield stress, volumetric weight, friction angle) are typically assumed time-invariant variables and not changing in time. Apart from subsidence, the Markermeerdijken are prone to shear strength degradation by oxidation of peat, however this would be limited to a small top layer. Therefore, it is not expected that this degradation has a significant effect on the stability factor of the two test cases. Especially for the observation situation based on the current daily conditions, it does not play a role.

Geohydrological parameters

The height of the phreatic line is dominated by rainfall, however the outside water level is too. Therefore the phreatic line is *partly* correlated with the outside water level, but not correlated in time.

The polder water level is influenced by other variable random variables (*e.g.* pumps), and therefore assumed to be not correlated in time.

Leakage lengths are calculation parameters that influence pore water pressures in the aquifer and blanket and are directly related to permeability and thickness of blanket and aquifer. On the engineering time scale and given no large changes in the circumstances are expected, the variable is correlated in time. However it cannot be excluded that such changes will not be present in the future. So, if this is the case, the assumption should be reconsidered.

The intrusion length depends on the duration of the load condition, so it can differ from observation and assessment. Yet, the load conditions considered in this study are all very long duration events, so we assume similar intrusion length for assessment and observation and hence, time-invariance.

Loads

The outside water level is varying randomly and is typically an irreducible variable. The traffic

load is also a varible which is time-invariant. Both are are modelled uncorrelated in time.

B.2 Starting points of the prior analysis

The proposed basic schematisations are based on data from the project *Dijken op Veen* II (DoV2) and FUGRO reports (references along the following section). In general, the starting points for the analyses are chosen as close to Halter *et al.* (2015) and WBI2017 (Van Deen and Van Duinen, 2016) as possible. As result, this includes specific choices:

1 The starting point for the **geometry and soil layering** is the available measured profile and original schematisation found from the year 2000. Ever since, the dike and subsoil have been subject to subsidence. Compared to AHN lidar measurements (AHN, 2015) approximately from the year 2012, a difference with the original profile is 15 cm on average. This is in accordance with the average subsidence of 1 cm per year, along the Markermeerdijken, according to van Zee (2014). Since the measured profile dates from the year 2000 and the considered assessment for the year 2023, subsidence is taken into account. This subsidence is considered as a reduction of the thickness in the most shallow peat layer, since this is the most compressible layer. The expected subsidence for the assessment situation (between 2000 and 2023) is of 25 cm. The polder water level is also lowered to the same extent. The soil layering, in both test cases, is consistent with available cone and ball penetration tests. The foreshore is not regarded, since this is not relevant for inner slope stability.

2 The **Critical State Soil Mechanics** (CSSM) framework is used. This includes the use of undrained shear strength in the sliding model, according to WBI2017 (Van Deen and Van Duinen, 2016). A stress-dependent ratio S ('Su-Calculated') is used to calculate the shear strength for different water levels and effective stresses. Furthermore, a strength exponent m and over consolidation ratio OCR[1] is needed. D-Geo Stability software is used for the slip plane analyses (using the method of slices, as described in Deltares (2016). The Uplift-Van model and the correspondent model uncertainty are considered for all the analysis performed. The model uncertainty follows a lognormal distribution with mean value $\mu_{m_d} = 1.005$ and standard deviation of $\sigma_{m_d} = 0.033$ (Van Duinen, 2015). In Dutch semi-probabilistic safety assessments model uncertainty is covered by a model factor γ_d, equal to the 5% conservative bound of the model uncertainty; γ_d is 1.06 based on Van Duinen (2015). This results in a design value of the stability factor $SF_d = SF/\gamma_d$, see Table B.6.

3 In the schematisations, the same soil materials/types are present at different depths. Because the distributions of the **strength parameters** S and m are derived per soil type, regardless the depths of the samples, the strength parameters S and m are taken into account fully correlated between soils of the same type.
Furthermore, soil layers are split up because of different volumetric weights and different state conditions next to and below the dike. The strength parameters S and m of these layers are also fully correlated.

4 **Uncertainties in strength parameters** are from the experimental research of Halter *et al.* (2015). This includes S-ratios, m-exponents and POP values for different types of peat and clay from a regional dataset of direct simple shear tests (peat) and triaxial tests (clay). From the orignal data, distributions for the parameters are fitted and spatial averaging is applied. This is further described in section B.4.1. The use of a regional dataset with a limited number of representative laboratory tests per soil type leads to larger uncertainty in the parameters and therefore more conservative safety assessments, compared to if local data.

[1] $OCR = \frac{\sigma'_y}{\sigma'_v} = \frac{\sigma'_v + POP}{\sigma'_v}$, where σ'_y is the yield stress, σ'_v is the vertical effective stress and POP is the pre-overburden stress.

5 The schematisation of **pore water pressures** is done using the available software (Waternet creator (Kanning, 2016), (Van Duinen, 2014) in accordance with assumptions in DoV2 - Zwanenburg (2014a), which based on Halter and Effing (2011). This includes leakage lengths instead of hydraulic head response factors applied to the head of the aquifer layer. Both the phreatic line (PL1) and the head in the aquifer, during daily and increased water level conditions (PL2 and PL3), are schematised according to DoV2.

In the test cases the phreatic line is schematised at locations (a), (b) and (c), as shown in Figure B.1. Note that since the base case has a lowered surface level, the polder water level is also lowered to the same extent (in comparison with the original schematisation). For the test cases, the phreatic level in the dike in daily and assessment situations is higher than the outside water level. This difference, between the outer water level (a) and the phreatic level in the dike body (b), varies from the two situations, therefore, this relation and corresponding interpolation is made using linear relation. All values mentioned for the schematisation of the pore water pressures are according to Zwanenburg (2014b). The intrusion length is taken as 2.3 m with a standard deviation of 0.69 m, according to section B.4.5, i.e. CoV =0.3.

Regarding the pore water pressure distribution in the vertical direction, reference is made to section B.4.5

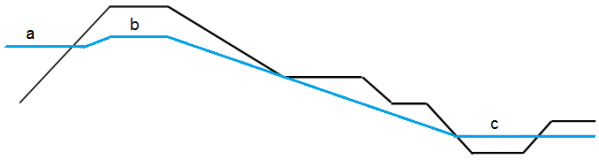


Figure B.1: *Schematisation of the phreatic level in the dike.*

6 **Uncertainties in pore water pressures** are taken into account according to Kanning (2016), using the uncertainties proposed by Rozing (2015), see details in section B.4.5.

7 The effect of **uncertainties in volumetric weight** is checked by means of a sensitivity study for high and low values (see appendix D). Theoretically, the volumetric weight can be also taken into account by means of random variables. However, this is not in accordance with current practice, since the influence of the volumetric weight can be both positive and negative. Furthermore, if using a POP in order to determine the yield stress and OCR, there is an effect of the schematisation of the effective in situ stress. This might end up taking the uncertainties (*in yield stress and volumetric weight*) into account twice. For these reasons, no uncertainty on the volumetric weight is taken into account.

8 The probability density distribution of the **water level** was derived from the annual maximum of the lake level and will be applied for both cases, see section B.4.6.

9 **Traffic load** is taken into account in the base case for the assessment situation only, according to the statutory Safety Assessment (VTV, 2007). Its value is 13.3 kN/m^2 over a width of 2.5 meters at the crest of the dike (Kremer *et al.*, 2001). For the applied (temporary) traffic load, an angle of dispersion(default=0) and a degree of consolidation are specified. The degree of consolidation defines the (excess) pore water pressures, as result of the load, per soil layer. For both cases, the degree of consolidation was taken as 20% for the clay and peat layers, according to the choices made in the schematisation of DOV. In current practice, 0% is the general applied conservative value.

10 **Unsaturated strength** is not taken into account in this study. The unsaturated zone affecting the slip plane is typically small for the Markermeerdijken and analyses for similar conditions at the IJsseldijken showed that the effect is negligible.

11 **Houses** in the dike are schematised as a gap in the dike body. For a house with a shallow footing, additional house load can be applied, however it was decided by the Expert Team Bewezen Sterkte (Utrecht, March 23th, 2016) to not include an additional house load in the assessment situation. This is a conservative assumption in this case, since it is expected that even if a house will collapse at some point, the material (weight) is still present at the location. Reasons a house load could be absent are e.g. fire and the demolition of a house without the water board being advised.

    In the case of a house with pile foundation, all weight is transferred to the Pleistocene sand layer and therefore there is no additional load of the house. The additional strength by pile is discussed in section 6.3. Other objects in or along the dike body are not considered for the cases.

12 Prior and posterior calculations are **probabilities of failure per year**, since it corresponds to the definition of the Dutch safety assessment norms (OI, 2015). The probability of failure (or in terms of reliability) is computed according to the method fragility curves, described in the background report Schweckendiek and Kanning (2016). A specific description of the work-flow and used software is included in appendix A. A more practical document describing the steps to take in order to perform the reliability updating with fragility curves is the Manual (NL: 'Werkwijzer'). This is a product of POV-M (Project Overstijgende Verkenning Macrostabiliteit), however it is still under development.

For further detail on the shear strength parameters and uncertainties, pore pressures and respective uncertainties and also the assumptions for the load (*i.e.* water level), reference is made to section B.4.

B.3 Starting points of the reliability updating

Evidence for reliability updating can be found in survived observations. The daily situation is one of such survived conditions. The reliability updating of the assessment base case will be done with past performance information from the observation base case. Various parameters are assumed to be fully correlated in time and therefore equal in the assessment and base case (time invariate variables, as described in section B.1). However, some variables could have been different in the observation, compared to the the assessment situation. These variables are summed below. The changed values (time variate variables) are described in chapter 4 and chapter 5.

1 **Subsidence** is also taken into account for the observed situation. As mentioned in section B.2, the subsidence is about 1 cm per year. So compared to the original geometry of the year 2000, subsidence of 15 cm is taken for the observation, taken into account as a reduction of the thickness of the most shallow peat layer, since this is the most compressible layer. The polder water level is also lowered to the same extent.

2 The schematisation of the **phreatic line** is done by the best estimate for the assessment and observation, as already described in section B.2. This includes that the phreatic level is dependent on the outside water level, since both are caused by heavy rainfall. Obviously, there are also other influences on the phreatic level, which will be discussed in section 6.3.

3 **Traffic load** is taken into account for the assessment situation only. In the base case of the observed situation, traffic load is not taken into account.

4 Dikes could have survived due to a positive effect of a house in the dike. So, for a house with a shallow footing, an additional **house load** will be applied in the observed situation. This is decided in by the Expert Team Bewezen Sterkte (Utrecht, March 23th,

2016). The considered house load is referred to chapter 5.

B.4 Further elaboration on points mentioned in section B.2 and B.3

B.4.1 Shear strength parameters and uncertainties

The shear strength parameters S, m and POP are derived from regional sampling (Halter *et al.*, 2015). Therefore, the parameters are the same for both test cases along the Markermeerdijk. Local averaging (along a slip plane) is accounted for, using eq.(C.5) (TAW, 2001). Furthermore, uncertainty by the limited number of samples is also accounted for. In Table B.1, the obtained values are shown. Table B.3 presents the uncertainties for the strength drained parameters of sandy layers present in the area. The correspondent coefficients of variation are presented in Tables B.2 and B.4 respectively.

In the following tables μ stands for mean and σ for standard deviation of the correspondent probability density function (PDF).

Table B.1: *Lognormal distribution parameters of the undrained shear strength parameters*

Soil type (Kr, BiT, A)	$S[-]$		$m[-]$		$POP[kPa]$	
	μ	σ	μ	σ	μ	σ
Klei, antr	0.359	0.100	0.881	0.059	31.8	7.8
Klei, hum	0.269	0.049	0.902	0.029	26.7	11.0
Klei, siltig	0.260	0.026	0.863	0.049	20.7	7.1
Klei, zandig	0.260	0.026	0.863	0.049	20.7	7.1
Veen	0.524	0.057	0.888	0.061	15.4	13.6
Veen, kleiig	0.407	0.024	0.925	0.026	24.2	7.9

Table B.2: *Coefficient of variation (CoV) for the undrained shear strength parameters*

Soil type (Kr, BiT, A)	$CoV(S)[-]$	$CoV(m)[-]$	$CoV(POP)[-]$
Klei, antr	0.28	0.07	0.24
Klei, hum	0.18	0.03	0.41
Klei, siltig	0.10	0.06	0.34
Klei, zandig	0.10	0.06	0.34
Veen	0.11	0.07	0.88
Veen, kleiig	0.06	0.03	0.33

Table B.3: *Lognormal distribution parameters of the drained shear strength parameters*

Soil type	$c'[kPa]$		$\phi'[^\circ]$	
	μ	σ	μ	σ
Zand, antr	0	0	35	1.8
Zand, calais	0	0	35	1.8
Zand, pleist	0	0	35	1.8

Table B.4: *Coefficient of variation (CoV) of the drained shear strength parameters*

Soil type	$CoV(c)[-]$	$CoV(\phi)[-]$
Zand, antr	0	0.05
Zand, calais	0	0.05
Zand, pleist	0	0.05

The measured uncertainty of the POP values is very high (CoV>0.3). Therefore, it is desirable to derive local POP values from CPT tests. Since not many CPT's are available, the uncertainty is taken from the available test sample results, shown in Table B.1.

B.4.2 CPT's to determine soil layering case house

Figure B.2 shows available penetration tests which were used to determine the soil layering. In between (i.e. below the inner toe) the profile is interpolated between the two profiles.

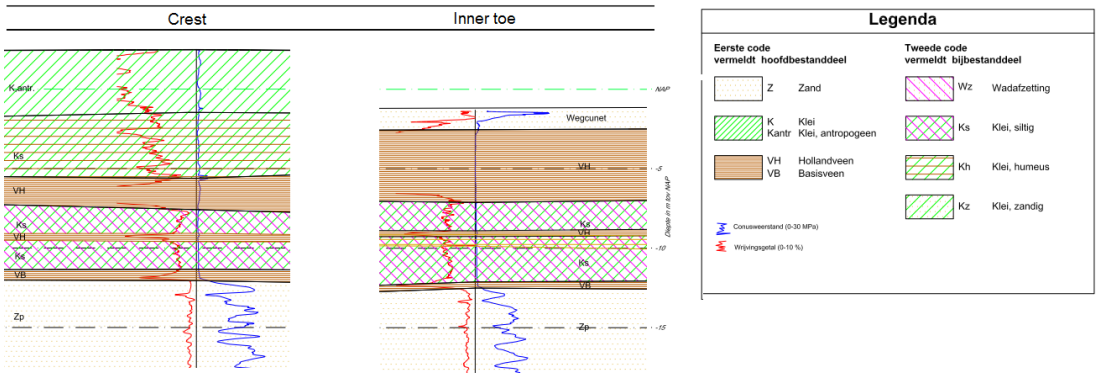


Figure B.2: *Case house: Soil layering below crest and inner toe*

B.4.3 Leakage lengths

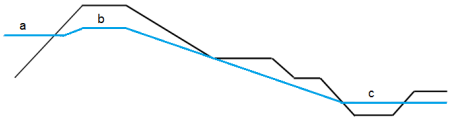
The inner side leakage length is estimated using eq.(B.1). For the 'case house' a λ_{in} of 3200 m is achieved, based on a permeability of the aquifer k_h = 20 m/d and thickness D =10 m (TNO), while for the weak soil layers, an average permeability k_v =2 mm/d (Halter, 2008) and thickness of d =10 m are considered. The measured response factor in head level in the aquifer is 0.40 (mentioned in section B.2), as such, an outer side leakage length of λ_{out} = 4800 m is derived.

$$\lambda_{in} = \sqrt{\frac{k_h D d}{k_v}} = \sqrt{\frac{20^m/_d \cdot 10m \cdot 10m}{0.2 \cdot 10^{-3} m/d}} \approx 3200m \tag{B.1}$$

B.4.4 Geohydrological properties in the observation

For the geohydrologycal parameters, the same assumptions in the observation as for the assessment conditions are considered - see starting points in section B.3. Hence, best estimates are used for the observation. The phreatic line is schematised at locations (a), (b) and (c), as shown in Table B.5 - example of 'case house'. Since the base case has a lowered surface level (due to subsidence), the polder water level is also lowered to the same extent and therefore, this is the only difference between the geohydrological schamatisation of the assessment and the observation. At the three locations, the phreatic line is estimated for both the daily conditions and a high water situation. This is done based on measurements (see section B.3). In between these points the phreatic line is estimated at the three points as a function of the outside water level, using interpolation between the daily and high water situation.

Table B.5: *Case house: schematization of the phreatic level observation*



Phreatic level in [m] above NAP per location			
	a	b	c
Daily	-0.40	+0.85	-1.90
High water	+1.15	+1.48	-1.90

B.4.5 Waternet Creator parameters and uncertainteis

Pore water pressures can have a significant impact on the inner slope stability of dikes. Ignoring uncertainties in pore water pressures can thus lead to a serious over- or under estimation of slope stability. Here it is explained the general implementation of pore water pressure uncertainties. For specific implications of pore water pressure uncertainties in Reliability Updating, reference is made to Kanning (2016). The choices recommended in Kanning (2016) are made taking into account the following considerations: (1) Uncertainties in the pore water pressures should be well reflected by the parameters that are chosen as random variables, (2) keep the implementation as simple as possible and (3) use the values of Rozing (2015) where applicable.

The pore water pressures are modelled within D-Geo Stability using the Waternet Creator, specific input for the considered test cases:

- **Type of dike**: both test cases here presented have an impermeable clay core and are situated on an impermeable subsoil ('clay on clay', type 1A).
- **Phreatic line** (PL1) definition, see schematisation per case in chapters 4 and 5.
- **Head level** in the aquifer/sand layer during daily (stationary) conditions (PL2), see schematisation per case in chapter 4 and chapter 5.
- **Leakage lengths**, which indicate how large the influence of an outside water level increase is on the head level in the aquifer (so-called response factor) and indicate how far this head level increase propagates in the sand layer in horizontal direction. The considered response factor used in these test cases is from the observed for the Markermeer area, with a value of 0.40 (Spits, 2012). The coefficient of variation on the leakage lengths 20% and a lognormal distribution.
- **Intrusion length**, which indicates how far do pore water pressures from the sand layer intrude in the clay/peat: mean value of 2.3 meters and a coefficient of variation of 30% truncated normal distribution [2].

Regarding the pore water pressure distribution in the vertical direction: the pore water pressures are taken hydrostatic (head PL1) until the bottom of the first peat layer (*Botoom hydrostatic zone*). From this level, the pore water pressures are interpolated linear until the top of the *Penetration zone* (head PL2). From this point onwards the pore water pressures are interpolated linear until the top of the aquifer (between PL2 and PL3). This is shown in Figure B.3.

For more explaining pictures, reference is made to Kanning (2016). Specific values of the variables per test case will be presented at the corresponding chapter.

[2] This distribution avoids the intrusion line goes beyond the *bottom hydrostatic zone* and beyond the aquifer.

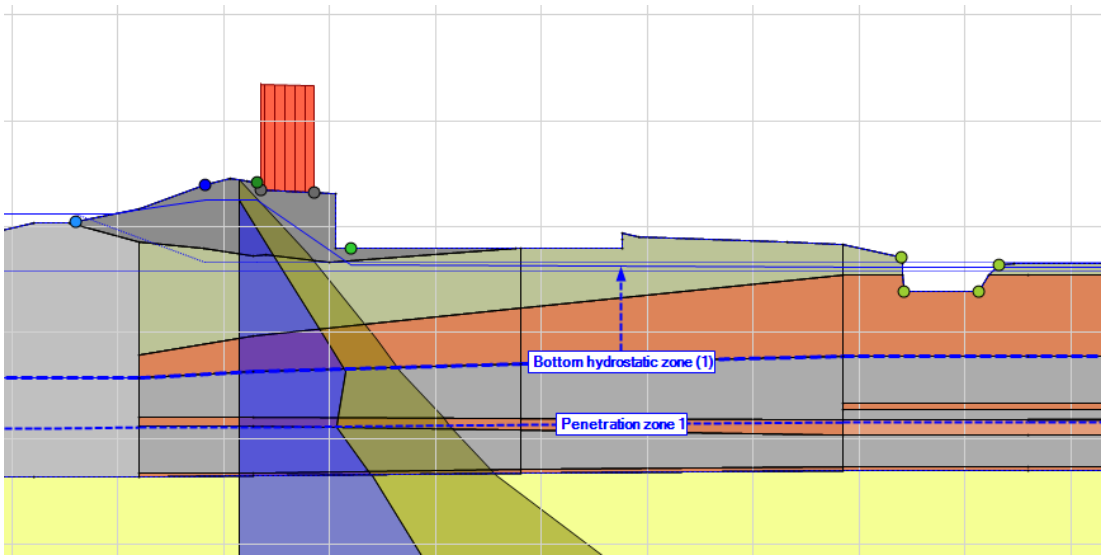


Figure B.3: *Pore water pressure development over vertical.*

B.4.6 Load: water level conditions

The applied probability distribution for the water level is shown in Figure B.4. The Gumbel distribution is fitted to the annual maximums of the daily average (winter) water level between the years 1975 and 2012, see Schweckendiek and Van der Krogt (2015).

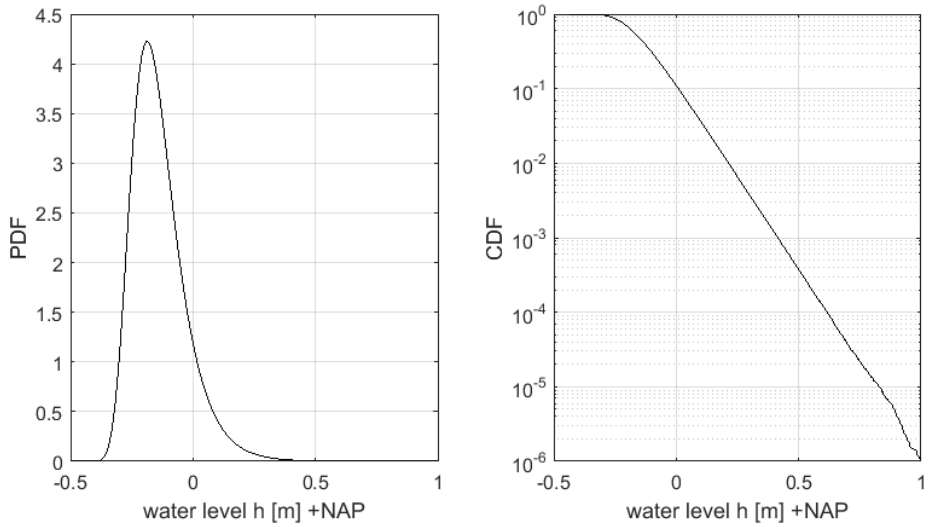


Figure B.4: *Probability distribution of the water level (Gumbel distribution fitted to annual maximums of the daily average water level at the lake, Gumbel location parameter* $a = -0.187$*, Gumbel scale parameter* $b = 0.087$*)*

B.4.7 Summary table of the uncertainties

In Table B.6 a summary is given for the uncertainties of different parameters used. Here μ stands for mean and σ for standard deviation of the correspondent probability density function (PDF).

Table B.6: *Summary table of the uncertainties considered in the reliability analysis for both test cases*

Parameter	PDF type	PDF parameters	Reference
Outside water level h	Gumbel	$loc = -0.187, scale = 0.087$	
Intrusion length IL: shift	Truncated normal	$\mu = 0, \sigma = 0.69$	Rozing (2015)
Leakage length λ	Lognormal	$CoV = 0.3$	Rozing (2015)
model uncertainty, m_d **	Lognormal	$\mu = 0.995, \sigma = 0.033$	Van Duinen (2015)
S-ratio for cohesive soils	Lognormal	see Table B.1	Halter *et al.* (2015)
m-exponent for cohesive soils	Lognormal	see Table B.1	Halter *et al.* (2015)
σ'_y, yield stress points	Shifted lognormal	*	Halter *et al.* (2015)
$tan(\phi')$ for sandy layers	Lognormal	$\mu = 0.699, \sigma = 0.042$	Halter *et al.* (2015)

*) The mean value of the yield stress point is computed by $\sigma'_y = \sigma'_v + POP$, where σ'_v is computed for the conditions when POP was measured, *i.e.* daily conditions. Furthermore, a shift to the lognormal distribution is added in order to avoid simulating yield stresses lower than the actual effective stress.
**) The definition of m_d in Van Duinen (2015) is the inverse of the definition used in this report and the values of van Duinen are adapted correspondingly.

C Spatial variability and averaging

Concerning the derived standard deviation of the undrained parameters, in the report Calle and Kanning (2013) one can consult a comprehensive overview of the different aspects of spatial variability on the modelling of geotechnical strength of flood defences. One important aspect of the spatial variability is the averaging. Averaging refers to the observation that soil properties fluctuate rapidly in the vertical dimension relative to the dimension of the failure plane, which results in partial averaging of the uncertainty.

There are various formulas with respect to averaging in slope stability, but as we typically deal with regional datasets, a stochastic model was developed in Calle and Kanning (2013), see Figure C.1. This model basically says that a part of the regional variance is due to local fluctuation of the sheer strength and a part is due to fluctuations in the local mean - eq.(C.1).

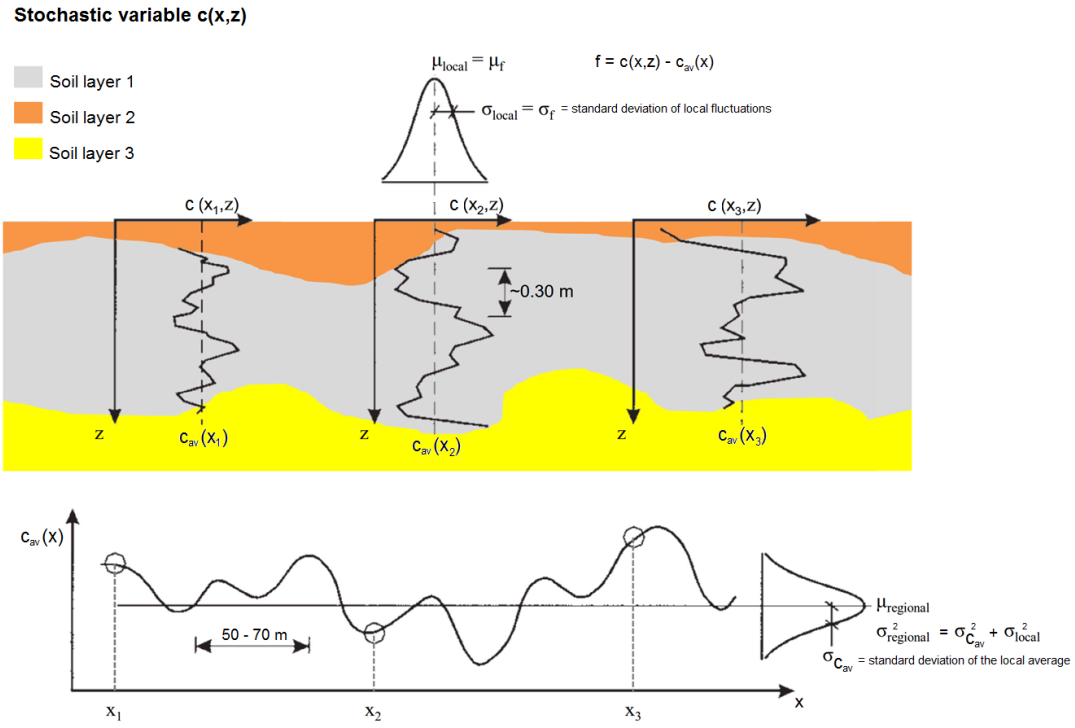


Figure C.1: *Stochastic model for regional and local standard deviations, adapted from Calle and Kanning (2013)*

$$\sigma_{reg}^2 = \sigma_{loc,aver}^2 + \sigma_{loc}^2 \tag{C.1}$$

There are three effects that determine the local, average standard deviation ($\sigma_{loc,aver}$), that should be an input into the computation based on the measured data (σ_{reg}):

- Incorporate the relation between regional and local variability: the a factor, eq.(C.3),
- Incorporate local averaging along failure plane: the γ_d factor,
- Incorporate the effect of limited number of measurements: n.

This can be summarized by the following eq.(C.2), an equation with unknown origin but used in the (e.g.) 'schematiseringshandleiding'.

$$\sigma_{loc,aver} = \sigma_{reg}\sqrt{(1-a) + a \cdot \gamma_d + 1/n} \tag{C.2}$$

with:

$$a = \frac{\sigma_{loc}^2}{\sigma_{reg}^2} \tag{C.3}$$

$$\gamma_d = min(D_v\sqrt{\pi/d}, 1) \tag{C.4}$$

where: Assuming that all local variances averages ($\gamma_d = 0$), the eq.(C.2) reduces to:

σ_{reg}	is the standard deviation of the regional variation,
a	is the portion of the total variability stemming from local variability (Eq.C.3), and $(1-a)$ the fluctuations of local means [default: $1-a = 0.27$]. In case of a local dataset $a = 1$,
γ_d	is the variance reduction factor (Eq.C.4), which takes value $zero$ when it is assumed that all local variances averages,
D_v	is the vertical correlation length,
d	is the layer thickness,
n	is the number of the regional samples.

$$\sigma_{loc,aver} = \sigma_{reg}\sqrt{(1-a) + 1/n} \tag{C.5}$$

D Sensitivity analysis with deterministic and stochastic soil volumetric weight

D.1 Introduction

This appendix presents a sensitivity analysis concerning the use of deterministic or stochastic soil volumetric weight (γ) for the study of the reliability index of a slope stability. Note that this choice has influence in the generated vertical stresses (total and effective) and therefore also in the generated yield stress point, which is initially set as the vertical effective stress at daily water level conditions added by the pre-overconsolidation pressure POP value (measured at the same water level conditions) - eq.(D.1). To understand the influence of such a choice (stochastic or deterministic), the reliability of the slope stability was performed with two approaches:

a the yield stress points are adjusted based on the changes in γ,
b the yield stress points are kept constant (corresponding to the mean values of γ).

Approach (a) is actually representing in this case what is happening in reality, since the calculated yield stresses is dependent on the initial schematisation of effective stresses added by the POP measured in the lab tests (the higher the volumetric weight is, the higher the effective stress and therefore yield stresses). However, approach (b) can also be representative of reality, for example, in case that the uncertainty that is considered for the yield stress points includes already some variation in the volumetric weight. In such a case, approach (a) would double count the uncertainties in the volumetric weight.

$$\sigma'_y = \sigma'_v + POP \tag{D.1}$$

Where σ'_y [kPa] is the yield stress and σ'_v [kPa] is the vertical effective stress for the water level conditions on which POP [kPa] is measured.

It is important to mention that the influence of the soil volumetric weight on the reliability is heavily connected with the modelling and schematisation of yield stress points. Also that, from the physics, the yield stresses being a measured state parameter of the soil it should not change with increasing/changing volumetric weights during the analysis. However, in this sensitivity analyses, the yield stress points are not measured parameters, but schematised by adding a POP value per soil to the schematized effective stress.

D.2 Input

The contemplated cross section was an input case for the research of Schweckendiek and Van der Krogt (2015) and it does not coincide with the current report's test cases. The schematised cross section is along the the Markermeerdijken and can be seen in Figure D.1.

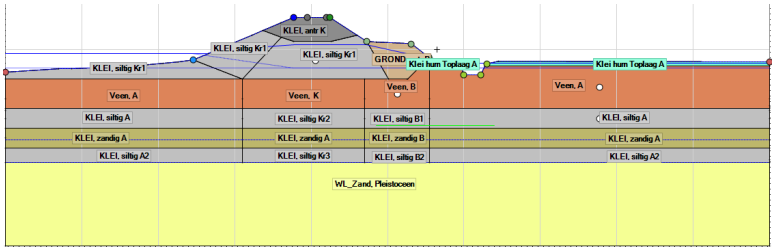


Figure D.1: *Schematization of the cross section used for the senstivity analyses of the volumteric weight.*

Apart from the soil volumetric weight (γ) and the yield stress points (σ'_y), the variables that were considered as random are the S-ratio and the m-exponent of the undrained shear strength model (layers Klei and Veen, i.e. Clay and Peat). A lognormal distribution was used for all the random variables. Mean and coefficient of variations were taken from Schweckendiek and Van der Krogt (2015). Moreover, a model factor was considered as a random variable, lognormally distributed, with a mean value of 0.995 and a standard deviation of 0.035.

As mentioned, the yield stress points are computed based on the effective vertical stress and the POP values measured per layer. Their mean values and uncertainties are presented in Table D.1.

Table D.1: *POP values and their standard deviation for each soil layer (according to Schweckendiek and Van der Krogt (2015)).*

Soil layer	POP mean $\pm$ standard deviation [kPa]
KLEI, antr K	10.0 $\pm$ 4.0
KLEI, siltig, Kr1	10.0 $\pm$ 4.0
Veen, K	20.0 $\pm$ 8.0
Veen, B	20.0 $\pm$ 8.0
Veen, A	10.0 $\pm$ 2.5
KLEI, siltig B1	10.0 $\pm$ 4.0
KLEI, siltig A	5.0 $\pm$ 2.5

Also concerning the schematisation of the phreatic line, the same assumptions of Schweckendiek and Van der Krogt (2015) were taken. More specifically, an unfavourable scenario was chosen for the minimum phreatic level (Dupuit level) and the phreatic line offset in the inner and the outer side of the dike that are described in the table below:

Table D.2: *Minimum phreatic level and the PL1 offset (input in the Waternet Creator according to Schweckendiek and Van der Krogt (2015)).*

DupuitLeft	NAP +1.2 m
DupuitRight	NAP +1.2 m
OffsetLeft	NAP -0.70 m
OffsetRight	NAP -0.70 m

D.3 Sensitivity analysis with soil volumetric weight as deterministic variable

Firstly, the variation of the volumetric weight has been applied deterministically. More precisely, three different values for the volumetric weight of each soil layer were investigated:

- Mean values;
- 10% higher values than the mean values;
- 10% lower values than the mean values.

It should be mentioned that a minimum value for the volumetric weight of 9.85 kN/m^3 has been used, so as to be higher than the default water weight which is 9.81 kN/m^3. A lower volumetric weight might cause problems in the D-Geostability calculations. The 10% higher or lower volumetric weights are an exemplary value in order to get a first impression of the volumetric weight's influence in the reliability index of slope stability.

Notice that the soil volumetric weights in the Netherlands are usually in the lower bounds. As referred, this section only shows a first impression of the volumetric weight's influence in the reliability index. A further analysis with lower variations of the soil volumetric weight is presented in section D.4.

In Table D.3 the mean values of the volumetric weights for each soil layer are presented as well as their values with adding/subtracting the 10%. The soil volumetric weights (γ) shown in Table D.3, concern the saturated volumetric weights. Furthermore, due to the saturated conditions in the Markermeerdijk, no distinction between the saturated and the unsaturated volumetric weight is made, therefore $\gamma_{sat} = \gamma_{unsat} = \gamma$.

Table D.3: *Different scenarios of the saturated volumetric weight for each soil layer [in kN/m^3], for the sensitivity analysis of the volumetric weight as deterministic parameter.*

Soil layer	mean γ	γ+10%	γ-10%
GROND, antr B	18.00	19.80	16.20
Klei hum Toplaag A	11.00	12.10	9.90
KLEI, antr K	15.00	16.50	13.50
KLEI, siltig A	14.10	15.51	12.69
KLEI, siltig A2	14.10	15.51	12.69
KLEI, siltig B1	14.00	15.40	12.60
KLEI, siltig B2	15.00	16.50	13.50
KLEI, siltig Kr1	15.00	16.50	13.50
KLEI, siltig Kr2	15.00	16.50	13.50
KLEI, siltig Kr3	16.00	17.60	14.40
KLEI, zandig A	17.30	19.03	15.57
KLEI, zandig B	16.40	18.04	14.76
Veen, K	12.00	13.20	10.80
Veen, A	10.00	11.00	9.85
Veen, B	11.00	12.10	9.90
WL_Zand, Pleistoceen	21.00	23.10	18.90

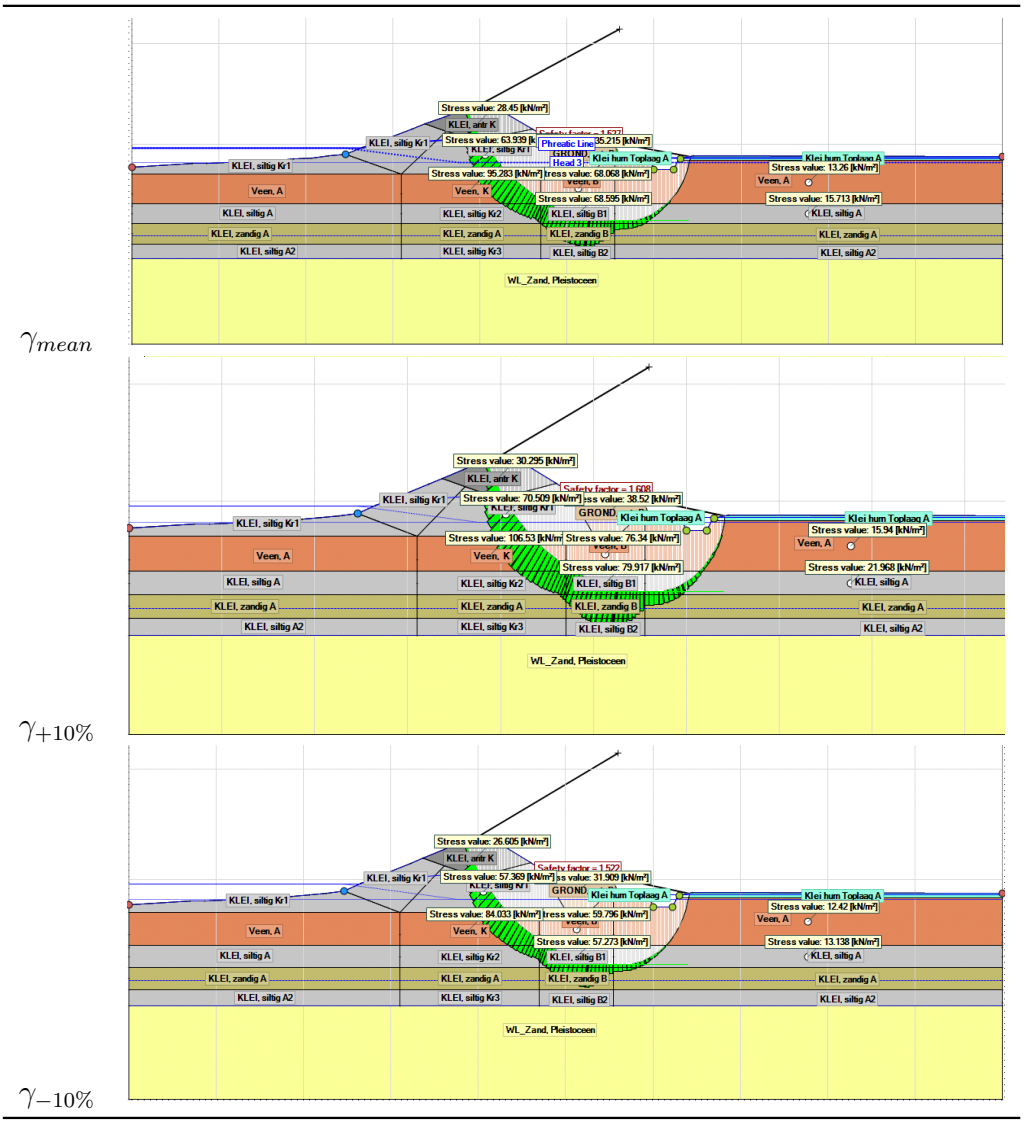
The following two sub-sections show the analysis, with the Table D.3 as starting point, and taking into account the approach (a) and approach (b), respectively in each sub-section.

D.3.1 Change in volumetric weight and in yield stress (approach a)

In the first sensitivity analysis, keeping the volumetric weight as deterministic variable, the yield stress values were adjusted - eq.(D.1) - for each scenario of the volumetric weight (mean $\pm$ 10%). Initially, a certain POP value was considered for each soil layer (see Table D.1) and then the yield stress points were re-calculated depending on the soil volumetric weight. The standard deviation of the yield stresses were taken equal to the standard deviation of the POP values (see Table D.1).

The yield stress points that were calculated for each scenario are shown in Table D.4. As it can be observed as the volumetric weight is increasing, the yield stresses are also increasing.

Table D.4: *Yield stress points for each scenario of the volumetric weight as deterministic parameter.*



The probabilistic analysis was performed with the FORM method and the results are presented in Table D.5. As it can be seen, the analysis was performed for different water levels, however, in this particular case, the cross section shows no sensitivity to the outside water level. As it can be observed, for lower volumetric weights (-10%), the reliability index does not change significantly when compared when mean values are used. However, for higher volu-

metric weights (+10%), the reliability index increases a factor of 1.2. Additionally, it observed that the lower the volumetric weights are, the lower the factor of safety is and thus the lower the reliability index is.

Table D.5: *Reliability analysis results, for the sensitivity analysis of the volumetric weight as deterministic parameter, in approach a.*

Water level [m+NAP]	Reliability index [-]		
	mean γ	γ+10%	γ-10%
-0.4	3.84	4.47	3.73
0.4	3.84	4.47	3.73
1.2	3.84	4.47	3.73

D.3.2 Change in volumetric weight without changing yield stress (approach b)

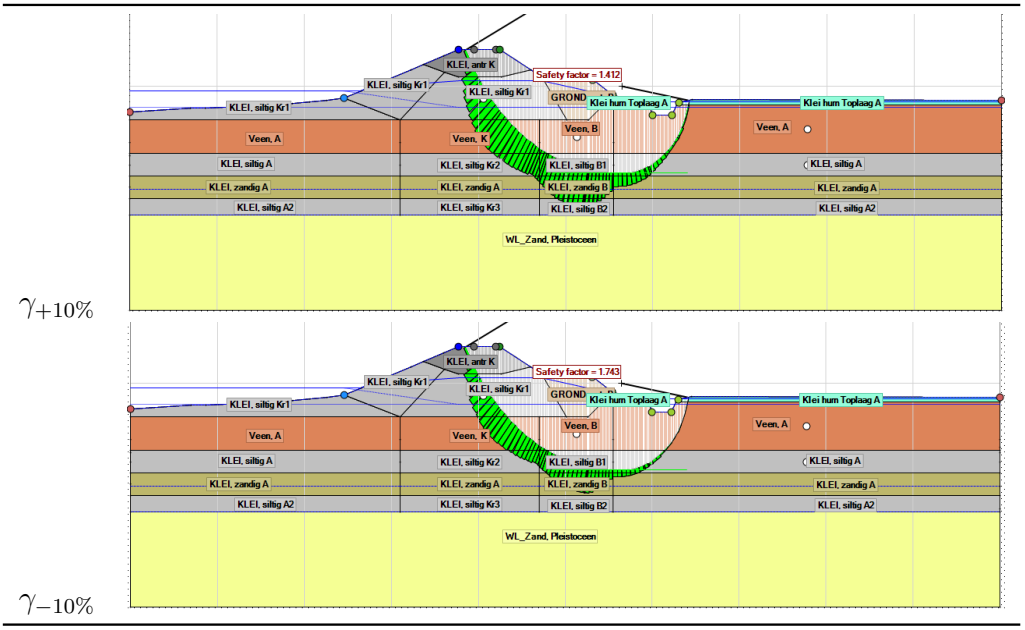
Again, the same scenarios of the volumetric weight were used (mean $\pm$ 10%), assuming the volumetric weight as a deterministic variable and keeping the values of the yield stress points the same as the base case (i.e. computed with the mean value of the volumetric weight - depicted in the top figure of Table D.4).

The results of the probabilistic analysis with FORM are presented in Table D.6. Again, the analysis was performed for different water levels. In this case, it can be observed that both for higher and lower volumetric weights, the reliability index can change considerably. In Table D.7, an illustration of the D-Geostability analysis (with the achieved safety factor), for the lower and the higher volumetric weights is shown. As it can be noticed, from the reliability results and the safety factors, the higher the volumetric weight is, the lower the factor of safety and thus the lower the reliability index is. This comes in contrast with what is happening in the first analysis in section D.3.1(approach a). This can be explained as follows: Increasing the volumetric weight, the shear strength is also increasing. This means that the active resisting moment is increasing as well. However, in approach (a), the yield stresses were also increased and this was an additional contributor to the resisting moment. On the other hand, in approach (b), the yield stresses were kept the same for all the scenarios and therefore, no additional resisting force was available in the higher volumetric weights scenario. This resulted to a more dominant driving moment. Consequently, a lower factor of safety is estimated.

Table D.6: *Reliability analysis results, for the sensitivity analysis of the volumetric weight as deterministic parameter, in approach b.*

Water level [m+NAP]	Reliability index [-]		
	mean γ	γ+10%	γ-10%
-0.4	3.84	3.15	5.14
0.4	3.84	3.15	5.14
1.2	3.84	3.15	5.14

Table D.7: *Factor of safety for each scenario of the volumetric weights (for the factor of safety with the mean values of volumteric weight, see Table D.4).*



D.4 Sensitivity analysis with with soil volumetric weight as stochastic variable

A sensitivity analysis of the cross section's reliability was carried out with the volumetric weight as a stochastic parameter. The analysis was carried out in Probabilistic Toolkit (PTK), following approaches (a) and (b). See in Figure D.2 the work-flow of each approach. For approach (a), since the yield stress points will change for every simulation of the stochastic volumetric weight, a so called model train[1] is considered, for approach (b) the yield stress points are fixed and based on the mean values of the volumetric weight. In both approaches, the volumetric weight was considered as a stochastic parameter, lognormally distributed, with a mean value specified in Table D.3 for each soil layer and coefficient of variation (CoV) of 5% and 10%, assumed to be reasonable variations according to engineering judgement.

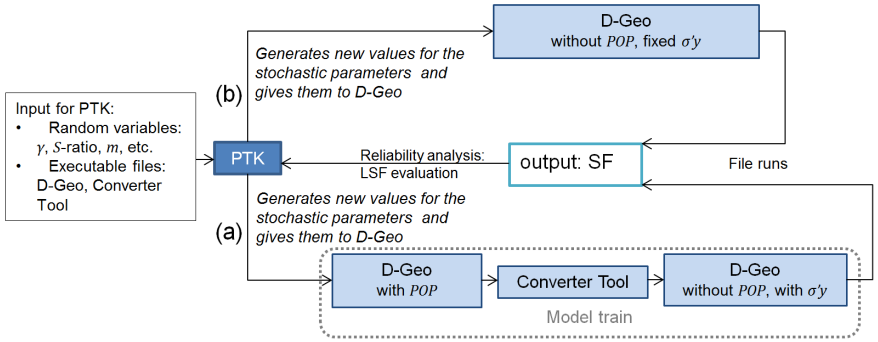


Figure D.2: *Calculation flow of PTK and D-Geostability (D-Geo) for approach (a) and (b) respectively. The output that PTK retrieves after each D-Geo calculation, is the safety factor (SF) in order to evaluate the limit state function (LSF).*

[1] A Converter Tool is being used so as to transform a *.dsx file with POP values to a *.dsx file with yield stress points. For that, the PTK is adjusted so as to call both D-Geostability and the Converter Tool one after the other.

In Table D.8, the results of the analyses are presented:

- The 1st column shows the different water levels for which the analyses were carried out,
- the 2nd refers to the analysis with the volumetric weight (γ) being considered as deterministic (base case) - result of section D.3,
- 3rd and 4th columns show the results, varying the yield stresses according to the changes of the volumetric weight using the model train - approach (a),
- and the last two columns (5th, 6th) present the results, considering the volumetric weight as stochastic but keeping the yield stresses' mean value fixed - approach (b).

Table D.8: *Reliability analysis results for approaches a and b, for the sensitivity analysis of the volumetric weight as stochastic parameter.*

	Reliability index [-]				
Water level $[m + NAP]$	γ deterministic (mean values) (base case)	approach (a) - model train γ stochast		approach (b) γ stochast	
		CoV(γ)=5%	CoV(γ)=10%	CoV(γ)=5%	CoV(γ)=10%
-0.4	3.84	3.73	3.63	3.64	3.24
0.4	3.84	3.73	3.63	3.64	3.24
1.2	3.84	3.74	3.64	3.64	3.24

As it can be seen from the results in Table D.8, keeping the volumetric weight as deterministic gives a higher reliability index. It can be also noticed that the reliability index in case of the model train (approach a) tends to be larger than keeping the yield stresses fixed (approach b). This verifies the positive effect of changing the yield stresses together with the volumetric weight (model train). For approach (a), it was also observed that the influence coefficients of the yield stresses are higher than these of the volumetric weights ($\alpha^2 = 0.14 > \alpha^2 = 0.02$), while for approach (b) these are similar ($\alpha^2 = 0.11$). The variables with the most influence in the probabilistic analysis are the S-ratio and the model uncertainty.

It is worth to notice that the analysis with CoV=10% for approach (b), had quite some difference from the base case (3.24 vs 3.84, respectively). In this approach, it was noticed that the volumetric weights of the dike material had a higher influence coefficient in comparison with the other soil layers' volumetric weights. Moreover, the design point showed that the volumetric weight of the dike material was increased compared to the mean value which consequently resulted to a higher driving force and thus a low reliability index.

D.5 Conclusions & Recommendations

This appendix analysed the influence of the volumetric weight on the reliability for slope stability. Initially, the sensitivity analysis of the volumetric weight's influence was carried out deterministically. This was achieved by assigning mean, a minimum and a maximum value to the volumetric weight of each soil layer. According to Table D.5 and Table D.6, we noticed that the soil volumetric weight can have a considerable influence, however, these are very extreme values, since all layers are weakened or strengthened at the same time.

Note that the influence of the volumetric weight is depending on the modelling and schematisation of yield stress points. From the physics, the yield stress is a measured state parameter of the soil which should not change with the analysis of increasing/different volumetric weights - approach (b) is the most suitable. However, in this particular case, the yield stress is not measured, but schematised by adding a POP value per soil to the schematised effective stress. In this particular situation, only POP values were determined with lab tests, so no local measurements of the yield stresses were available. Therefore it is justified to change the value of the yield stress points with changes in volumetric weight (and thus the schematisation of the in situ stress) - approach (a), model train, can be considered.

Considering approach (a), which uses the model train, as the most suitable for this sensitivity analysis, it was shown that changing the mean value of the volumetric weight with 10%, can have an influence on the reliability (between -0.1 and +0.6). However, if the volumetric weight is made uncertain by means of a stochastic variable with CoV of 10%, the decrease of the reliability is negligible (order -0.2, Table D.8). For this reason, it is a reasonable choice to not include volumetric weight as stochastic parameter in this specific case. Of course, this cannot be generalized.

E Case Green

This appendix presents more detailed results for the 'case green' probabilistic analysis, such as the value of each variable at the design point, the correlation coefficient achieved and the prior and posterior results considering the traffic load at the berm.

E.1 Design point prior analysis - case green

Table E.1: *Case green: calculated design points for the prior analysis.*

Variable	Design point final
tanphi ant	0.698
tanphi ant2	0.698
tanphi cal	0.698
tanphi zand	0.698
WNC intr shift	-0.0144
S KleiAnt	0.325
S KleiHum	0.264
S KleiSilt	0.247
S Veen	0.507
S VeenKl	0.404
m KleiAnt	0.876
m KleiHum	0.901
m KleiSilt	0.856
m Veen	0.876
m VeenKl	0.924
m KleiZand	0.862
S KleiZand	0.259
WNC LL out	2942
WNC LL in	1961

Variable	Design point final
1271.Yield	47.2
1273.Yield	52.0
1275.Yield	39.2
1277.Yield	26.0
1279.Yield	18.5
1281.Yield	28.7
1283.Yield	15.2
1285.Yield	26.1
1287.Yield	25.8
1289.Yield	131.6
1291.Yield	94.0
1293.Yield	90.0
1295.Yield	78.1
1297.Yield	50.7
1299.Yield	66.8
1301.Yield	123.8
1303.Yield	82.7
1305.Yield	64.7
1307.Yield	70.7
1309.Yield	54.0
1311.Yield	30.0
Model	0.984

E.2 Coefficient of correlation - case green

Table E.2: *Case green: FORM influence coefficients (α) for the base case assessment and also observation situations.*

	Assessment α_i^f			Observation α_i^p				$\alpha_i^p \alpha_i^f \rho_i^{p,f}$		
variable i	h=-0.4	h=0.4	h=1.2	h=-0.4	h=0.4	h=1.2	$\rho_i^{p,f}$	h=-0.4	h=0.4	h=1.2
tanphi_ant	0.003	0.002	0.002	0.003	0.003	0.003	1	0	0	0
tanphi_ant2	0	0	0	0	0	0	1	0	0	0
tanphi_cal	0	0	0	0	0	0	1	0	0	0
tanphi_zand	0	0	0	0	0	0	1	0	0	0
S_KleiZand_hlp	0	0	0	0	0	0	1	0	0	0
S_KleiAnt_hlp	0	0	0	0	0	0	1	0	0	0
S_KleiHum_hlp	0.024	0.024	0.024	0.024	0.024	0.024	1	0.001	0.001	0.001
S_KleiSilt_hlp	0.644	0.641	0.638	0.638	0.635	0.632	1	0.410	0.407	0.403
S_Veen_hlp	0.292	0.288	0.285	0.265	0.262	0.258	1	0.078	0.075	0.073
S_VeenKl_hlp	0.124	0.124	0.124	0.174	0.174	0.174	1	0.022	0.022	0.022
m_KleiZand_hlp	0	0	0	0	0	0	1	0	0	0
m_KleiAnt_hlp	0.045	0.046	0.047	0.038	0.038	0.039	1	0.002	0.002	0.002
m_KleiHum_hlp	0.012	0.012	0.012	0.012	0.012	0.012	1	0	0	0
m_KleiSilt_hlp	0.134	0.143	0.152	0.124	0.133	0.142	1	0.017	0.019	0.022
m_Veen_hlp	0.188	0.196	0.199	0.172	0.179	0.183	1	0.032	0.035	0.036
m_VeenKl_hlp	0.024	0.028	0.032	0.028	0.034	0.040	1	0.001	0.001	0.001
WNC.IL_shift	0.017	0.009	0.009	0.028	0.021	0.019	1	0	0	0
WNC.LL_out	0	0	0	0	0	0	1	0	0	0
WNC.LL_in	0	0	0	0	0	0	1	0	0	0
1271.Yield	0	0	0	0	0	0	1	0	0	0
1273.Yield	0	0	0	0	0	0	1	0	0	0
1275.Yield	0	0	0	0	0	0	1	0	0	0
1277.Yield	0.190	0.191	0.193	0.185	0.186	0.188	1	0.035	0.036	0.036
1279.Yield	0.033	0.033	0.034	0.034	0.034	0.034	1	0.001	0.001	0.001
1281.Yield	0.044	0.044	0.044	0.046	0.046	0.046	1	0.002	0.002	0.002
1283.Yield	0.221	0.223	0.226	0.220	0.222	0.225	1	0.048	0.049	0.051
1285.Yield	0	0	0	0	0	0	1	0	0	0
1287.Yield	0	0	0	0	0	0	1	0	0	0
1289.Yield	0	0	0	0	0	0	1	0	0	0
1291.Yield	0	0	0	0	0	0	1	0	0	0
1293.Yield	0	0	0	0	0	0	1	0	0	0
1295.Yield	0.198	0.199	0.201	0.197	0.198	0.200	1	0.039	0.039	0.040
1297.Yield	0.267	0.268	0.271	0.225	0.225	0.227	1	0.060	0.060	0.062
1299.Yield	0	0	0	0	0	0	1	0	0	0
1301.Yield	0	0	0	0	0	0	1	0	0	0
1303.Yield	0	0	0	0	0	0	1	0	0	0
1305.Yield	0	0	0	0	0	0	1	0	0	0
1307.Yield	0.213	0.212	0.212	0.272	0.272	0.272	1	0.058	0.058	0.057
1309.Yield	0.178	0.177	0.177	0.176	0.175	0.175	1	0.031	0.031	0.031
1311.Yield	0.102	0.102	0.102	0.086	0.086	0.086	1	0.009	0.009	0.009
Model unc.	0.377	0.375	0.373	0.396	0.394	0.391	1	0.149	0.148	0.146
							$\sum_i$	0.995	0.995	0.995

E.3 Prior and posterior results considering the traffic load at the berm

The road is sometimes located in the berm instead of the crest. The following graphs present the results of the same analysis carried out in section 4.3, where we consider three different values for the traffic load for both the assessment and observation conditions:

a. no traffic load (0 kN/m^2)

b. design value (13.3 kN/m^2)

c. intermediate value (7 kN/m^2, arbitrarily chosen)

Table E.3: *Case green: calculated prior reliability index (β) - assessment with different traffic load (T) considerations at the berm.*

T [kN/m^2]	$\beta[-]$
13.3	1.02
0	1.24
PMF	1.23

E.3.1 Assumption 1: No traffic load in assessment

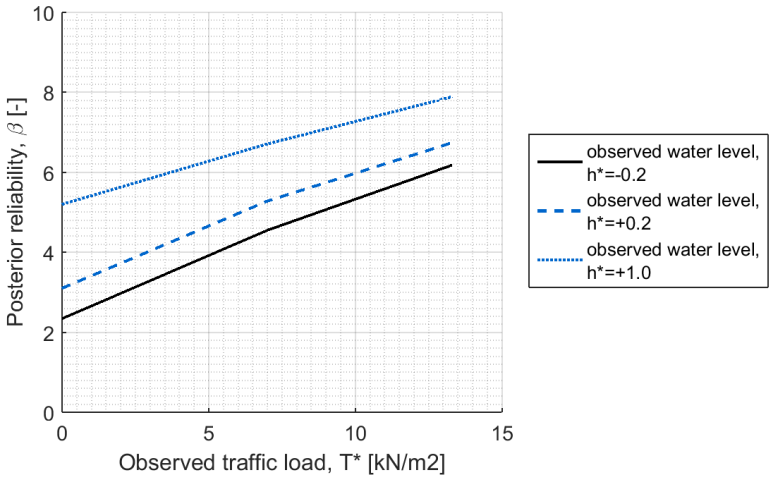


Figure E.1: *Case green: posterior reliabiltiy after RUPP with both observed water level (h^*) and observed traffic load (T*), where in the assessment situation the traffic load is T = 0 kN/m^2*

E.3.2 Assumption 2: Design value of the traffic load

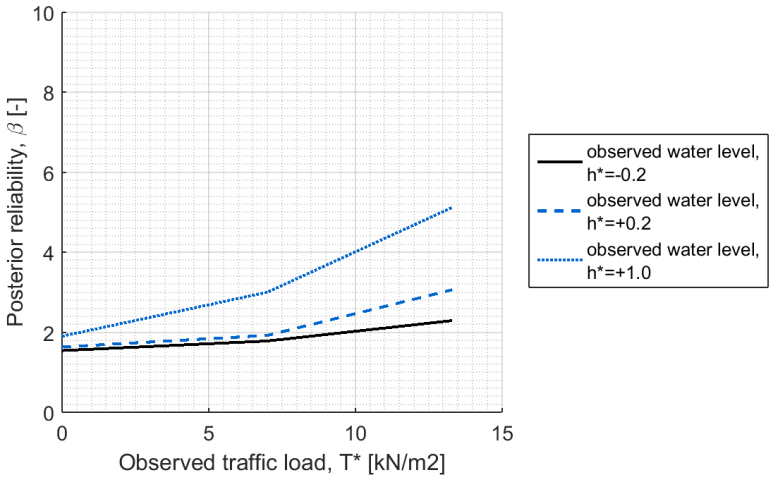


Figure E.2: *Case green: posterior reliabiltiy after RUPP with both observed water level (h^*) and observed traffic load (T*), where in the assessment situation the traffic load is T = 13.3 kN/m^2*

E.3.3 Assumption 3: Probability distribution for the traffic load

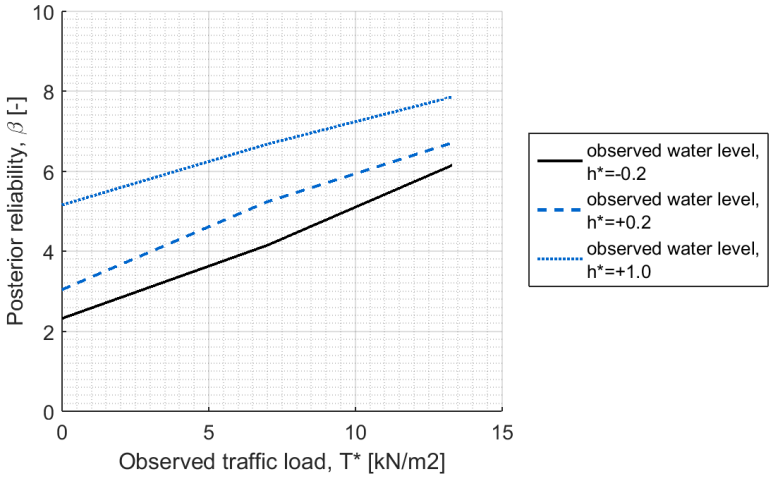


Figure E.3: *Case green: posterior reliabiltiy after RUPP with both observed water level (h^*) and observed traffic load (T*), where in the assessment situation the traffic load is considered with the PMF distribution shown in table 4.12*

F Case House

This appendix presents more detailed results for the 'case house' probabilistic analysis, such as the value of each variable at the design point, the correlation coefficient achieved and the load of a house with shallow footing.

F.1 Design point prior analysis - case house

Table F.1: *Calculated design points prior analysis.*

Variable	Design point final	Variable	Design point final
tanphi	0.698	1024.Yield	24.9
intr.shift.hlp	0.165	1026.Yield	7.92
S.KleiAnt.hlp	0.232	1028.Yield	26.6
S.KleiHum.hlp	0.185	1030.Yield	29.6
S.KleiSilt.hlp	0.191	1032.Yield	63.9
S.Veen.hlp	0.427	1034.Yield	62.0
m.KleiAnt.hlp	0.860	1036.Yield	54.7
m.KleiHum.hlp	0.898	1038.Yield	40.5
m.KleiSilt.hlp	0.830	1040.Yield	35.0
m.Veen.hlp	0.860	1042.Yield	40.1
WNC.LL.out	4710	1044.Yield	101
WNC.LL.in	3140	1046.Yield	104
Model.Factor	1.08	1048.Yield	89.0
1014.Yield	35.5	1050.Yield	80.7
1016.Yield	40.4	1052.Yield	58.8
1018.Yield	29.2	1054.Yield	63.6
1020.Yield	35.8	1056.Yield	57.0
1022.Yield	26.3		

F.2 Coefficient of correlation - case house

Table F.2: *Case house: FORM influence coefficients (α) for the base case assessment and also observation situations.*

	Assessment α_i^f			Observation α_i^p				$\alpha_i^p \alpha_i^f \rho_i^{p,f}$		
variable i	h=-0.4	h=0.4	h=1.2	h=-0.4	h=0.4	h=1.2	$\rho_i^{p,f}$	h=-0.4	h=0.4	h=1.2
tanphi	0	0	0	0	0	0	1	0	0	0
S_KleiAnt_hlp	0.261	0.260	0.198	0.223	0.207	0.130	1	0.058	0.054	0.026
S_KleiHum_hlp	0.360	0.348	0.323	0.364	0.338	0.295	1	0.131	0.118	0.095
S_KleiSilt_hlp	0.533	0.536	0.555	0.537	0.552	0.653	1	0.287	0.296	0.362
S_Veen_hlp	0.328	0.329	0.355	0.356	0.366	0.288	1	0.117	0.121	0.102
m_KleiAnt_hlp	0.057	0.067	0.078	0.051	0.060	0.062	1	0.003	0.004	0.005
m_KleiHum_hlp	0.024	0.026	0.028	0.024	0.026	0.026	1	0.001	0.001	0.001
m_KleiSilt_hlp	0.122	0.124	0.128	0.112	0.114	0.116	1	0.014	0.014	0.015
m_Veen_hlp	0.079	0.083	0.084	0.073	0.075	0.071	1	0.006	0.006	0.006
WNC.IL_shift	-0.021	-0.026	-0.049	-0.037	-0.043	-0.020	1	0.001	0.001	0.001
WNC.LL_out	0	0	0	0	0	0	1	0	0	0
WNC.LL_in	0	0	0	0	0	0	1	0	0	0
1014.Yield	0	0	0	0	0	0	1	0	0	0
1016.Yield	0	0	0	0	0	0	1	0	0	0
1018.Yield	0	0	0	0	0	0	1	0	0	0
1020.Yield	0	0	0	0	0	0	1	0	0	0
1022.Yield	0	0	0	0	0	0	1	0	0	0
1024.Yield	0.126	0.126	0.124	0.126	0.124	0.117	1	0.016	0.016	0.014
1026.Yield	0.206	0.207	0.205	0.202	0.200	0.177	1	0.042	0.041	0.036
1028.Yield	0	0	0	0	0	0	1	0	0	0
1030.Yield	0.008	0.008	0.008	0.010	0.010	0.000	1	0	0	0
1032.Yield	0	0	0	0	0	0	1	0	0	0
1034.Yield	0	0	0	0	0	0	1	0	0	0
1036.Yield	0	0	0	0	0	0	1	0	0	0
1038.Yield	0.238	0.242	0.240	0.232	0.229	0.226	1	0.055	0.055	0.054
1040.Yield	0	0	0	0	0	0	1	0	0	0
1042.Yield	0	0	0	0	0	0	1	0	0	0
1044.Yield	0	0	0	0	0	0	1	0	0	0
1046.Yield	0	0	0	0	0	0	1	0	0	0
1048.Yield	0	0	0	0	0	0	1	0	0	0
1050.Yield	0.157	0.157	0.180	0.154	0.174	0.203	1	0.024	0.027	0.036
1052.Yield	0.216	0.218	0.211	0.214	0.205	0.184	1	0.046	0.045	0.039
1054.Yield	0.212	0.214	0.221	0.213	0.213	0.222	1	0.045	0.045	0.049
1056.Yield	0.120	0.117	0.098	0.110	0.102	0.071	1	0.013	0.012	0.007
Model.Factor	0.373	0.372	0.374	0.378	0.380	0.372	1	0.141	0.141	0.139
							$\sum_i$	0.999	0.997	0.988

F.3 Load of a house with shallow footing

For an indicative calculation of the weight of a house on a shallow footing, the following basic assumptions are made:

- Size of the house (L x W) 13.5m x 6m
- First floor 2.5m high, second floor 5m high.
- Angled roof 60 degrees
- Concrete foundation slab 0.4m
- Double bond brickwork

An indicative weight calculation is shown in Table F.3.

Table F.3: *Calculated design points prior analysis.*

Element	Unit weight	Amount	Total weight
Tiled roof	$0.65kN/m^2$	$162m^2$	$105kN$
Floor construction wood	$0.30kN/m^2$	$162m^2$	$50kN$
Brickwork	$2.8kN/m^2$	$157.5m^2$	$440kN$
Foundation	$25kN/m^3$	$32.4m^3$	$810kN$
Total weight			$1.4MN$
Uniform load			$17.3kN/m^2$

G Reliability updating according to method TRAS

G.1 Introduction

These products are envisaged to complement and partially replace in the so-called TRAS (Dutch: Technisch Rapport Actuele Sterkte; ENW, 2009).

Earlier guidance on reliability updating with past performance (Calle, 2005) is reported in the technical report "Technisch Rapport Actuele Sterkte (TRAS)" (ENW, 2009), in the remainder called "TRAS method". The method is based on the same basic theory of Bayesian updating as described in the background report (Schweckendiek and Kanning, 2016), yet it applies an additional first order approximation in combining the assessment and the observation limit state. We have applied the approach described in sec.6.3 of the TRAS to 'case green' and compared the results with the proposed approximation with Fragility Curves (FC), as well as with the results obtained by directly applying Monte Carlo Simulation (MCS).

The approach requires the following steps:

Step 1: Determine the FORM influence coefficients from the prior probability of failure calculation for the observation (α_p) and assessment (α_f).

Step 2: Calculate the correlation between observation (p:past) and assessment (f:future) by:

$$\rho\left(Z_p, Z_f\right) = \rho = \sum_i \alpha_i^p \alpha_i^f \rho_i^{p,f} \tag{G.1}$$

where $\rho_i^{p,f}$ is the correlation coefficient between the variable i in assessment and observation (we assume independence between the random variables and only contemplate autocorrelation in time).

Step 3: Determine the posterior reliability index ($\beta''_{f|Z_p>0}$) - eq.(G.2), using the prior reliability index of the assessment $\left(\beta'_f\right)$, the reliability index of the observation $\left(\beta'_p\right)$ and correlation ρ:

$$\beta''_{f|Z_p>0} = \frac{1}{\sqrt{1-\rho^2}}\left(\beta'_f + \rho\frac{\varphi\left(\beta'_p\right)}{\Phi\left(\beta'_p\right)}\right) \tag{G.2}$$

where φ and Φ are the standard normal density and cumulative distribution functions, respectively.

G.2 Comparison

Two variations for 'case green' are elaborated by performing the three steps mentioned above. Afterwards, a comparison is made between the posterior reliability index according to the TRAS method and the FC method (section 4.2.4) as well as with results from a (direct) MCS (section 4.6). The following variations are included:

- Variation 1: Assessment traffic load T = 13.3 kN/m^2, observation traffic load T = 0 kN/m^2 (base case),
- Variation 2: Assessment traffic load T = 0 kN/m^2, observation traffic load T = 0 kN/m^2.

G.2.1 Variation 1 (base case)

For each fragility point in the assessment and the observation, the FORM influence coefficients conditional to water level are determined. In order to get the influence coefficients in the design point, the individual influence coefficients for the resistance parameters are interpolated at the design point of the water level and normalized for the influence coefficient of the water level (in order for the squared alpha values to sum to 1), see eq.(G.3). In Table G.1, the list of parameters with corresponding influence coefficients in the design point is shown for the observation and the assessment. Variables with alpha equal to 0 are excluded.

$$\alpha_{i,norm} = \sqrt{\frac{\alpha_i^2}{\sum_{i=1}^{n} {\alpha_i}^2}} \tag{G.3}$$

Table G.1: *Case green: FORM influence coefficients for variation 1 (base case).*

	Assessment		Observation	
Variable	α_i	α_i^2	α_i	α_i^2
tanphi_ant	0.003	0.000	0.003	0.000
intr_shift_hlp	0.015	0.000	0.027	0.001
S_KleiAnt_hlp	0.276	0.076	0.239	0.057
S_KleiHum_hlp	0.025	0.001	0.025	0.001
S_KleiSilt_hlp	0.572	0.327	0.580	0.336
S_Veen_hlp	0.292	0.085	0.266	0.071
S_VeenKl_hlp	0.126	0.016	0.177	0.031
m_KleiAnt_hlp	0.059	0.004	0.049	0.002
m_KleiHum_hlp	0.012	0.000	0.012	0.000
m_KleiSilt_hlp	0.136	0.018	0.126	0.016
m_Veen_hlp	0.185	0.034	0.171	0.029
m_VeenKl_hlp	0.025	0.001	0.029	0.001
1277.Yield	0.189	0.036	0.184	0.034
1279.Yield	0.033	0.001	0.034	0.001
1281.Yield	0.045	0.002	0.047	0.002
1283.Yield	0.216	0.047	0.216	0.047
1295.Yield	0.197	0.039	0.197	0.039
1297.Yield	0.262	0.068	0.222	0.049
1307.Yield	0.214	0.046	0.272	0.074
1309.Yield	0.177	0.031	0.176	0.031
1311.Yield	0.134	0.018	0.113	0.013
Model.Factor	0.387	0.150	0.406	0.165
Water Level	-0.020	0.000	0.000	0.000
Sum		1.000		1.000

In this case it is assumed that $\rho_i^{p,f} = 1$ for all resistance parameters, exactly as in the elaboration of the 'case green' in the report. For the water level (h) no correlation between assessment and observation is assumed, $\rho_h^{p,f} = 0$. Elaborating Equation G.1 yields ρ = 0.994. The prior reliability index of the cross section β'_f was calculated equal to 0.91 (Table 4.4). The reliability index in the observation β'_p for an observed water level h^* of -0.2m + NAP was calculated equal to 1.54 (Table 4.7). Hence, the posterior reliability becomes $\beta''_{f|Z_p>0}$ = 9.7, when applying eq.(G.2)

G.2.2 Variation 2

For variation 2, the same work-flow is followed as for variation 1. In Table G.2, the list of parameters with corresponding influence coefficients in the design point is shown for the observation and the assessment. Again, variables with alpha equal to 0 are excluded.

Table G.2: *Case green: FORM influence coefficients for variation 2.*

	Assessment		Observation	
Variable	α_i	α_i^2	α_i	α_i^2
tanphi_ant	0.003	0.000	0.003	0.000
intr_shift_hlp	0.017	0.000	0.027	0.001
S_KleiAnt_hlp	0.242	0.059	0.239	0.057
S_KleiHum_hlp	0.025	0.001	0.025	0.001
S_KleiSilt_hlp	0.584	0.341	0.580	0.336
S_Veen_hlp	0.265	0.070	0.266	0.071
S_VeenKl_hlp	0.170	0.029	0.177	0.031
m_KleiAnt_hlp	0.051	0.003	0.049	0.002
m_KleiHum_hlp	0.013	0.000	0.012	0.000
m_KleiSilt_hlp	0.134	0.018	0.126	0.016
m_Veen_hlp	0.173	0.030	0.171	0.029
m_VeenKl_hlp	0.031	0.001	0.029	0.001
1277.Yield	0.184	0.034	0.184	0.034
1279.Yield	0.034	0.001	0.034	0.001
1281.Yield	0.047	0.002	0.047	0.002
1283.Yield	0.208	0.043	0.216	0.047
1295.Yield	0.198	0.039	0.197	0.039
1297.Yield	0.224	0.050	0.222	0.049
1307.Yield	0.264	0.070	0.272	0.074
1309.Yield	0.178	0.032	0.176	0.031
1311.Yield	0.116	0.013	0.113	0.013
Model.Factor	0.404	0.164	0.406	0.165
Water Level	-0.018	0.000	0.000	0.000
Sum		1.000		1.000

Again, it is assumed that $\rho_i^{p,f} = 1$ for all resistance parameters except for the water level, where $\rho_h^{p,f} = 0$. Elaborating eq.(G.1) yields ρ = 0.9996. The prior reliability index of the cross section β'_f was calculated equal to 1.50 (Table 4.11). The reliability index of instability in the observation β'_p for an observed water level h^* of -0.2m + NAP was calculated equal to 1.54 (Table 4.7). Hence, the posterior reliability becomes $\beta''_{f|Z_p>0}$ = 59.8.

G.3 Conclusion

Table G.3 presents the results for a the two variations studied for the 'case green'. The general observation is that the TRAS method can severely over-estimate the posterior reliability if we use the MCS results as reference. Meanwhile, the approach with FC seems to give reasonably accurate results. The potential reasons and implications are discussed further in section 6.4.

Table G.3: *Case green: comparison of posterior reliability index calculated with different methods. FC: Fragility Curve, MCS: Monte Carlo Simulation, TRAS: Technisch Rapport Actuele Sterkte.*

	Traffic load [kN/m^2]		Prior β	Posterior β		
	Assessment	Observation		FC	MCS	TRAS
Variation 1 (base case)	T = 13.3	T = 0	0.91	1.14	1.38	9.7
Variation 2	T = 0	T = 0	1.50	2.52	2.77	59.8