Decisions on life-cycle reliability of flood defence systems

Klerk, W.J.

DOI
10.4233/uuid:877bed45-d775-40bb-bde2-d2322cb334f0

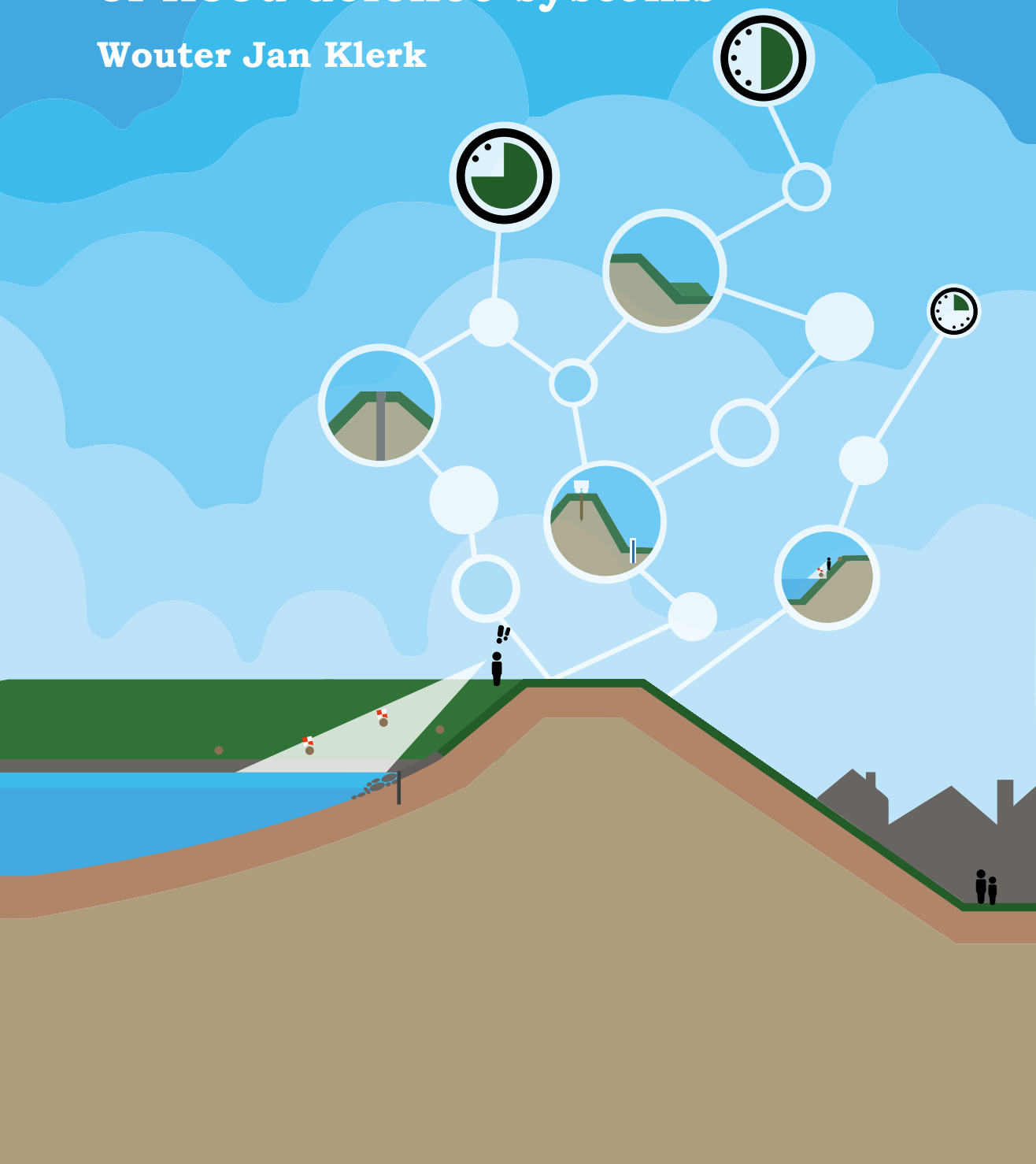
Publication date
2022

Document Version
Final published version

Citation (APA)
Klerk, W. J. (2022). *Decisions on life-cycle reliability of flood defence systems*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:877bed45-d775-40bb-bde2-d2322cb334f0

Decisions on life-cycle reliability of flood defence systems

Wouter Jan Klerk



Decisions on life-cycle reliability of flood defence systems

Decisions on life-cycle reliability of flood defence systems

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen;
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
28 maart 2022 om 15:00 uur

door

Wouter Johannes KLERK

civiel ingenieur, Technische Universiteit, Delft, Nederland,
geboren te Wageningen, Nederland.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof.dr.ir. M. Kok	Technische Universiteit Delft, promotor
Prof.dr.ir. A.R.M. Wolfert	Technische Universiteit Delft, promotor
Dr.ir. W. Kanning	Technische Universiteit Delft, copromotor
Onafhankelijke leden:	
Prof.dr.ir. S.N. Jonkman	Technische Universiteit Delft
Prof.dr.ir. R. Dekker	Erasmus Universiteit Rotterdam
Prof.dr. D. Straub	Technische Universität München
Ir. R.E. Jorissen	Rijkswaterstaat
Prof.dr.ir. P.H.A.J.M. van Gelder	Technische Universiteit Delft, reservelid

Keywords: flood defences, levees, asset management, risk-based decision making, Bayesian decision theory, inspection, maintenance, uncertainty reduction, reinforcement, optimization, reliability

Printed by: Gildeprint

Front & Back: Stefan Glas

ISBN 978-94-6384-313-3

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

All streams flow into the sea, yet the sea is never full.
To the place the streams come from, there they return again.

Ecclesiastes 1:7 (NIV)

Summary

Many countries rely on flood defence systems to prevent economic damage and loss-of-life due to catastrophic floods. Asset managers of flood defence systems need to cope with the consequences of structural degradation, and changing societal and environmental conditions, in order to satisfy performance requirements and optimize societal value of flood defence assets. This is a continuous effort of planning, executing and evaluating a variety of different system interventions. These can be aimed at both reducing the uncertainty on (e.g., inspection or monitoring), or improving the performance of a flood defence system (e.g., reinforcement). Performance is typically expressed as the reliability on a system level, which in this thesis is interpreted as the life-cycle reliability: the estimated reliability with all foreseen interventions in time. The key objective of this thesis is to improve decisions on life-cycle reliability of flood defence systems. This is elaborated for three key topics, with a focus on earthen flood defences (also known as levees or dikes):

1. Optimization of flood defence reinforcement design from a system perspective.
2. Valuation of the reduction of uncertainty in reliability estimates.
3. Managing uncertain degradation processes through Inspection & Maintenance (I&M).

Flood defences in the Netherlands have to meet requirements based on loss-of-life criteria and cost-benefit analysis of flood risk reduction for flood defence segments (i.e. a levee of approximately $20\,\mathrm{km}$). Asset managers of flood defence systems need to meet performance requirements in an efficient and effective manner, amongst others through timely and cost-effective reinforcement projects. However, for flood defence reinforcements typically an approach based on cross-sectional reliability requirements is used. As such, a system perspective is lacking.

We demonstrate that optimal allocation of resources for flood defence system reinforcements in space (i.e., different flood defence sections) and time (i.e., different times of reinforcement) significantly reduces Total Cost (flood risk and investment costs), especially compared to existing approaches based on cross-sectional reliability requirements. The size of flood defence systems, combined with the variety of potential reinforcement interventions, leads to very large optimization problems. To obtain optimal solutions, a greedy search algorithm was developed, validated and applied to an on-going reinforcement project in the Netherlands. The approach allows for inclusion of a variety of reinforcement measures, such as diaphragm walls, berms and stability screens. The result is a practical approach that leads to a significant reduction of reinforcement costs (≈40 % in the considered case), while

reliability requirements at a system level are still satisfied.

A key aspect of flood defence asset management is dealing with uncertainty in reliability estimates. We distinguish both time-invariant uncertainty in parameters, and uncertainty in time-variant processes: the first mostly concerns (geotechnical) strength parameters, while the latter mainly concerns degradation processes.

Reduction of time-invariant uncertainty is evaluated through Bayesian pre-posterior analysis. We consider proof load testing to reduce uncertainty in strength parameters, combined with pore pressure monitoring aimed at reducing uncertainty in pore water pressures, in the context of a flood defence reinforcement. Additionally we consider reducing uncertainty on aquifer permeability (for piping erosion), and its effects on optimal system reinforcement. It is demonstrated that the Value of Information of reducing time-invariant uncertainty depends on a variety of factors, such as the available methods for reinforcement (at a system level), whether and in what sequence different uncertainty reduction methods are combined, and the reliability requirement that has to be satisfied.

Uncertainty in time-variant degradation processes (e.g., animal burrowing) is often reduced through visual inspection. The accuracy of visual inspection is unknown, but important: if damage due to degradation processes remains undetected, flood defence reliability during a flood event is worse than expected. The accuracy and consistency of condition inspections is evaluated in an experiment where 4 flood defence sections are inspected by 22 inspectors in 14 rounds. The accuracy (Probability of Detection) differs strongly per inspector (e.g., 20-80% for essential damage spots) and per type of damage. The consistency is found to be limited, especially for the severity of damage: in over half of the registrations the reported severity was different from the reference. Key reasons for inaccuracy and inconsistency are complexity of the inspection task and ambiguity in inspection guidelines (unclear classification, overlapping parameters).

Finally, this thesis investigates the impact of degradation and Inspection & Maintenance (I&M) on the reliability of a flood defence segment. Degradation data from inspections suggests that most degradation (e.g., animal burrowing and bare spots) is caused by random shock-based processes. Using a Dynamic Bayesian Network we investigate the effect of degradation and a typical maintenance concept on flood defence segment reliability. Degradation is found to increase the segment failure probability by an order of magnitude compared to a standard assessment without considering damage. Next we assess how other maintenance concepts and structural upgrades influence structural robustness and Total Cost. Additional inspections in autumn are identified as a cost-effective way to reduce the impact of degradation, while structural robustness can be further increased by targeted measures to mitigate the effects of for instance animal burrowing. Finally it is investigated what requirements are to be imposed to new inspection techniques such as drone observations to outperform existing visual inspections.

For each of the three considered key topics the approaches from this thesis have been demonstrated to both increase efficiency and effectiveness of decisions, and increase transparency in decisions. As such this thesis can contribute to a better utilization of a risk-based approach for asset management of flood defence systems.

Samenvatting

Veel landen vertrouwen op systemen van waterkeringen om economische schade en verlies van mensenlevens als gevolg van grootschalige overstromingen te voorkomen. Beheerders van waterkeringen hebben te maken met de gevolgen van degradatie, veranderende maatschappelijke eisen, en verandering van de omgeving. Om aan geldende prestatie-eisen te kunnen voldoen, en de maatschappelijke waarde van waterkeringen te optimaliseren, vereist dit een continue inspanning van het plannen, uitvoeren en evalueren van veel verschillende soorten maatregelen. Zulke maatregelen kunnen gericht zijn op het verminderen van onzekerheid over (bijv. door inspecties of monitoring), of het verbeteren van de prestatie van een waterkering (bijv. door dijkversterking). Meestal wordt prestatie uitgedrukt in termen van systeembetrouwbaarheid, in dit proefschrift wordt dat geïnterpreteerd als de levenscyclusbetrouwbaarheid: de geschatte betrouwbaarheid, met alle geplande ingrepen, in de tijd. Het hoofddoel van dit proefschrift is het verbeteren van beslissingen over levenscyclusbetrouwbaarheid van waterkeringsystemen. Dit is uitgewerkt voor drie hoofdonderwerpen, met een focus op dijken:

1. Optimaliseren van dijkversterkingen op systeemniveau.
2. Het waarderen van reductie van onzekerheid in betrouwbaarheidsschattingen.
3. Omgaan met onzekerheid in degradatie door inspecties en onderhoud.

In Nederland zijn deze betrouwbaarheidseisen van dijken gebaseerd op een maatschappelijke kosten-baten analyse, en criteria voor aantallen slachtoffers. Deze eisen zijn gedefinieerd op systeemniveau, bijvoorbeeld voor een dijktraject van 20 kilometer. De hoofdtaak van beheerders van waterkeringen is om op effectieve en efficiënte wijze te voldoen aan (vaak wettelijke) prestatie-eisen, bijvoorbeeld door tijdige en kosteneffectieve dijkversterkingsprojecten. Echter, voor dijkversterkingen wordt vaak een aanpak gehanteerd gebaseerd op doorsnede-eisen, waardoor het systeemperspectief op dergelijke projecten ontbreekt.

In dit proefschrift wordt aangetoond dat het optimaal verdelen van middelen voor dijkversterkingen in de ruimte (over verschillende dijkvakken) en in de tijd (verschillende tijdstippen van dijkversterking) tot een grote reductie van totale kosten leidt, vooral in vergelijking tot de in de praktijk gebruikte methoden gebaseerd op doorsnede-eisen voor de betrouwbaarheid. De omvang van dijktrajecten, in combinatie met de variëteit aan mogelijk versterkingsmaatregelen leidt tot zeer grote optimalisatieproblemen. Om optimale oplossingen te verkrijgen is een zoekalgoritme ontwikkeld, gevalideerd en toegepast op een actueel dijkversterkingsproject in Nederland. De aanpak maakt het mogelijk om zeer veel verschillende versterkingsmaatregelen zoals diepwanden, bermen en stabiliteitsschermen mee te nemen. Het

resultaat is een praktische aanpak die leidt tot een grote reductie van dijkversterkingskosten (ongeveer 40% voor het beschouwde project), terwijl alsnog aan de betrouwbaarheidseisen wordt voldaan.

Een belangrijk onderdeel van asset management van waterkeringen is omgaan met onzekerheid in betrouwbaarheidsschattingen. We maken onderscheid in tijdsonafhankelijke onzekerheden in parameters, en onzekerheid in tijdsafhankelijke processen. Het eerste gaat vooral over (geotechnische) sterkteparameters, het tweede over degradatieprocessen.

Keuzes voor het reduceren van onzekerheden in parameters kunnen worden afgewogen met Bayesiaanse pre-posteriori analyse. In dit proefschrift wordt, in de context van een dijkversterking, onderzocht of proefbelasting om onzekerheid in sterkteparameters te verkleinen, al dan niet gecombineerd met waterspanningsmetingen om de onzekerheid in waterspanningen te verkleinen, een effectieve strategie is. Aanvullend wordt een soortgelijke analyse gedaan voor het verkleinen van onzekerheid in de doorlatendheid van het watervoerende pakket (t.b.v. falen door piping), en de invloed op de dijkversterking op systeemniveau. Het is aangetoond dat de waarde van informatie van het verkleinen van dergelijke onzekerheden afhangt van een aantal factoren, zoals de beschikbare dijkversterkingsmethoden (op systeemniveau), of en in welke volgorde de methoden voor het verkleinen van onzekerheden worden gecombineerd, en aan welke betrouwbaarheidseis moet worden voldaan.

Onzekerheid in tijdsafhankelijke degradatieprocessen (bijv. dierlijke graverij) wordt vaak gereduceerd door visuele inspecties. De nauwkeurigheid van dergelijke inspecties is niet bekend, maar wel van belang: als schade door degradatie niet wordt waargenomen is de betrouwbaarheid van de waterkering tijdens een hoogwater lager dan verwacht. De nauwkeurigheid en consistentie van inspecties is onderzocht in een proef waar 4 dijkvakken zijn geïnspecteerd door 22 inspecteurs in 14 inspectierondes. De nauwkeurigheid (detectiekans) verschilt sterk per inspecteur (bijv. 20-80% voor de meest essentiële schadepunten) en per soort schade. De consistentie is beperkt, vooral voor de ernst van schade: in meer dan de helft van de meldingen was de gerapporteerde ernst anders dan de referentie. Belangrijkste oorzaken voor inconsistentie en onnauwkeurigheid zijn de complexiteit van de inspectie als geheel, en de onduidelijkheid in de gehanteerde inspectierichtlijnen (bijv. onduidelijk klasse-indelingen en overlappende parameters).

Tot slot wordt gekeken naar de invloed van degradatie, inspectie en onderhoud op de betrouwbaarheid van een waterkeringsysteem. Degradatiedata van inspecties laat zien dat de meeste degradatie (bijv. dierlijke graverij en kale plekken) veroorzaakt wordt door 'schokken' die willekeurig in de tijd optreden. Met behulp van een Dynamisch Bayesiaans Netwerk is gekeken naar de invloed van degradatie en bestaande inspectie- en onderhoudsregimes op de betrouwbaarheid van waterkeringen. Alles bij elkaar is het effect hiervan dat de systeemfaalkansen met een orde grootte toenemen in vergelijking tot een gebruikelijke beoordeling waarbij dit niet wordt meegenomen. Vervolgens is gekeken naar de invloed van andere inspectie- en onderhoudsregimes en/of constructieve aanpassingen op de totale kosten en constructieve robuustheid van het systeem. Extra inspecties in het najaar blijken

een kosten-effectieve manier om de invloed van degradatie te beperken, en de constructieve robuustheid kan verder worden vergroot met gerichte maatregelen om de gevolgen van schade in te perken, zoals maatregelen gericht op het beperken van de gevolgen van dierlijke graverij. Tot slot is beschouwd welke eisen moeten worden gesteld aan nieuwe inspectietechnieken zoals drones, zodanig dat ze beter presteren dan de bestaande visuele inspecties.

Voor elk van de drie hoofdonderwerpen leiden de analyses uit dit proefschrift tot betere en meer herleidbare beslissingen. Zodoende kunnen de bevindingen uit dit proefschrift bijdragen aan het beter benutten van een risicobeandering voor het asset management van waterkeringsystemen.

Acknowledgments

Als er een bever in de dijk bij school graaft, dan hebben we een heel groot probleem!

Amélie Klerk (5 jaar)[1]

Acknowledgements in a PhD thesis can never cover everyone's contribution, but at most highlight some of the most influential ones. First of all, Matthijs, thanks for the opportunity. I'm glad that we could seize the momentum of AllRisk, resulting in this thesis, which carries one name as author, but also the fingerprints of your comments, additions and suggestions during the many conversations we had.

Talking about fingerprints, Wim thank you for the many hours that we met via Skype and Teams. With you as supervisor living in Geneva from the start of my PhD, we were definitely well prepared for 'working from home', with abundant experience in online meetings. Obviously about content, but also about riding bikes (you even bought a road bike during this period!) and many other things. Whenever I wrote something you responded in detail, and helped me focus when required. An invaluable contribution!

Rogier, you compensated your relative inexperience on how the Dutch deal with flood defences (which is a peculiar business) with lots of practical advice on how to apply systems thinking in this thesis. A contribution that is impossible to grasp from literature. Thanks a lot!

One of the things that I'm most proud of in this thesis is the connection to practice. However, as researcher you can never make this work without people who see opportunities and dare to take a leap. The contents of Chapter 3 would never have been the same without Nelle-Jan van Veen, who stuck out his neck and was a tremendous sparring partner and end user within the SAFE project. Similarly, Chapter 6 would have never materialized without the enthusiasm and perseverance of Jaap Bronsveld. Many thanks to both of you!

In the first years, before a forced retreat to the attic, lots of work was done in the great atmosphere of room 3.58 at the TU. Chris, thanks for the self imported rum from Barbados, Arny, for the chats about your crazy long mountain walks which make my cycling endeavours modest at best. Joost, thanks for scheduling me for the graveyard shift at your test at Flood Proof Holland. I maybe wasn't too enthusiastic at the time, but it is a lifetime memory! Mark, thanks for the great collaboration and I sincerely hope we can continue this at Deltares.

[1] If a beaver burrows in the levee near school, we have a major problem!

Deltares has been really supportive throughout these years. Specifically Frank den Heijer, thanks for stimulating my development in the projects we did on asset management and encouraging me to take it a step further in this PhD. Annemargreet and Judith, for the support before and during my PhD. Andre van Hoven, for the projects on revetments that helped me materialize some of the ideas that are also in this thesis. And of course Ferdinand for his n weekly check-up calls during lockdowns.

Obviously there are many other colleagues who deserve a mention as well: special thanks to Vera for the collaboration on the last chapter, and to AllRisk in general for the nice and stimulating atmosphere within the programme. Juan Pablo, thanks for the passionate discussions about inspections and inspection data, and Timo for fundamental discussions about utility formulations, I'm sure we can conclude those some day. And of course to all those colleagues who helped out with the field test at Tiel: Robert, Stephan, Danny, Mark, Joost and Ligaya from the TU. Richard, Matthijs and Kim from AllRisk. This could not have been done without you!

Maarten, thanks for the discussions (sometimes rants) about flood defences and how things are but should not be done. I'm sure we can have lengthy ones soon again! Rens, thanks for the great local experience in Hanoi, but great to have you back: it is much easier to have a 'BBQ'tje at night' now.

Pa en ma, dank voor jullie steun tijdens alle jaren van naar school gaan, studeren en verder, en in het bijzonder ook voor het oppassen in de laatste maanden van het onderzoek toen de tijdsdruk hard toenam. Het is een luxe om van jongs af aan zo gesteund en gestimuleerd te worden, zeker geen vanzelfsprekendheid.

Het beste van dit dankwoord heb ik voor het laatst bewaard. Elzelien, wat heb je me enorm maar dan ook enorm geholpen en gesteund in de afgelopen jaren. Dank je, voor het zijn van 'thuiscollega'. Dank je voor de waarschuwingen als je mijn stressniveau voor ik het zelf door had zag toenemen. Dank je, dat je technische monologen over kansen, updates en lengte-effecten braaf knikkend aanhoorde. Dank je, dat je me aanmoedigde, steunde, je armen om me heen sloeg als de route naar de eindstreep ineens wel heel ingewikkeld leek. Sorry, dat ik mezelf soms overschat als ik bedenk dat ik naast een proefschrift schrijven ook wel even de zolder eigenhandig kan gaan verbouwen. Het resultaat is een proefschrift en een zolder, maar ik moet je voor deze ene keer toch gelijk geven dat het alles bij elkaar misschien wat veel was :)

Amélie, wat ben jij ontzettend groot geworden. Toen ik begon met het schrijven van mijn boek kon je net lopen. En nu leg je me uit over de dijk bij school en kun je dit stukje zelf lezen. Wat heb jij ontzettend veel geleerd in deze tijd, veel meer dan papa in dit boekje kon opschrijven. Joas, kleine eigenwijze krullenbol. Papa moppert altijd wel als je weer komt vragen of er heel misschien 'taatuus' gekeken kunnen worden terwijl papa aan het werk is, maar stiekem vind ik brandweerauto's ook veel mooier dan programmeercode! Rhodé, wat kan ik genieten van je gulle lach, je onschuldige heldere ogen en hoe je onbevangen de wereld ontdekt. Wat ben ik enorm gezegend met jullie!

Wouter Jan Klerk
Delft, February 2022

Contents

1

Introduction

1

1.1. Background and rationale

Increasing population and economic value in rapidly urbanizing delta areas have increased global flood damages over the past decades [1]. Over time, flood defence systems have been constructed in order to protect over half a billion people who live in vulnerable areas around the world [2]. Such systems typically consist of earthen levees, sandy coasts and storm surge barriers that shield low-lying regions from flood hazards. While these vitally important systems have provided protection and enabled significant socio-economic growth in deltas, their failure can lead to catastrophic flooding.

To prevent such disasters, asset managers (e.g., water authorities or levee districts) are tasked with ensuring that requirements for flood defence system performance are met throughout their life cycle, with the available public funds. However, the unpredictable effects of climate change may increase flood probability, existing assets may degrade and become unreliable, and societal developments might result in increasingly stringent flood defence system performance requirements.

Generally speaking, the key objective of flood defences is to reduce flood risk. This is the case in all countries where flood defence assets are present [3]. Nevertheless, the approach towards achieving this goal, and balancing the performance of flood defences with costs differs per country. For instance, while the UK uses a benefit-cost based appraisal to flood risk reduction investments, in the USA funding is arranged based on a 'levee safety action classification' [3]. Both are implementations of the key principle of asset management, which ISO55000 defines as the continuous 'coordinated effort to realize value from assets' [4], which is also a key principle for decision making used in this thesis. In the Netherlands, whilst investments are prioritised based on their estimated effectiveness in terms of reducing flood risk, performance requirements for flood defences are also defined by law since 1996. Up to 2017 annual exceedence probabilities of hydraulic loads were the basis for performance requirements in the Netherlands. After 2017, the requirements for the Dutch flood defences were revised and changed to probability of flooding or failure probability requirements, based on a national assessment of loss-of-life risk, and societal costs and benefits [5]. This has resulted in the definition of target reliability requirements for flood defence segments (with a length of 5-25 kilometres) in the Netherlands, and has been accompanied by updated assessment manuals in order to assess flood probabilities using (semi)probabilistic assessment methods [6], thus connecting the Dutch approach to the practice of structural reliability analysis.[1] It should be noted that while requirements are defined for flood defence segments, reliability analysis and design decisions are typically considered at the level of a flood defence section of approximately 1 km length.

In reliability analysis one has to account for all uncertainties, both in strength and load parameters [7]. This means that asset management efforts to reduce uncertainty, such as monitoring campaigns and inspections, can have a direct influence on reliability estimates [8]. This is of importance to the asset manager, as not meeting the performance requirement is not necessarily due to the flood defence

[1] Within the context of this thesis flooding probability, failure probability and reliability are interchangeable terms unless stated otherwise.

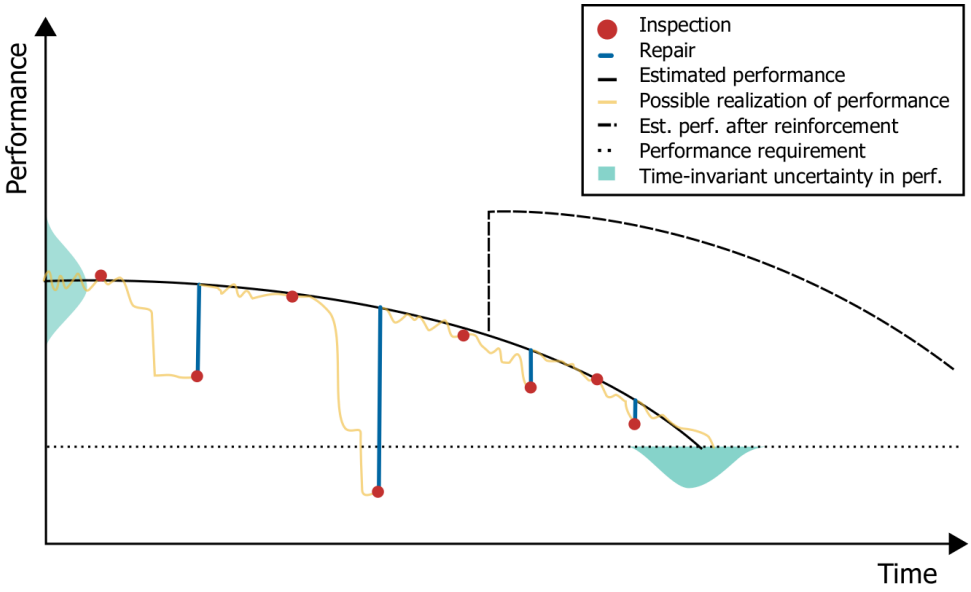


Figure 1.1.: *Schematic representation of life-cycle performance. Black line denotes the performance as typically considered for long term investments, whereas the yellow line shows a possible realization such as dealt with in inspection (red dots) and maintenance (repairs represented by blue vertical lines). Life-cycle performance is also subject to time-invariant uncertainty (green distribution). In time, opportunities might arise for reinforcement (black dashed line).*

being insufficiently safe, but can also be caused by a lack of information on for instance the strength of a flood defence. Consequently, the action perspective of flood defence asset management has changed with the new approach.

The general principles of asset management as outlined in ISO55000 can be applied at different levels of decision-making — for instance, when determining efficient performance requirements or making choices about inspection and maintenance [4]. However, it is vitally important to align these different levels in order to optimize the engineering asset management process, and add the concreteness that lacks in general guidelines such as the ISO55000. One recent effort on this front is the development of the FAIR framework for adaptive asset management of flood protection [9]. Here the importance of accounting for the interdependence between strategic decisions on e.g., flood defence requirements and the required operational actions (e.g., the frequency and intensity of inspections) has been highlighted, but was also found to be missing in practice [10].

The life-cycle performance is the (expected) performance of a structure (or structural system) in time, with all scheduled interventions [11, 12][2], which is illustrated in Figure 1.1 by the black line. Here it is shown that a variety of factors influence

[2]It should be noted that this also includes interventions as part of for instance an I&M policy. E.g., 'if asset condition is x, do a repair'.

life-cycle performance, amongst others the presence of time-invariant uncertainty in performance estimates (turquoise distribution). A variety of degradation processes might lead to a specific realization of performance (yellow): shock-based degradation processes can lead to sudden drops in performance and continuous processes to a gradual decline [13]. The effects of these processes on performance are (possibly) detected in inspections (red dots) and resolved by repair works (blue lines). Additionally, intermediate larger interventions such as reinforcements can be done (dashed black line), either if requirements are no longer satisfied or if other opportunities for intervention arise (e.g., alignment with other activities [14]).

In a performance analysis, only a part of these processes and uncertainties might be included, and not all available actions might be considered. From an asset management perspective there are two key points that lead to efficient and effective strategies for balancing costs, risks and performance of flood defence systems: First, we need to ensure that all relevant uncertainties are accounted for in the decision analysis. Or, in the context of Figure 1.1, we should substantiate whether our approximation of the performance of a flood defence (system) is sufficiently accurate. Secondly we need to consider all relevant actions, both for a flood defence section and the system as a whole, and substantiate why certain decisions are in or out of the scope of a decision analysis. The key theme of the topics considered in this thesis is to underpin decisions on interventions to improve and/or to reduce uncertainty on life-cycle reliability of flood defence systems.

1.2. Key topics

This thesis focuses on three key topics where current (practical) limitations hamper efficient and effective asset management of flood defence systems. More specifically, as explained in the chapters that follow, each topic relates to inefficiencies or unknowns in the practical translation of performance (requirements) to decisions.

1. Optimization of system design
   While ample effort has been put into optimizing the safety targets of flood defence systems [5, 15–18] and the design of sections based on a representative cross-section [19], limited attention has been paid to optimizing system design in the context of flood defence reinforcement projects. A typical approach to translate system reliability targets to reliability targets for sections is to use standardized failure probability budgets. This can lead to inefficient investments in flood defence reinforcements. By developing an approach for optimizing design at a flood defence system level, the translation of the performance requirement (black dotted line in Figure 1.1) from segment to section level will be optimized in space and time, and as such the efficiency of investments in flood defence reinforcements will be improved.

2. Reduction of time-invariant uncertainty
   If uncertainties are considered explicitly, reduction of uncertainty is a potential means to improve reliability estimates [e.g. 20, 21]. Time-invariant uncertainties in flood defence reliability are for instance uncertainties in soil parameters,

which are typically fully correlated in time. In the context of both flood defence reinforcement projects and of asset management in general, investments in reducing such uncertainties through for instance Structural Health Monitoring (SHM) can yield large benefits for structural integrity [21–23]. However, even if reducible uncertainties are dominant in reliability analysis, reduction of uncertainty is not always considered. Many reliability estimates of flood defences are dominated by reducible time-invariant uncertainties in strength parameters. By including investments in reduction of time-invariant uncertainty in flood defence asset management decisions, timing and efficiency of investments can be improved through the reduction of performance uncertainty (denoted by the green distribution in Figure 1.1).

3. Managing uncertainty in time-variant processes
   Regular inspections are essential to manage uncertainty in time-variant processes, most notably degradation. While such uncertainties might not be fully reducible, inspections are typically required to mitigate their influence on reliability. The importance of (risk-based) optimization of Inspection & Maintenance (I&M) efforts has been demonstrated for many applications [e.g. 24–27]. However, I&M of flood defence systems, in relation to system reliability targets, has received limited attention. As a result, the design and assessment of flood defences is not well-connected to I&M [10], and only a limited selection of time-variant processes is considered. This is a potential problem, as some types of degradation for flood defences can cause large (and sudden) drops in performance (e.g. animal burrowing [28]). By elaborating the relation between I&M and the reliability estimates from for instance safety assessments, these can be made more consistent with reality (yellow and black lines in Figure 1.1).

By advancing on the above topics we can improve the asset management of existing flood defence systems. It should be noted that also for large future challenges such as adaptation to climate change or large socio-economic developments each of the above points is of importance: system optimization will result in lower investment costs in general, explicitly considering reduction of uncertainty will lead to flood defence systems that are understood better, and connecting inspection and maintenance to system reliability targets will result in a better alignment of investment decisions and actual performance. As such, by better understanding our existing systems, adaptation of these systems will become more efficient, and the technical feasibility of adaptation will also increase.

1.3. Objective & outline

The main objective of this thesis is as follows:

Improving decisions on interventions in flood defence systems subject to uncertain future performance.

For each of the key topics in this thesis we will both develop the necessary tools to deal with the issues outlined in the previous section, and in relation to that answer the following questions:

1. How can we optimize investment planning of flood defence system reinforcements?
2. How can we incorporate uncertainty reduction in flood defence (system) reinforcement decisions?
3. What are the long-term benefits of uncertainty reduction for flood defence asset management?
4. How accurate are visual inspections of flood defences?
5. How can we account for Inspection & Maintenance in reliability estimates, and how does this translate to asset management decisions?

Figure 1.2 presents the general structure of this thesis. After this introduction, Chapter 2 will set the scene for this thesis with an outline of the key concepts and methods used, along with some substantiation of assumptions made in subsequent chapters. After this general background, each of the Chapters 3 to 7 will answer one of the five research questions, related to the key topics (KT) which are also shown in Figure 1.2. Chapter 3 presents a methodology for optimizing dike reinforcement projects at a system level (KT1). Chapter 4 considers the influence of reducing time-invariant uncertainty (KT2), and connects this to optimal design for both a dike section and a dike segment (KT1). While Chapter 4 considers investments for a time horizon of 50 years, Chapter 5 considers the benefits of reducing time-invariant uncertainty (KT2) for investments at a dike section for a period of 200 years (KT1). Uncertainty in time-variant degradation processes of flood defences are monitored through visual inspections. In Chapter 6, we investigate the accuracy of these visual inspections (KT3), in order to determine the probability that degradation remains undetected. This is an important input parameter for Chapter 7 where we investigate the influence of Inspection & Maintenance (KT3) on asset management strategies for flood defence segments. In particular we jointly evaluate the effects of decisions on I&M and structural adaptation of the flood defence structure (KT1). Chapter 8 presents key conclusions, recommendations for future research and engineering practice, as well as some more general closing remarks. This thesis will mostly consider cases that originate from Dutch practice, but as the core principle is risk-based decision analysis it is applicable to other practical contexts as well.

1.4. Contribution

The main contribution of this thesis is that it improves different asset management efforts in the context of life-cycle reliability on three key topics. The first key topic is that, in practice, relatively simple approaches are used for determining performance

requirements for design in flood defence reinforcements. In Chapter 3 we present a method for optimizing flood defence reinforcement projects of large flood defence systems, which results in significantly lower total life-cycle cost. The method is practically applicable and can be directly utilized to optimize for instance the approximately 400 million € per year invested in flood defence reinforcements as part of the Dutch Flood Protection Program (HWBP).

Life-cycle costs are strongly influenced by uncertainties in flood defence performance, which can be both on time-variant processes and time-invariant parameters. The second key topic is how reduction of time-invariant uncertainty can be considered in flood defence design and investments. Chapter 4 considers efforts to reduce time-invariant uncertainty in the context of flood defence reinforcements, and Chapter 5 in the context of long-term investments (multiple reinforcements). The methods developed in these chapters can be used to incorporate investments in uncertainty reduction in flood defence reinforcement decisions, as well as for long-term flood defence asset management strategies.

The third key topic is that many uncertainties in time-variant processes are not considered in reliability estimates of flood defence systems. This mainly concerns uncertainty in de occurrence of damage to a flood defence (e.g., through animal burrowing). Reduction of uncertainty on such time-variant processes is mostly done through visual inspection. In Chapter 6 we present the first quantification of the accuracy of flood defence inspections. In Chapter 7 we translate this to reliability estimates for a flood defence segment, and demonstrate that including such uncertainty is pivotal in obtaining accurate reliability estimates. With the developed approach, combinations of Inspection & Maintenance policies and structural upgrades can be evaluated.

1

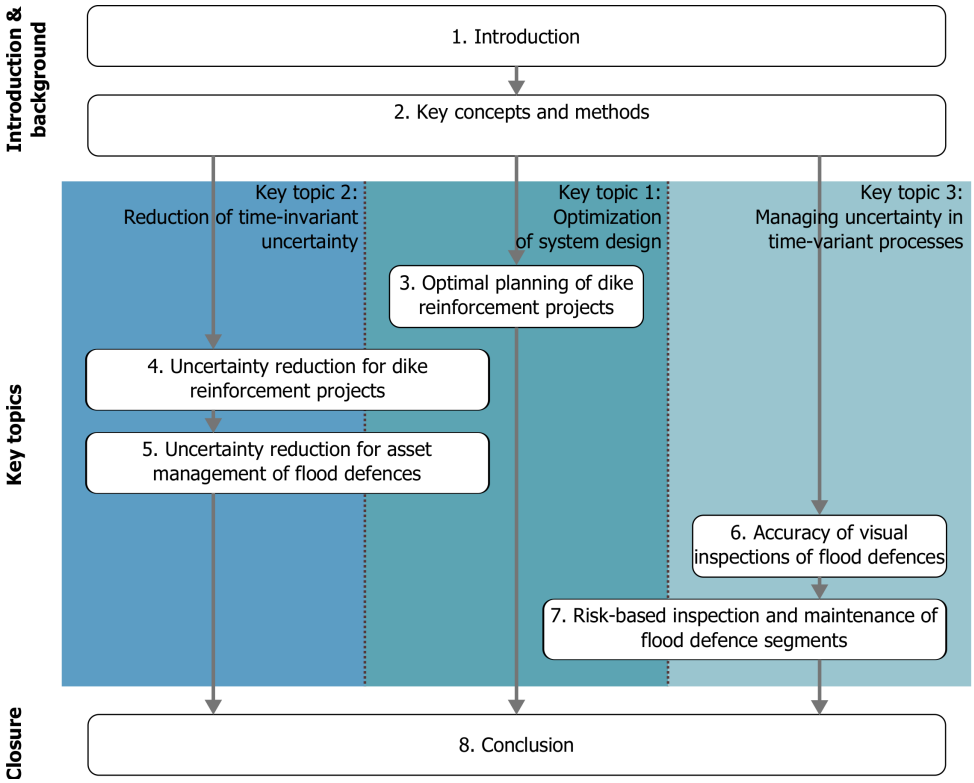


Figure 1.2.: *General outline of this dissertation. Each chapter is related to one of the research questions and (a combination of) the key topics. Descriptions in boxes are chapter titles.*

2

Key concepts and methods

The framework is science, the filling-out art, the whole being rounded off by science in the form of the computer calculating expectations and searching for maxima.

Dennis V. Lindley

This thesis revolves around the three key topics outlined in Chapter 1, and how improvement of these can lead to more efficient and effective flood defence asset management. This chapter outlines the key concepts and methods that are used in this thesis. The key focus of this chapter is how general asset management concepts (Section 2.1) relate to performance requirements for flood defence systems (Section 2.2). Next we discuss reliability estimates of flood defence systems: first we discuss the different types of uncertainty that might impact reliability estimates (Section 2.3). Subsequently we substantiate some key assumptions on the translation of reliability estimates in space and time (Sections 2.4 and 2.5) that are made in this thesis.

Section 2.6 considers a taxonomy with along which different interventions for flood defences can be distinguished. Section 2.7 introduces decision theory, and most notably Bayesian decision theory, which will be a key concept for evaluating asset management strategies in many chapters of this thesis. Sections 2.8 and 2.9 provide some general considerations on how to utilize decision theory for the key topics in this thesis.

Parts of Section 2.7 are published in Klerk, W. J. et al. Value of Information of Structural Health Monitoring in Asset Management of Flood Defences. *Infrastructures* **4,** 56. doi:10.3390/infrastructures4030056 (2019)

2.1. Asset management of flood defence systems

The key goal of engineering asset management, is to optimally balance the (life-cycle) performance, cost and risk of engineering assets. For flood defences, this means that an optimal balance between flood risk costs, costs of interventions and performance levels is sought, such that both (statutory) performance requirements are met and value to users/stakeholders is maximized [9, 30, 31].

Interventions as part of an asset management strategy can be very diverse, and can be done both at system and component level. For flood defence systems, an investment in a storm surge barrier can be considered a system intervention, while reinforcement of a specific flood defence section is an intervention at a single component of a larger system. Each of these interventions contributes to the goals of flood defence asset management, which are typically defined on a system level: i.e., maximizing system value and meeting performance requirements. Vonk et al. [9] identified that this involves different contexts (i.e., strategic, tactical and operational) of decisions, of which the alignment is typically lacking in practical implementation. This is confirmed by the state-of-practice at some of the key topics discussed in this thesis, for instance the lack of system optimization in design, and the disconnect between Inspection & Maintenance (I&M) interventions and reliability estimates (see Chapter 1).

Quantitative analysis of asset management decisions often take an analytical thinking approach, where a system is treated as 'a whole to be taken apart' [14]. In this reductionist approach it is assumed that solutions for parts of a system translate to a solution for the whole [32]. A system thinking approach considers a part of a system as 'a part of the containing whole'. This means that for assessing the performance and effect of interventions the relation of system elements are of equal or more relevance than the performance of the individual parts. For flood defence asset management, Sayers et al. [10] identifies that a major issue is that each institutional silo optimizes their own decisions. By adopting a system thinking approach, one takes into account that maximizing the performance of individual parts, might not lead to the optimal performance of the system [32]. In this thesis we will come across a number of cases where this is the case (a.o. Chapters 3 and 7). However, first we explore some basic concepts of flood defence systems, and their performance.

2.2. Performance requirements for flood defence assets

There are several indicators with which performance of engineering assets can be measured. The most common ones are reliability, availability, maintainability and safety [33]. Reliability is the ability to perform a specific function over a given time period, availability is the ability of an asset to remain 'usable over a period of time, maintainability indicates how easily the performance can be restored or inspected and safety indicates the probability that an asset will harm people. The key objective of flood defence systems is to reduce the risk from flooding, and this has to be realized for often lengthy periods of time. As such, the key performance measure for flood defence systems is their reliability. It should be noted that while in this thesis we mainly look at reliability, other performance indicators such as environmental impact can also influence optimal decisions.

The sophistication of methods used for determining flood defence reliability requirements has developed significantly over time, and led to a variety of similar indicators in different countries [see e.g. 3]. This development is nicely illustrated by the development of requirements for flood defences in the Netherlands. Whereas in the early 20th century flood defence height was based upon whether it survived the most recent storm, after the 1953 flood, performance requirements were based upon a risk-based consideration (i.e., probability times consequences) [34]. This has led to requirements formulated as exceedence frequencies of extreme loads — in the context of the legal status they obtained as part of the 1995 Flood Defence Act this was combined with assessment rules that ensured sufficient performance for loads with such exceedence frequencies. However, increases in population and economic value made the risk-consideration outdated, and as technical advancements now enabled further computation and optimization of risk-based standards, in 2017 the Water Act[1] was revised, and new risk-based standards were introduced. These standards should be evaluated considering all relevant uncertainties that contribute to flood defence failures [35]. However, flood risk is multi-dimensional — aside from various quantitative measures for loss of life and economic damage, also environmental risk indicators can be considered [36]. For the flood defence standards in the Netherlands, 3 types of risk indicators are considered [5, 37]:

- Individual risk: the risk of a single individual dying from a flood.
- Societal risk: the risk of large numbers of people dying in a flood event.
- Societal cost-benefit analysis: the optimal balance between cost of flood protection and (societal) risk from flooding.

Based on the risk analysis for these three criteria, requirements in terms of allowable probability of flooding (or failure probability) have been derived for flood defence segments (5-25 kilometres length).[2] The maximum allowable flooding probabilities for the Netherlands are displayed in Figure 2.1.

[1] Note that the Water Act was the successor of the Flood Defence Act.

[2] In this thesis we will refer to flood defence segments as flood defence (or levee/dike) systems that can be represented by a series system of independent elements (dike sections). We will refer to the

2

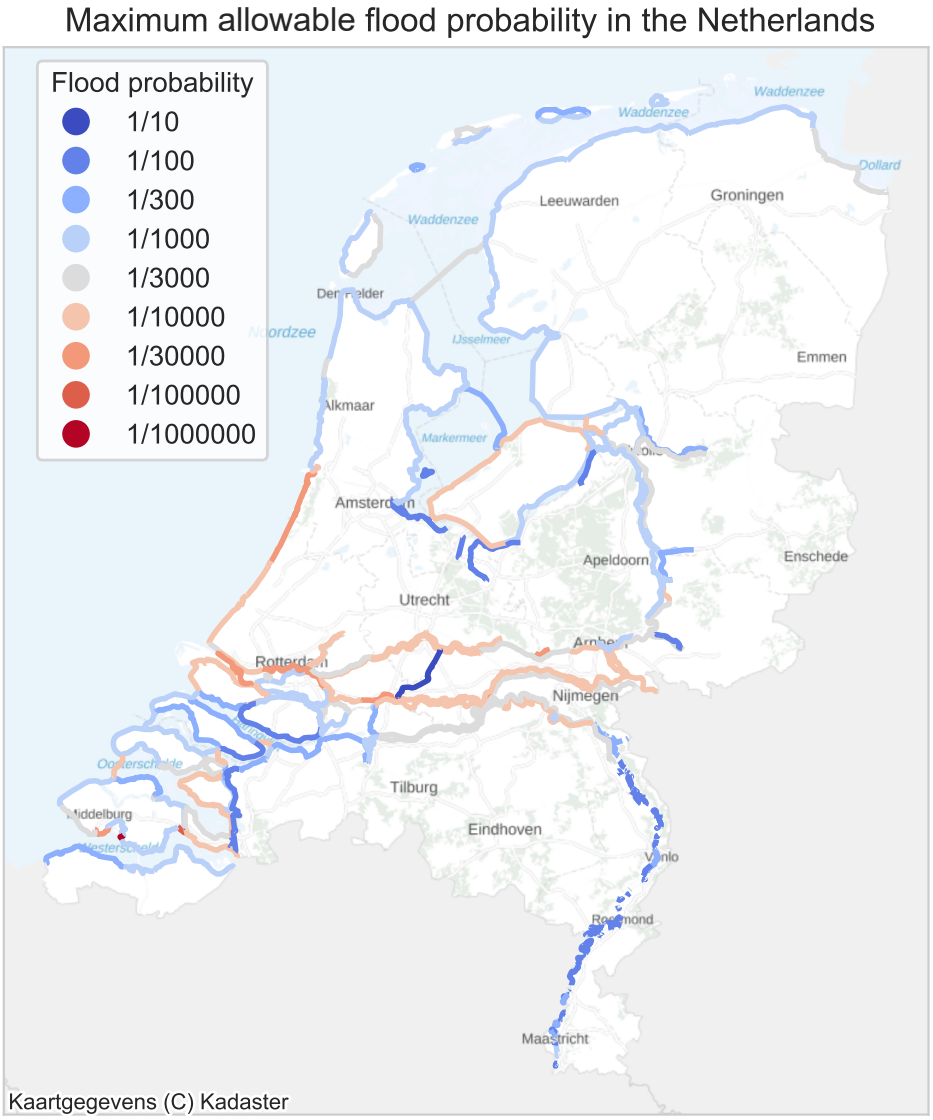


Figure 2.1.: *Maximum allowable probabilities of flooding per year for flood defence segments in the Netherlands as described in the Water Act.*

The change in requirements has been accompanied by updated manuals that enable assessment of failure probability requirements using (semi-)probabilistic assessment methods [6]. There are two important remarks to this approach which will be considered explicitly in this thesis:

1. The definition of reliability (requirements), and especially how these are formulated and transformed in space and time builds upon a number of assumptions with regards to the influence of spatial and temporal correlation on reliability, as well as the general influence of degradation on reliability. This is further discussed in Sections 2.4 and 2.5.

2. An important aspect of estimating failure probabilities is that all uncertainties should be quantified, and taken into account [7]. This means that reducing uncertainty —such as through monitoring or site investigation— has a direct and explicit influence on failure probability estimates [e.g. 8]. However, this also means that overlooking or ignoring uncertain factors —parameters or processes— leads to an inaccurate failure probability estimate, which means that decisions might be wrong or at least suboptimal. This is further discussed in Section 2.3.

In evaluating structural reliability, typically a structure is assumed to be in a certain state. In this thesis we will also consider the influence of structural degradation or damage. Such damage is typically not considered in the state considered in reliability analysis of flood defence structures. There are a few concepts specifically suitable for assessing the structural performance in case of such damage. Lind [38] formulated vulnerability V as the ratio of failure probability with and without damage to the structure, in an attempt to obtain an indicator that represents the possibility that small damage leads to disproportionate consequences (e.g., progressive collapse of a structure). Vulnerability is formulated as follows:

$$V = \frac{P(F|D)}{P(F|\overline{D})}, \tag{2.1}$$

where $P(F|D)$ and $P(F|\overline{D})$ are the probabilities of failure with, respectively without damage. Baker et al. [39] developed an approach for assessing structural robustness. Structural robustness is defined as 'the ability of a structure to withstand events like fire, explosions, impact or the consequences of human error, without being damaged to an extent disproportionate to the original cause' [40]. Baker et al. [39] formulates the robustness index I_R from a decision theory perspective as:

$$I_R = \frac{R_{\mathrm{D_i}}}{R_{\mathrm{D_i}} + R_{\mathrm{I_i}}}, \tag{2.2}$$

where $R_{\mathrm{D_i}}$ are the direct consequences following situation (e.g., with initial damage i), and $R_{\mathrm{I_i}}$ are the indirect consequences. As such, for a small damage with low

independent elements as sections, which are parts of the flood defence of ≈1 kilometre with relatively homogeneous strength properties, that can be represented by a cross-section. See Figure 2.3 for a schematization.

direct consequences and large indirect consequences $I_R \approx 0$, the structure is not robust to damage. The principles of vulnerability and robustness are related to other structural concepts: e.g., structural systems that are more ductile and have higher redundancy are also more robust (less vulnerable) [39]. Structural systems that can recover from disturbances are more resilient, and as these systems have less indirect consequences are thus typically also more robust. The Eurocode advises risk-based robustness design for structures with high consequences of failure [40].

It should be noted that there are many different robustness concepts used in literature, and that structural robustness is distinctly different from decision robustness. For instance, Maier et al. [41] defines robustness as the insensitivity of a decision to changes in future conditions, which is a clear definition of decision robustness. Baker et al. [39] and European Committee for Standardization [40] consider robustness in the context of structural failure: Structural robustness is defined as that 'the consequences of structural failure should not be disproportional to the effect causing the failure'. This is similar to Mens et al. [42], where system robustness is defined as 'the ability of a system to remain functioning under disturbances'. In this thesis we will use the general definition by European Committee for Standardization [40], as further specified by Baker et al. [39].

2.3. Uncertainty in flood defence reliability estimates

2.3.1. Estimating flood defence reliability

For flood defence structures the failure probability is typically defined as the probability that a structural failure results in a flooding of the protected hinterland [35]. The failure probability of a flood defence with uncertain parameters $\mathbf{X}$ can be described as follows:

$$P_f(\mathbf{X}) = P(g(\mathbf{X} < 0)) = \Phi(-\beta(\mathbf{X})) \tag{2.3}$$

where $P_f(\mathbf{X})$ is the failure probability given uncertain input parameters $\mathbf{X}$, $g(\mathbf{X})$ is the limit state function that relates input parameters to failure, and β is the reliability index of the flood defence. It should be noted that in specific cases the limit state function is often represented as $Z = R - S$, where R represents the strength, and S the load on the structure, and failure occurs of $Z < 0$. Note that for a flood defence typically multiple limit state functions have to be evaluated for different (combined) failure modes. Figure 2.2 shows the failure modes considered in this thesis, namely overtopping erosion, piping erosion, inner slope instability and erosion of the outer slope. For more information on this and other failure modes the reader is referred to the International Levee Handbook[3] and the concerning chapters.

2.3.2. Classification of uncertainty

Many of the parameters $x \in \mathbf{X}$ are uncertain, but the class and type of this uncertainty can differ per parameter. This classification is of importance, as it determines

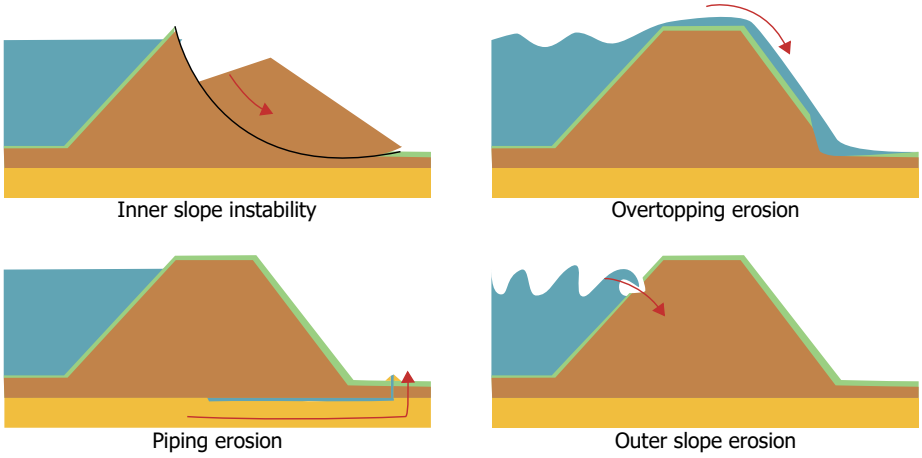


Figure 2.2.: *The 4 failure modes considered in this thesis. More information can be found in concerning chapters and the International Levee Handbook [3].*

the action perspective of an asset manager in coping with these uncertainties. Various authors have developed structures and systems for classifying uncertainty. A common distinction between different types of uncertainty is that of epistemic and aleatory — or knowledge and inherent — uncertainty. For instance Van Gelder [43] distinguishes between these two main categories, where epistemic uncertainty is split between statistical and model uncertainties, and inherent uncertainties are those caused by randomness. Examples of the first are uncertainty due to lack of sufficient data, or lack of understanding of physical behaviour, whereas examples of inherent uncertainties are for instance the extreme loads to be withstood in the next year. Typically the assumption is that epistemic uncertainties are reducible, whereas aleatory uncertainty is not. However, as is stated by Kiureghian & Ditlevsen [44] the classification of uncertainty is only relevant and consistent within the confines of a defined model of analysis. From a decision perspective this could be extended even further to the statement that the scope of the decision analysis determines the reducibility of uncertainty. Drawn to the (theoretical) extreme, one could say that all uncertainties are reducible in the context of a decision problem with infinite time and money (which is an impracticable definition for a decision problem). In practice, there will likely be some money and some time, which will determine the methods that can be used for reducing uncertainty, and thus which uncertainties are reducible. In determining which methods are most suitable, it should be noted that in the application of such methods there are also many types of (at least partially) reducible uncertainties — Phoon et al. [e.g. 45] distinguish between amongst others measurement, transformation and model uncertainties. Each of these can be reduced, but each might need a different approach. Examples of this will be elaborated in a.o. Chapters 4 and 5.

2.3.3. Levels of uncertainty in decisions

One of the core ideas of risk-based decision making is that all (known) uncertainties are quantified and taken into account. In most risk-based decisions this is not the case, for which Walker et al. [46] identifies 4 levels of uncertainty: statistical uncertainty, scenario uncertainty, recognized ignorance and total ignorance. Especially the latter two categories are often not taken into account in decision analysis. For 'total ignorance' this makes sense, as the decision maker is unaware of it.[3]

For uncertain variables in the category of recognized ignorance this is more tricky for two reasons. First of all, Raiffa & Schlaifer [47] already stated that every decision analysis starts with setting a scope in which some (in)formal choices are made to make the analysis feasible at all. However, through such choices that limit the scope of the decision analysis, the solution space is limited (an example will be discussed in Chapter 3). Hence, also informal choices in the definition of the decision problem have to be properly substantiated [47], especially if the uncertainties involved in that choice are difficult to quantify (i.e., they fall in the category of 'recognized ignorance'). Secondly, a decision maker might not include uncertainty as he is unable to properly grasp or quantify it. This can be reasonable if it is likely that it is also of limited influence, but if this is not the case it is doubtful whether the analysis will lead to a sound decision. This can be a major issue if there is insufficient knowledge about a potentially influential uncertain variable. In such cases using additional performance indicators, such as structural robustness or vulnerability (see Section 2.3) can be of value, as will be demonstrated in Chapter 7.

2.4. Reliability of flood defence sections and segments

Figure 2.3 displays a typical flood defence segment. In the Netherlands, the reliability requirement for such a system is defined at a segment level, which is typically a stretch of flood defence of 5 to 25 kilometres in length. For practical reasons, such a segment is usually subdivided in sections with relatively homogeneous geometry and strength parameters. Such sections, typically with a length of 500-1500 meters, can then be represented by a single representative cross section.

Reliability estimates as described in Section 2.3 need to be evaluated based on the reliability requirement of the flood defence segment. In order to compute the system reliability, one therefore has to combine the evaluations of all limit state functions for all failure modes for all (cross) sections. If such a cross-sectional evaluation of the failure probability is done, one has to properly account for the spatial transformation of the different input parameters, such that the cross section is indeed representative for the section it is meant to represent. This includes defining the input parameters at the proper scale, and accounting for aspects such as spatial averaging in case of for instance slope stability, where the scale of a

[3] An illustrative example: in 2019, a year before the COVID-19 pandemic, a decision analysis on investing in company assets likely did not include a substantial amount of assets related to working at a distance compared to the same analysis at the time of writing this thesis.

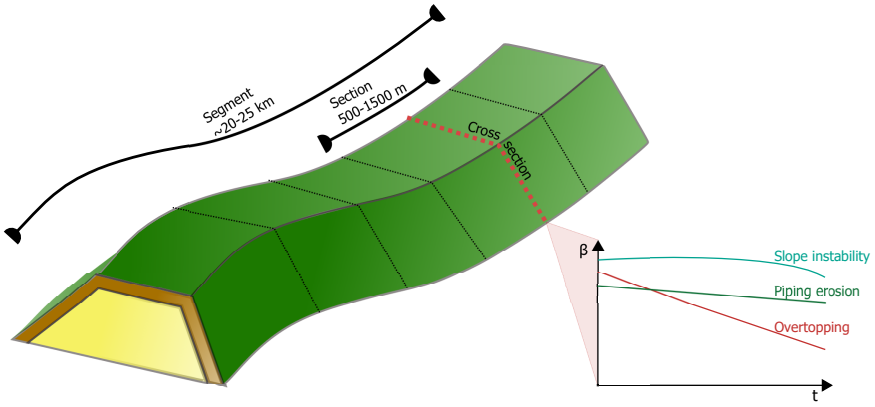


Figure 2.3.: *Schematization of a flood defence segment. Segment, section and cross section are displayed as they are used in this thesis. Cross-sectional life-cycle reliability estimates for different mechanisms are also illustrated.*

failure is much larger than that of a cross section. This is further discussed in van der Krogt et al. [48], and details are beyond the scope of this thesis.

A flood defence segment can be schematized as a series system, where the system reliability is determined by the correlation of uncertain random variables of different elements and the failure probability of those elements. The correlation between parameters in the spatial dimension is often modelled using (squared exponential) autocorrelation functions [6, 49]. Typically the correlation length of subsoil parameters is rather short (order of a section), and the correlation length of load variables is relatively long, depending on the type of system. As such, there is typically partial correlation between flood defence sections. One of the methods to obtain this system reliability is by combining the limit state functions based on the correlation of different (shared) random variables between failure modes and flood defence sections using for instance the Equivalent Planes method [50].

The system reliability can be quite well approximated by two bounds— the system failure probability of a series system with N fully independent elements is approximated by:

$$P_{\mathrm{f,sys}}(t) \approx 1 - \prod_{i}^{N} 1 - P_{\mathrm{f,i}}, \tag{2.4}$$

and of N fully dependent elements is approximated by:

$$P_{\mathrm{f,sys}} \approx \max_{i \in N} P_{\mathrm{f,i}}. \tag{2.5}$$

It was demonstrated by Vanmarcke [51] that even for systems with equally reliable components and a relatively strong correlation ($\rho > 0.8$) the reliability can be approximated quite accurately by Equation (2.5), and similarly, for systems with $\rho < 0.8$ by Equation (2.4). Therefore, while not exact computations, such formula-

tions still provide rather accurate results for two reasons:[4]

- In many cases a limited number of weak spots dominates the reliability of a flood defence segment, regardless of the exact correlation between segments. In such cases the exact correlation between elements has limited influence on the segment reliability.
- For systems with equally reliability components, most failure modes behave along the lines of 1 of the two bounds given in Equation (2.4) and Equation (2.5). Concretely, for geotechnical failure modes such as piping erosion and slope instability the elements only have relatively limited partial correlation ($\rho < 0.8$) due to the influence of strength uncertainty. In such cases Equation (2.4) provides an accurate (slightly conservative) estimate of system reliability. If correlation is higher, such as for load-dominated mechanisms like overflow the bound in Equation (2.5) provides an accurate estimate.

To illustrate this we consider a simple example of a series system with 10 elements. We evaluate the system reliability by randomly sampling outcomes from a correlated (u_1) and uncorrelated (u_2) standard normal random variable. As such, we can compute the limit state function for each element as follows:

$$Z = \beta_{\mathrm{i}} - \rho \cdot u_1 + \sqrt{1-\rho^2} \cdot u_2, \tag{2.6}$$

where ρ is the correlation coefficient, and β_{i} the reliability of the element.

The left pane in Figure 2.4 shows the system reliability β_{sys} for a system of 10 elements of equal reliability ($\beta_{\mathrm{i}} = 3.5$) for different values of the correlation ρ. The right pane shows the same system, but now one of the elements has a reliability of $\beta_{\mathrm{i}} = 2.5$. It can be clearly observed that the weak spot dominates the system reliability.

In many cases it is desired to also obtain target failure probabilities for a (cross) section. To that end, a failure probability budget has been defined [6]. This approach distributes the allowed failure probability among failure modes and sections, based upon general assumptions of the correlations and typically dominant failure modes. The main principle is that the total failure probability for the segment is translated to requirements for cross sections that each represent an independent dike section. If the requirements are met for each failure mode at each cross section, the overall target is also met. The requirement for an independent section for failure mode m ($P_{\mathrm{T,m,cs}}$) is given by:

$$P_{\mathrm{T,m,cs}} = \frac{\omega_{\mathrm{m}} \cdot a_{\mathrm{m}} \cdot P_{\mathrm{T,seg}} \cdot L}{b_{\mathrm{m}}}, \tag{2.7}$$

where $P_{\mathrm{T,seg}}$ is the maximum allowable failure probability for the segment with length L in meters, ω_{m} is the fraction of the total failure probability that is allocated for failure mode m, a_{m} and b_{m} are length effect factors. a_{m} represents the fraction

[4] It should be noted that, while being more accurate, also more advanced methods such as the Equivalent Planes method only provide approximations of system reliability.

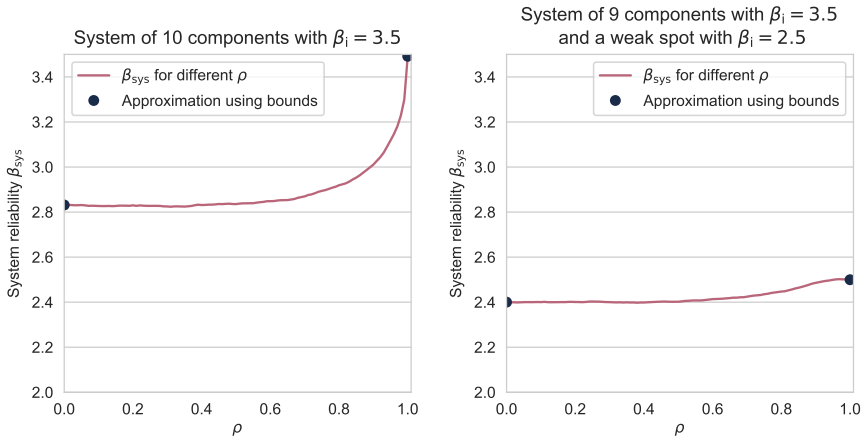


Figure 2.4.: *System reliability for a system of 10 elements. Left pane: all elements have* $\beta_i = 3.5$*. Right pane: one element has* $\beta_i = 2.5$*.*

Table 2.1.: *Values used for determining the cross-sectional target reliability* $P_{T,m,cs}$ *[52].*

Failure mode	a_m	b_m	N_m	ω_m
Overflow	1	-	1, 2 or 3	0.24
Inner slope stability	0.033	50	-	0.04
Piping erosion	0.9	300	-	0.24

of the dike segment that is sensitive to failure mode m, and b_m represents the length of an independent equivalent section in meters. For overflow L/b_m is substituted by a length-independent factor N_m that depends on the segment considered. For more details and default values see Table 2.1 and Jongejan et al. [6].

2.5. Temporal aspects of reliability

The reliability of (deteriorating) structures in time can be described as a first-passage problem governed by a mixture of stochastic processes and random variables [53]. A first-passage problem means,that reliability is determined by the time t at which the resistance of a structure R is lower than the load S for the first time. In the context of time-variant reliability, the key question thus is what the probability is that this happens at some time t, or at the interval $[0...t]$. This probability is not only influenced by the probability at $t = 0$, but also by the correlation between time-invariant random variables and the distributions of time-variant random processes. It should be noted that such time-variant reliability problems can be difficult to solve computationally [53].

There is a variety of patterns that describe the failure of structures over time.

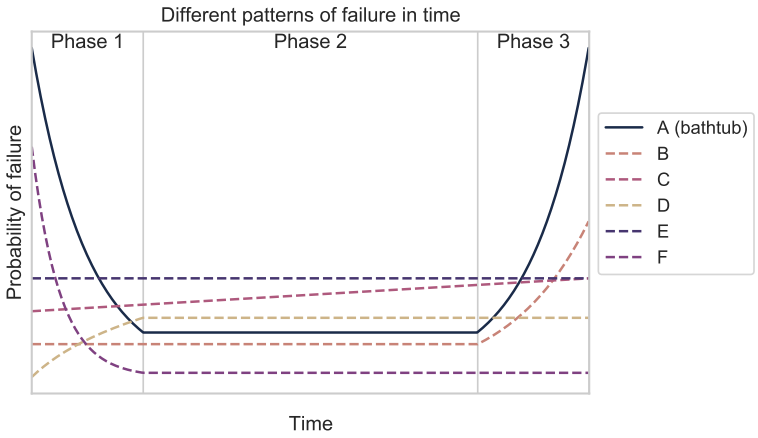


Figure 2.5.: *Six patterns of failure observed in aircraft maintenance: the bathtub curve (solid) and 5 different patterns encountered in practice (dashed). Reproduced and adapted from [54]. Values on both axes are arbitrary.*

Moubray [54] describes the development of views on how assets fail, mainly in the context of aircraft maintenance. In the last part of the 20th century, there was a big leap in quantitative analysis of maintenance, as some of the assumptions on failure patterns from the past did not align with what was encountered in practice. A common assumption before that period was that all asset failures were distributed like a bathtub curve (curve A in Figure 2.5). However, it was found in practice that nearly 70% of aircraft parts behaved like pattern F, and that all 6 patterns in Figure 2.5 were encountered in practice. Obviously this misunderstanding caused many ineffective replacements and additional failures, as a different pattern requires an entirely different maintenance and replacement policy. Looking at the bathtub curve (pattern A in the figure) there are three phases distinguished:

- Phase 1 is the so-called burn in phase, where infant defects are detected as assets fail early after installation or construction. Such infant defects become less likely over time. Consequently, (part of) the time-invariant uncertainty in their performance is reduced over time.
- Phase 2 is a phase with a more or less constant failure rate. Infant failures have occurred, and the failure rate is typically constant, where failures are for instance characterized as a Poisson process with a constant rate [55].
- Phase 3 is the phase where the influence of physical degradation starts to increase.

In practice, the three phases can not be distinguished for all assets. In some cases there might be hardly any infant failures in phase 1, whereas some assets hardly suffer from the degradation in phase 3, and in some cases the underlying

mechanisms of the phases might overlap. In Section 2.5.1 we will go into the principles underlying phase 1, Section 2.5.2 will go into the principles of phase 3.[5]

2.5.1. Temporal dependence in reliability

The behaviour observed in the first phase of the bathtub curve depends on the degree of temporal correlation in reliability estimates. This is something not always taken into account in structural and flood defence design. As this thesis will contain many analysis that consider flood risk cost over a longer time period, in this section we will illustrate in what situations including temporal dependence in reliability analysis is necessary.

Requirements for flood defences are typically formulated and assessed for a reference period of a year. In the Eurocode standards, reliability requirements are provided both for a year and 50 years [56, Table C2]. The underlying assumption of this table is that the years are independent, such that the reliability of a structure for a reference period t_{ref}, $\beta_{t_{\text{ref}}}$ is defined as [57]:

$$\beta_{t_{\text{ref}}} = \Phi^{-1}([\Phi(\beta_1)]^{t_{\text{ref}}}) \tag{2.8}$$

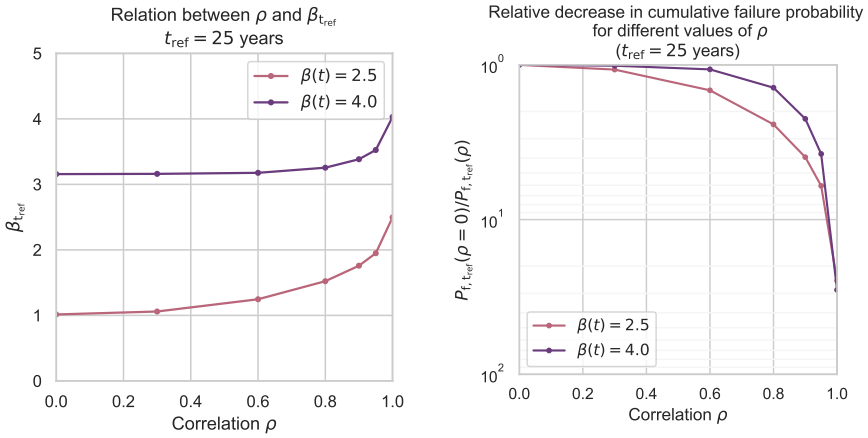
Where β_1 is the annual reliability requirement and $\Phi(\cdot)$ is the inverse cumulative distribution function of the standardised normal distribution.

However, there can be a dependence if uncertainty in time-invariant parameters contributes to the overall failure probability, as was argued by amongst others Vrouwenvelder [58] and Roubos et al. [59], and which is also shown in the different patterns of failure in phase 1 of Figure 2.5. One of the classes of structures where this is of particular interest are geotechnical structures, where it is not uncommon that 80-90% of the uncertainty in the design point estimate in a probabilistic computation originates from time-invariant (soil) parameters. The failure rate over time can be expressed both conditional and unconditional upon preceding years [60]. If the failure rate at time t is expressed as $P(Z(t) < 0)$, the conditional failure rate is given by:

$$P(Z(t) < 0 | Z(1..t-1) > 0) = P\left(\frac{Z(t) < 0 \cap Z(1..t-1) > 0}{Z(1..t-1) > 0}\right), \tag{2.9}$$

which essentially means that the failure rate at t is determined by those cases where the structure fails at t, and has not failed before. As such, the knowledge of a structure not failing in certain years can be used to update the parameters governing Z [53]. As was demonstrated by Roubos et al. [59], for a case where there is full dependence between years, 1 year of survival means that a structure will survive forever (provided that all parameters are time-invariant). In such a case the failure pattern in Figure 2.5 would be a sharply decreasing probability of failure in phase 1, and no failures in phase 2 and 3 (provided there is no degradation). This is a bit extreme, but in the context of this thesis it is valuable to look into the potential influence of such temporal correlation. As such we can assess the

[5] We will not go further into phase 2 as there is no time-dependent influence on reliability



(a) *Reliability index β for different values of temporal correlation ρ* (b) *Change in cumulative failure probability compared to the case with $\rho = 0$*

Figure 2.6.: *Relation between reliability β for a reference period t_{ref} = 25 years. Left pane shows relation between β and ρ for cases with time-invariant $\beta(t)$=2.5 and 4.0, the right pane shows the relative change in failure probability for the reference period.*

potential influence of this conditionality on reliability estimates for flood defences, which are often dominated by time-invariant uncertainties.

We consider two simple examples of (partially) correlated Z-functions where the reliability index is 2.5 and 4.0 (per year). It should be noted that we consider the case to be fully time-invariant, which —due to degradation— is not the case in practice (e.g. in Chapter 3). For both cases we use a Monte Carlo simulation to compute the reliability index for a reference period $t_{\text{ref}} = 25$ years, for different values of the correlation between years ρ. Figure 2.6a shows that for a relatively low reliability index (2.5), the difference in $\beta_{t_{\text{ref}}}$ is larger than for the case with a higher reliability index (4.0) where the difference is marginal. Additionally, the relative difference (see Figure 2.6b) only becomes significant for the case with $\beta = 4.0$ if the correlation between subsequent years $\rho \gtrapprox 0.9$, which is a rather extreme case. Henceforth, we can conclude that for cases where the annual reliability index is low, and where there is extremely strong correlation among years, temporal correlation influences reliability estimates. The cases that we consider in this thesis do not fall into this category.

It should be noted that a practical situation that might arise in practice is one where an actual survival of an extreme event was observed. In such a case one can use this to obtain updated posterior current and future reliability estimates [e.g. 20, 21]. Worthwhile to note is that the basic principles are the same of that outlined in Equation (2.9).

2.5.2. Influence of degradation on reliability in time

Next to the temporal correlation in reliability, flood defences also change physically in time due to a variety of degradation processes (phase 3 of the bathtub curve). In this section we will outline the main distinctions in modelling such degradation. Examples of degradation processes are the degradation of asphalt revetments, settlement of the crest and various kinds of damage that can impact the performance of revetments [3], such as animal burrowing, rutting or weeds. It should be noted that in scientific literature terms such as degradation, deterioration and damage are used as substitutable terms. In this thesis we will refer to the decrease in performance due to time-variant processes as degradation.

The processes underlying different types of degradation can differ. Sanchez-Silva et al. [13] state that structural degradation is caused by a combination of progressive and sudden events. From this, we can distinguish two main types of degradation processes:

- Continuous progressive degradation: this covers those processes that are typically relatively slow and can in many cases be described by a predictive model. Examples of continuous degradation are corrosion, settlement and degradation of asphalt revetments [61].

- Shock-based degradation: shock-based processes are processes where individual shocks cause a sudden decrease in performance. Shocks can be distributed randomly in time, and can differ in size. In some cases a single shock might result in a large decrease in performance, in some cases many successive shocks can cause failure.[6] Examples of a failure through a single shock are extensive damage due to an earthquake, or in the case of flood defences large scale burrowing damage. Smaller shocks can for instance be degradation due to fatigue or in the case of flood defences wash-out of revetment joint fill due to recurring wave loads. It should be noted that if shocks have very limited impact on structural performance, the practical implications of degradation are similar to progressive degradation — in such cases the degradation of a structure can be monitored by regular inspections.

In practice, such processes might also be related, leading to combined degradation [13]. However, while Sanchez-Silva et al. [13] only look at the influence of different degradation processes, Buijs et al. [62] elaborated this for flood defences and distinguished three types of features that lead to a certain degradation of flood defence properties:

- Excitation features — the properties or circumstances that initiate the time-variant process. Examples are third party interference or wave loads.

- Ancillary features — the structural properties that transform the excitation into damage. For instance the slope of a revetment.

[6] Note that in this context failure is not necessarily a flood, but rather that the state of a flood defence is such that the target reliability is not met.

- Affected features — the actual structural parameters that change due to the combination of excitation and ancillary features.

This gives a more physically sound basis for the consideration of shocks versus continuous degradation: in some cases there might be multiple excitation features that affect the same features — for instance missing blocks in a revetment might be caused both by vandalism and by extreme or recurring wave loads, and might be enhanced by the ancillary feature of a bad state of maintenance (e.g., lack of joint fill material).

The interplay between these different features makes modelling damage difficult: for instance Buijs et al. [62] models cumulative effects of piping erosion as a hierarchical gamma process where the effective seepage length decreases over time. However, in a follow-up of the test described in Pol et al. [63] it was found that pipes can partially recover which — in practice — might be further enhanced by traffic-induced vibrations or other ancillary features [64]. This only emphasizes the complexity and uncertainty in time-variant reliability estimates for flood defences. Fortunately there are many measures to deal with this uncertainty, the next section provides a taxonomy of such measures.

2.6. A taxonomy of interventions for managing life-cycle reliability of flood defences

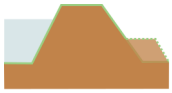
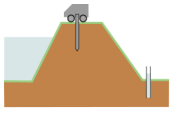
From the perspective of Bayesian decision theory, which we will introduce further in Section 2.7, there are two generic types of decisions [47]: actions to obtain information about the state of the world (experiments), and actions to improve that state (terminal acts). In the context of flood defences we can decide upon actions to improve the flood defence reliability, or reduce uncertainty in reliability estimates.

There is a variety of larger and smaller interventions that can improve flood defence reliability, such as relatively small repair works (e.g., repair of a grass revetment) or large-scale reinforcement measures (e.g., enlarging the berm or placing a diaphragm wall). From the perspective of uncertainty, we distinguish two types of methods for improving reliability: interventions where relevant uncertainties are mitigated, and those where other uncertainties become dominant due to a change in structural behaviour. An overview of these interventions/measures is given in Table 2.2.

Mitigation measures concern for instance enlargement of the berm, repair works to a revetment and other measures where the structural behaviour of the flood defence does not change, except for a general increase in reliability. This can for instance be determined by evaluating the relative uncertainty contribution of different parameters — for a mitigation measure the relative uncertainty contribution is similar after improving the mean structural capacity.

Substitution measures are different in the sense that the intended structural behaviour of the system is adapted, which not only influences the mean structural capacity, but also the relative contribution of different uncertain random variables.

Table 2.2.: *Three types of measures as considered in this thesis*

Type of intervention	**Definition**	**Examples in this thesis**	
Mitigation measure	Measure to improve reliability without any significant changes to the structural behaviour.	• Berm reinforcement • Crest heightening • Repair of revetment	
Substitution measure	Measure to improve reliability where the structural behaviour is altered such that other uncertainties become dominant.	• Diaphragm Wall • Vertical Geotextile • Stability screen	
Uncertainty reduction	Investment in reducing epistemic uncertainty through obtaining additional data or knowledge on the structural system.	• Pore pressure monitoring • Inspection • Proof load tests	

For instance, when a (self-retaining) diaphragm wall is installed, reliability for piping erosion and inner slope instability are no longer dominated by the site-specific ground properties, but by the structural capacity of the diaphragm wall. Hence, the uncertainties in ground properties are eliminated from the decision problem, and substituted by the uncertainty in structural capacity. Such measures can be specifically interesting if existing uncertainties in for instance subsoil permeability require extremely large piping berms. Eliminating this uncertainty by substituting it for e.g., the performance uncertainty of a Vertical Geotextile can be a very effective measure to improve reliability. [7]

Alternatively, if dominant uncertainties are reducible, reducing these measures through for instance pore pressure monitoring or site investigation is often a worthwhile effort. One specific characteristic is that the outcome of such an action is not known a priori, which is why, along with Raiffa & Schlaifer [47] and many others we will use pre-posterior analysis to evaluate the benefits of such measures (see Section 2.7 for details).

When determining a strategy for reducing uncertainty, it is important to distinguish between uncertainty in time-invariant parameters, and uncertainty in time-variant processes. This is illustrated in the top pane in Figure 2.7. Here we consider uncertainty in the shear strength of the subsoil and uncertainty in the pore pressure response. For reduction of time-invariant uncertainty it is important to distinguish between time-dependent efforts to reduce uncertainty (e.g., pore pressure monitoring [21]) and time-independent efforts (e.g., site investigation[8]). In the example, the information obtained from pore pressure monitoring depends on observed water levels, which is an ergodic signal.[8] Consequently, the information obtained from

[7] A Vertical Geotextile is a vertically inserted geotextile that blocks piping erosion. See for more information Chapter 3 and Koelewijn et al. [65].

[8] For ergodic variables, the average over time is the same as the average of the underlying probability distribution.

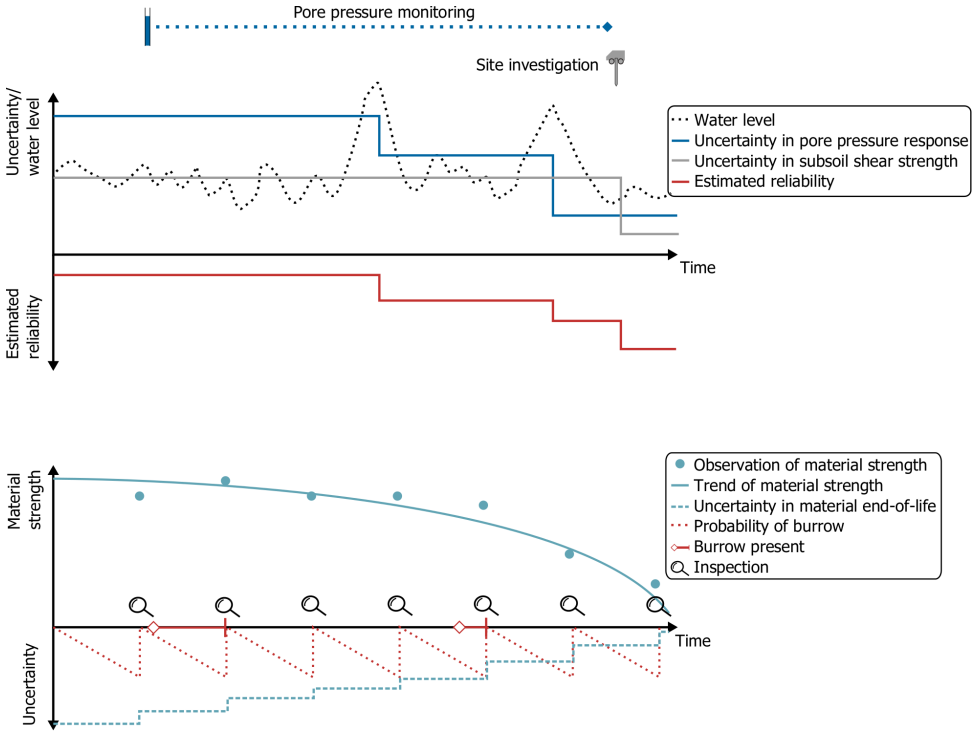


Figure 2.7.: *Illustrative figure for different types of uncertainty reduction. Top pane: reduction of time-invariant uncertainty through a time-dependent measure (pore pressure monitoring) and a time-independent measure (site investigation) and relation to reliability estimate. Bottom pane: reduction of uncertainty in time-variant animal burrowing and material degradation through periodic inspections (note that burrows are repaired upon inspection). Note that all values are relative and illustrative.*

pore pressure monitoring is time-dependent: a reduction of uncertainty (drop of the blue line) coincides with a measurement from an extreme water level (dotted line). As such, the acquired information depends on monitoring observations, and thus the duration of monitoring. Contrary to this, in case of the shear strength, one can immediately execute Cone Penetration Tests (CPTs) in order to reduce uncertainty. Such efforts are time-independent.

Reduction of uncertainty in time-variant processes is illustrated in the bottom pane of Figure 2.7. Here we consider uncertainty in continuous (non-ergodic, see Section 2.5.2) material degradation such as an asphalt revetment (blue lines). Upon each observation, the additional information can be used to update the end-of-life estimate of the material. For shock-based processes, such as the occurrence of animal burrows (red lines) behaviour is different. For such processes, random shocks occur in time, and the day after an inspection such a shock might occur. As such, the uncertainty on the presence of burrows increases in between inspections,

which makes the frequency of inspection an important parameter in decisions on uncertainty reduction for such processes.

Lastly, within the context of measures available to manage life-cycle reliability, one can also combine different measures. For instance, during construction one can combine observations of the load and pore water pressures to decrease uncertainty and thus use the Observational Method to prove the new embankment is sufficiently safe [66]. After construction, monitoring can be used to demonstrate that for instance the drainage capacity of a relief well is still sufficient [67]. The variety of different potential measures highlights the need for an objective and formal evaluation method for decisions in flood defence asset management. In the next section we introduce Bayesian decision analysis, which will be a core principle in the remainder of this thesis.

2.7. Bayesian decision analysis

In the preceding sections we have outlined some of the aspects in estimating the current performance of flood defence systems. However, in order to translate this to efficient and effective asset management decisions, we need to evaluate and compare different strategies in order to determine which strategy achieves the best balance between performance, cost and risk. The rational analysis of decisions was formalized in the utility theory developed by von Neumann & Morgenstern [68], and related actions to a set of potential (uncertain) outcomes. Subsequently, the action with the largest expected utility is assumed to be the optimal course of action.[9] This principle has been applied widely in engineering problems, for instance in cost-benefit optimizations of flood defence systems [e.g. 34, 71].

The utility theory by von Neumann & Morgenstern [68] was further extended by Raiffa & Schlaifer [47], who provided a set of formalized procedures to evaluate (sets of) decisions based on preferences and (subjective) judgments by decision makers. To properly execute such an analysis a decision maker needs to be capable of two things:

1. Assign a utility (or cost) to each decision and all of its possible outcomes,
2. Assign probabilities to all uncertain variables involved.

The main advantage of such an approach is that it makes the often intuitive and subjective assumptions that underlie decisions transparent.

Bayesian decision models are based upon an important distinction between the state of a system and the belief of a decision maker about that state [47]. By relating both the belief and the state to decisions, the sets of actions for acquiring information and actions based upon this information, can be translated into estimates for utility. We distinguish 3 types of decision analysis:

[9] In this context it should be noted that it was shown by Kahneman & Tversky [69] that the economic rationality implied by the expected utility theory of von Neumann & Morgenstern [68] does not align well with human behaviour and preferences (a.o., due to risk averseness). van Erp [70] demonstrated that by use of an alternative position measure, this can also be implemented in the mathematical foundations of Bayesian decision theory, see also chapters 13 and 14 in van Erp [70].

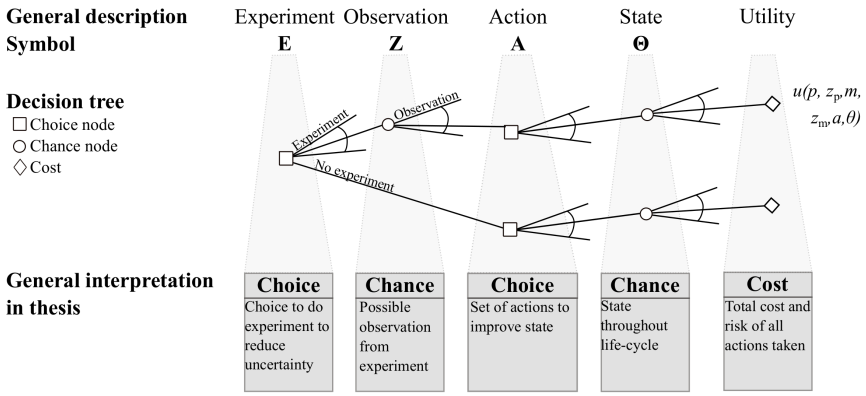


Figure 2.8.: *Decision tree for the choice whether to obtain information through experiment $e \in E$, modified from [47]. Levels E and A are choices by the decision maker, whereas levels Z and Θ are governed by chance. Level A consists of decision rules $d(z)$ which map actions to outcomes of z. The result is the cost over a combination of the levels $c(i, z, d(z), \theta)$.*

1. Prior decision analysis that focuses on the optimal action given the currently available information.

2. Posterior decision analysis (terminal analysis in Raiffa & Schlaifer [47]) that determines the optimal action given the available information supplemented with some additional obtained evidence.

3. Pre-posterior analysis to determine the optimal set of actions for obtaining new information, and other actions to change the system (e.g., a reinforcement measure).

The general goal of such models is to maximize the utility of decisions for a particular decision problem. It should be noted that, as Lindley [72] states "Today's posterior is tomorrow's prior" — at each point in time there is some evidence from the past that is used to underpin future decisions. In this thesis we are mostly interested in pre-posterior decision analysis, as such analysis enable quantifying the benefits of both single actions and sequences of actions (e.g., multiple inspections) [22]. A general formulation of such a model is given in Figure 2.8. In this decision tree we distinguish 5 levels:

- The action to acquire information through experiment $e \in E$, where E is the set of all possible information acquiring actions. Examples of such actions are inspections, monitoring and site investigation[e.g. 8, 73].

- The outcome of the experiment (e), $z \in Z$, where Z is the set of all possible outcomes. Note that z (if acquired) is used to update the belief of the decision maker about the state θ (see last bullet). Depending on the effectiveness of

experiment e, the belief z deviates more or less from the actual state θ. In case of perfect information (e.g., a perfect inspection) it holds that $z = \theta$. z can be both discrete and continuous.

- The action $a \in A$ following the obtained information, where A is the set of all possible actions. Here it should be noted that this can be formulated by a decision rule which maps different outcomes $a \in A$ to outcomes $z \in Z$. This yields a decision rule $d(z)$ that assigns an a to each z. Hence we use a set of decision rules $d(z) \in A$. An example would be a set of design rules that translate estimates of parameters z to (parametric) designs a, such that a given reliability target is met.
- The unknown state of nature $\theta \in \Theta$ where Θ is the a priori set of all possible states of nature. This unknown state determines the utility.[10]
- The utility given all preceding choices and chance outcomes $u(i, z, d(z), \theta)$.[11] For a flood defence utility u is governed by the expected cost of flooding given θ (probability times consequences), and the cost of all preceding actions (here e and $d(z)$).

Based on the input values for the different levels of the decision tree, one can compute the utility $u(e, z, d(z), \theta)$ at the last level for each combination of experiment and action (e and $d(z)$), and each realization for the state of nature θ. However, typically we do not know the outcome z of a yet to be obtained information, and here Bayesian pre-posterior decision analysis provides a structured framework to evaluate combinations of $e \in E$ and $d(z) \in D(z)$ and obtain the optimal combination with maximum utility based on the a priori information using [22]:

$$u(e, d(z)) = \max_{e \in E, d(z) \in A} E_\Theta[E_{z|\theta}[c(e, z, d(z), \theta)]]), \tag{2.10}$$

where A is the set of possible decision rules $d(z)$ and Θ is the prior distribution of state of nature. By comparing this utility to that without information e (bottom branch in Figure 2.8):

$$u(d(z)) = \max_{d(z) \in A} E_\Theta[u(d(z), \Theta)]), \tag{2.11}$$

one can obtain the Value of Information of the experiment e:

$$VoI = u(e, d(z)) - u(d(z)). \tag{2.12}$$

It should be noted that if decision trees are used in the context of an asset management strategy with many sequential decisions, they may quickly become

[10] Note that a slightly different approach is to weigh two branches: one branch with a failure and one without, where the probability of each branch is weighted by the probability of failure. This is however only a cosmetic difference which yields the same posterior utility.

[11] Note that we use utility here as it is more general, but in the context of flood defences cost is more adequate. The optimal combination of decisions maximizes utility or minimizes cost.

computationally intractable [e.g. 22, 26]. In such cases for instance use of influence diagrams with decision heuristics are a solution, as is the consideration of decisions using for instance Markov Decision Processes [e.g. 27].

As was stated in the beginning of this section, a proper pre-posterior analysis requires quantification of all relevant uncertain variables, and that the utility of each decision can be determined. Obviously this can be a daunting task, as is also stated in the introduction of the book by Raiffa & Schlaifer [47]. Nevertheless, it does provide a transparent approach, and while it might not lead to 'optimal' decisions, it can aid in finding 'satisficing' flood defence asset management strategies. The key point according to Raiffa & Schlaifer [47] is: "In most applied decision problems, both the preferences of a responsible decision maker and his judgments about the weights to be attached to the various possible states of nature are based on very substantial objective evidence; and quantification of his preferences and judgments enables him to arrive at a decision which is consistent with this objective evidence." In the next section we will consider some important aspects of defining these preferences and judgements.

2.8. Formal and informal choices in scoping flood defence asset management decision problems

Raiffa & Schlaifer [47] also highlight an important notion towards the practical applications of formal decision analysis: "in many situations informal analysis will quite properly reduce the field [of possible decisions] to a single contender and leave no scope for formal [decision] analysis at all". As such, based on the judgment of a decision maker, the formal decision problem is deemed irrelevant as there is only one course of action that aligns with the objective evidence. However, in practice decisions are not that clear, and for a flood defence the scope of a formal decision analysis on life-cycle reliability has three dimensions:

1. The temporal and spatial scope of the system under consideration.
2. A definition of the solution space (i.e., available actions and experiments)
3. A definition of what (types of) utility are considered for the evaluation.

In each of these dimensions, judgments of the (responsible!) decision maker determine what is to be incorporated in the decision analysis. The first dimension for instance entails that excluding part of a flood defence system in a decision analysis either means the objective judgment is that it will not influence the decision, or it is an uncertainty in the category of 'recognized ignorance' (i.e., we know it has effect, but do not objectively assess how much). The latter category should be avoided.

As for the second dimension, this often relates to the other dimensions, as is illustrated by the following example. We consider a case where both a diaphragm wall and berm widening lead to the same overall life-cycle performance. However, for constructing the berm, houses need to be removed which is not desirable. The

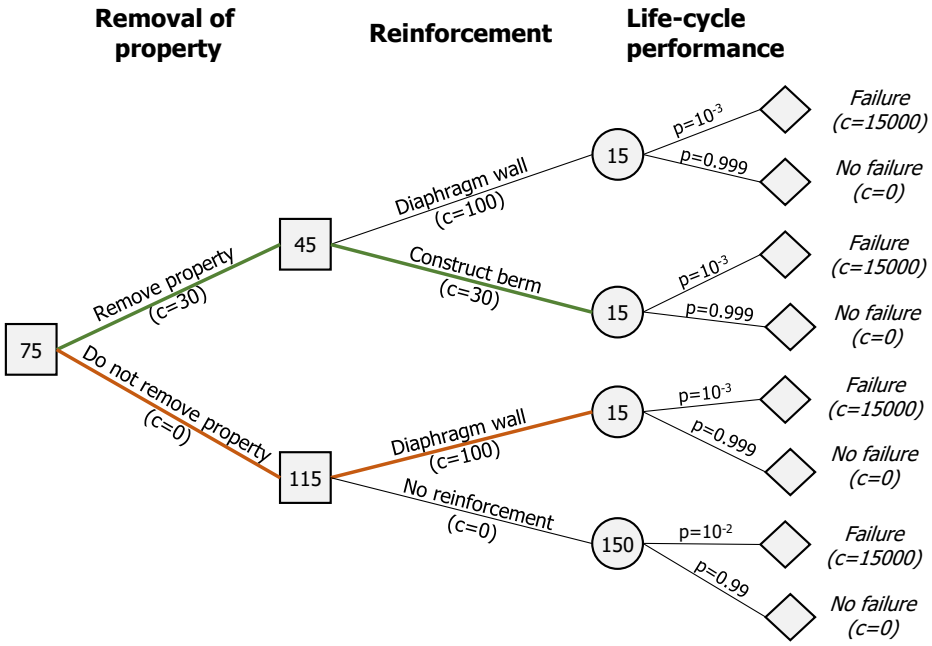


Figure 2.9.: *Decision tree for a reinforcement decision. The decision has 2 choice nodes (squares): a choice to remove adjacent property, and a choice of a reinforcement method. Values in the nodes indicate the total cost of the optimal decisions conditional on the choice in that node.*

decision tree including all assumed costs is shown in Figure 2.9. If all possible options are considered, including the cost for removal of the houses (compensation and removal), the economically optimal combination of choices would be to remove properties and construct a berm (green line). Now consider the case that we impose the constraint that no removal of properties is allowed — the optimal choice would now be to construct a diaphragm wall without removing property (orange line), which is more than 50 % more expensive than the unconstrained optimal choice. In other words: imposing the constraint of 'no property removal' — or framed differently, including the utility of removing properties — has a marginal cost of 40 (the difference between the branches). This demonstrates the importance of properly substantiating choices in utility measures and constraints to the solution space. It should be noted that incurring this marginal economic cost can be acceptable to the decision maker, but the consequences should be substantiated.[12]

Not in all cases the preferences and judgments of a decision maker can be fully expressed in economic utility, which has given rise to many different measures of performance and utility. In Section 2.2 we have already introduced structural robustness [39]. In literature also decision robustness has received ample attention, as a means to deal with large (future) uncertainty, for instance in estimates of future sea level rise [e.g. 41, 74, 75]. In such analysis a wide range of scenarios is considered. Such robustness indicators can be extremely useful in properly reflecting preferences and judgments, but should be used carefully. It was posed by Giuliani & Castelletti [76] that, as the different indicators for decision robustness have a large influence on the relative performance of decisions, the choice of measure should be an uncertain variable in the analysis itself. Next to that, as was noted by Maier et al. [41], even the most exhaustive exploratory scenario analysis is limited by human imagination and might cause overconfidence in a chosen strategy. Especially in view of the biases humans suffer from when estimating probabilities and possible scenarios, this should be kept in mind when evaluating additional performance measures, even if those are generally useful for formal decision analysis [77].

2.9. Trade-offs in solving decision problems

So far, we have mainly discussed decision problems along the lines of the decision tree proposed by Raiffa & Schlaifer [47]. However, in many cases, such decision trees might grow very large, due to a multitude of subsequent and dependent decisions throughout the life-cycle of a structure [22]. This hampers the evaluation of strategies for managing life-cycle reliability — especially when actions depend on previously obtained uncertain outcomes (e.g. inspections) one quickly encounters state space explosion. In such cases there are many techniques available that can aid in solving such problems, such as Dynamic Bayesian Networks [26], Markov Decision Processes [27, 78] and influence diagrams [79] combined with heuristics

[12] Note that here we translated the consequences to 'implied economic cost'. Obviously the consequences of removing properties could also be evaluated through separate criteria for intangible impacts of house removal, or by quantifying these intangible costs.

such as Direct Policy Search [80, 81].

The goal of this section is not to provide a lengthy generic description of all different methods available for solving such problems, each of the following chapters will discuss the applicable relevant literature by itself. However, there are a few key thoughts that will emerge multiple times in this thesis:

- Should we optimize or satisfice?
  In many cases [e.g. 26, 27, 81] it has been demonstrated that heuristic methods can find near-optimal strategies that are sufficient to base a decision on as will be demonstrated in a.o. Chapter 3. For a decision maker it typically suffices to obtain evidence that one decision is superior to other options considered. Consequently, a heuristic method that is more easy to solve and finds near-optimal strategies might be preferred for decision support.

- Abstract or realistic?
  While many studies in optimization of flood defence investments are available [e.g. 18, 71], there is often a focus on ensuring that optimal solutions are found (see also the remarks at the preceding point), at the expense of simplifying the real world problem. As such, the formal decision analysis might be accurate, but the judgements and preferences of the decision maker might not be fully reflected. Such simplifications hamper practical implementation, as there is typically no clear-cut approach for translating optimization results to practical solutions.

- What parts can be simplified?
  While realism in models is advisable, a realistic model is not necessarily extremely complex. Thorough analysis of influential uncertainties as well as sensitivity analysis such as global variance analysis [82] can aid in defining a model that is realistic but no more complex than necessary. As such, these analysis can aid in narrowing the scope of a formal decision analysis.

A final point of attention is the computation of economic cost which, in this thesis, is done by computations of life-cycle cost. However, it should be noted that in some cases using life-cycle cost might not yield all the desired information on what strategy is to be preferred. Woodward et al. [83] and van den Boomen et al. [84] propose to use real options analysis (combined with a decision tree approach) in some cases, especially when flexibility is of interest to the decision maker. In this thesis we focus on dealing with uncertainties in the performance of flood defence assets themselves, and not in uncertainties that are for example related to market variables as considered in [84]. Hence, use of life-cycle costing is sufficient for the scope of this thesis, but this might differ for decision problems with a different focus.

3

Optimal planning of dike reinforcement projects

The earth's fertility resembles a spring that is being pressed downwards... the effect of additional weights will gradually diminish.

Anne-Robert-Jacques Turgot

In Chapter 1 we have introduced three key topics that hamper efficient and effective asset management of flood defence systems. In the life-cycle of a flood defence system, reinforcement is the most impactful and costly intervention. Typically, every 20-50 years a major investment is required to counter effects of degradation, adapt to new insights, performance requirements, and meet societal demands. However, in practical situations, the translation of system reliability targets to interventions is suboptimal. For this reason, we introduce a new approach that optimizes this translation based on the available data on costs and benefits of reinforcement for the area considered. The approach uses a greedy optimization algorithm that is capable of finding near-optimal solutions for reinforcement of dike segments (length ≈ 20 kilometres) through allocation of investments in space and time. The case study presented has been carried out as part of the preparation of dike reinforcement project Streefkerk-Ameide-Fort Everdingen at one of the branches of the Rhine river.

The majority of this chapter has been published in Klerk, W. J. et al. Optimal planning of flood defence system reinforcements using a greedy search algorithm. *Reliability Engineering & System Safety* **207,** 107344. doi:10.1016/j.ress.2020.107344 (2021).

THROUGHOUT the life-cycle of flood defence systems, some of the largest investments are dike reinforcement projects, that typically recur every 20-50 year. These projects are required if dike segments do no longer meet the reliability requirements, for instance due to degradation (e.g., subsidence), increased loads due to climate change and changed societal requirements. Such projects are complicated: the impact on the surroundings is large, the costs are typically high, and the spatial scope is also large.

As was shown in Section 2.4, the reliability of a dike segment is determined by the reliability of the weakest sections. Additionally, the correlation between sections differs per failure mode — piping erosion and slope instability have very limited correlation due to the large influence of uncertainty in spatially heterogeneous ground parameters, while overtopping is strongly correlated due to the large influence of spatially homogeneous loads. Each of these failure modes has to be considered in a dike reinforcement, but typically the reliability of these failure modes varies per section.

The commonly used design approach to deal with this in flood defence reinforcement projects is a reliability-based design approach using cross-sectional reliability requirements for each failure mode [35]. These are determined based on the failure probability budget and (pre-defined) length effect factors discussed in Section 2.4. By ensuring that every section meets the reliability requirement for each failure mode, it is ensured that the dike segment as a whole also meets the target reliability. This is a typical form of 'analytical thinking' (see Section 2.1): it is presumed that considering parts of the system leads to an optimal (or good) solution for the whole.

However, such an approach towards translating system to element requirements has the disadvantage that it significantly restricts the solution space. A 'system thinking' approach would would not use cross-sectional reliability requirements but would rather look at the overall segment reliability and how this is impacted by interventions at different sections. A total cost optimization (cumulative flood risk and investment costs over time) could then allow for optimizing reliability along three dimensions:

- Optimal scheduling of interventions in time,
- Allocation of resources to those dike sections where the marginal utility is largest, [1]
- Allocation of resources to the failure modes which yield the highest marginal utility.

It should be noted, that solving such an optimization problem is rather complicated due to the large solution space. The main focus of this section is to investigate whether such an optimization problem can be solved, and to what extent this yields a more efficient reinforcement. We consider a variety of reinforcement measures

[1] Marginal utility is in this case the reduction of economic flood risk, which is the main driver for flood defence reinforcement projects.

for each section. Investments in uncertainty reduction in the context of dike reinforcement projects will be considered in Chapter 4.

In Sections 3.1 and 3.2 we describe the principles of (improving) dike segment reliability and how this can translate to heuristics for a greedy algorithm. In Section 3.3 we discuss previous research on optimization of system interventions, both for flood defences and other infrastructure. Section 3.4 provides a general description of the optimization problem for planning reinforcements of a dike segment. The greedy algorithm described in Section 3.5 is then used to analyse the design and planning of an actual riverine dike system consisting of 41 dike sections in Section 3.6. Sections 3.7 and 3.8 provide some additional discussion and conclusions on the approach, as well as how the findings relate to the general perspective of managing life-cycle reliability of flood defences.

3.1. Reliability estimates for dike segments

For riverine dike segments, typically the three most dominant failure mechanisms are piping erosion, inner slope instability and overtopping. In Section 2.4 it was already demonstrated that in most cases the spatial correlation between different sections for these mechanisms behaves in accordance with the bounds formulated by Vanmarcke [51] (Equation (2.5) and Equation (2.4)). These relations consider a single failure mode, if M mechanisms are considered the overall segment failure probability can be obtained from:

$$P_{\text{f,segment}}(t) \approx 1 - \prod_{m \in M} (1 - P_{\text{f,m,segment}}(t)). \tag{3.1}$$

Note that we assume that correlation among failure modes is not relevant, which is true if the dominant uncertain variables differ (no correlation), or if there is a large difference between the failure probability of different failure modes (correlation is not relevant). The latter is true for most sections in the considered case in Section 3.6.

Due to temporal changes in load and strength, reliability will change over time. Deterioration of the strength can consist of, amongst others, settlement of the dike crest resulting in additional overtopping, settlement of the inner toe resulting in larger hydraulic head for piping erosion and decreased slope stability. Temporal changes in loads are typically caused by climate change resulting in higher water levels and higher waves [86]. This has a direct and relatively large effect on overtopping reliability, but the influence on piping and slope instability is smaller as there is damping of these effects in the subsoil. Figure 3.1 shows the reliability in time for an example cross section (solid lines) as part of a larger dike segment. Here it can be seen that overtopping reliability drops faster than that of geotechnical parameters due to the larger influence of the load on overtopping reliability. Further details on how we derive the reliability in time for the different failure modes are given in Appendix B.

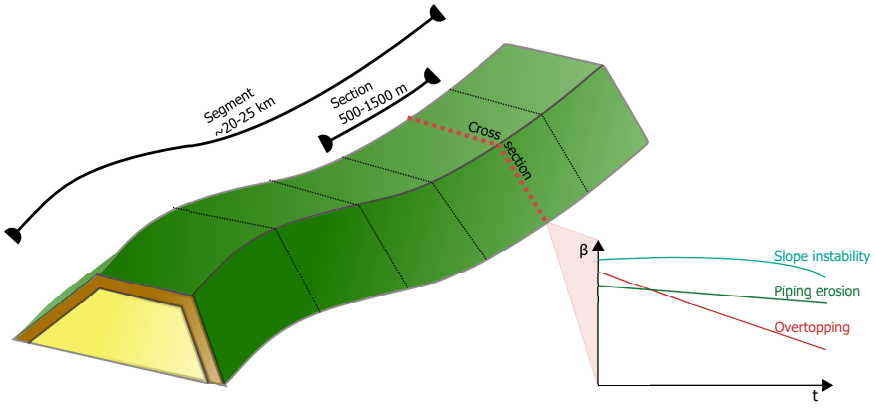


Figure 3.1.: *Relation between dike segment, dike section and cross section and schematic representation of the reliability in time for a single cross section for different failure modes as considered in this chapter.*

3.2. Improving reliability of dike segments

Dike segments in the Netherlands have to meet reliability requirements at a segment level. However, within the context at the time of writing, there are a number of considerations that dike asset owners have to deal with. Aside from the fact that dike segments consist of many non-homogeneous sections, dike asset owners in the Netherlands currently also have to deal with at least the following three considerations when planning dike reinforcement projects:

- The dike segment has to meet a reliability requirement, but they typically have time to achieve this.
- The strategy to achieve this has to be explainable to other stakeholders.
- They can choose between many different types of reinforcement methods, with different costs at each dike section.

Note that in practice also considerations of other functions and regional developments are of importance, but in this chapter we focus on achieving the required reliability in a cost-optimal manner.

Dikes provide utility in terms of reduction of economic damage and loss-of-life due to flooding [36]. Based on analysis of costs and utility, optimal target reliability requirements have been obtained that ensure optimal risk levels [e.g. 35] (see also Section 2.2). Ideally, any investment in dikes would be evaluated to yield minimal total cost (flood risk + investment), whilst meeting other criteria for individual and societal risk as well. The practical approach in the Netherlands is to first define optimal reliability requirements after which investments are aimed at meeting these. As a major part of the optimization is based on optimal total cost, these reliability requirements are typically very close to what would follow from a total cost optimization.

In the Netherlands, the overall goal is to meet the reliability requirements in 2050 for the entire country. The main reason is that it is not achievable to reinforce all dike segments to the required level before that time with the available budget and capacity. Practically this means that dike managers have to ensure two things:

- That they meet the target reliability in 2050.
- That they achieve this target in an efficient way within existing (budget) constraints.

Thus, we can distinguish two phases in the optimization problem: the period up to the year when the requirement has to be met (2050), and the period after that. In the first period it is of importance that investments are efficient in terms of total cost (investment and risk costs). It has been found that as for example loss-of-life risk is typically strongly correlated to economic damage, optimizing the total economic cost is a good approximation for other risk indicators as well [87]. Thus, we can use a total cost optimization to determine an optimal planning of reinforcement measures for the first phase. In the second phase, after 2050, a total cost optimization is still a good approach, but additionally it should be required that the reliability remains above the target level.

Secondly, aside from societally cost optimal, a planning of dike reinforcements needs to be explainable and transparent, as many stakeholders are involved. In the context of a dike reinforcement there is for instance involvement by financing organizations, inhabitants, nature preservation organizations, local farmers and local and regional governments. This means that in practice a decision on a dike reinforcement is risk-informed, rather than risk-based, and as was argued by Bohnenblust & Slovic [88] a technical analysis should aim to focus discussion between stakeholders on key issues rather than providing a clear-cut solution. For our analysis this means that an approach that not only gives an optimal planning but is also explainable to non-technical stakeholders is to be preferred.

Lastly, there are many techniques available for dike reinforcement, although typically dikes are reinforced by heightening the crest (to prevent overtopping) and widening berms (to counter instability and piping erosion issues) with additional soil material. However, also structural measures such as diaphragm walls and sheetpiles [89], as well as innovative measures such as Vertical Sandtight Geotextile (VSG) [65] are applied. Within the context of the distinction proposed in Section 2.6, reinforcements with soil can typically be considered as uncertainty mitigation measures, structural measures such as diaphragm walls and e.g., Vertical Sandtight Geotextile are uncertainty substitution measures. For instance, while a VSG largely eliminates the threat of piping and thus the influence of ground uncertainties, these uncertainties are substituted by uncertainties on the performance of the VSG that is installed. Such structural measures are specifically interesting for countering threats from instability and piping erosion in densely populated areas where large ground uncertainties result in low reliability estimates for slope instability and piping erosion. By substituting these large uncertainties such measures can ensure sufficient reliability using a limited spatial footprint. It should be noted that

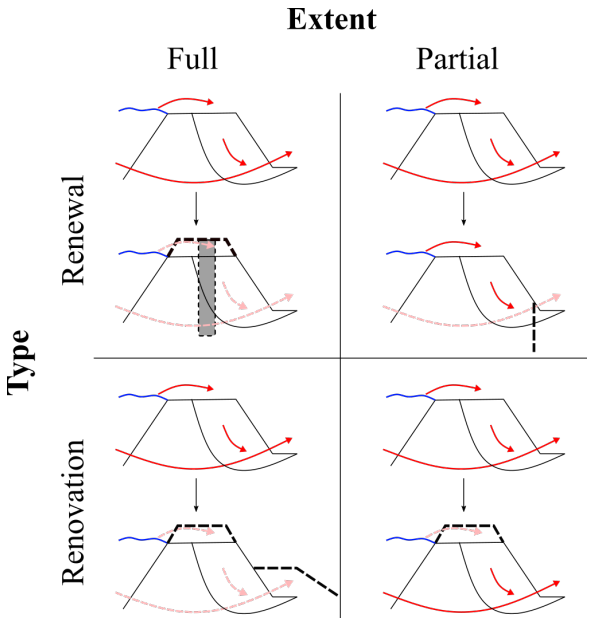


Figure 3.2.: *Categorization of types of reinforcement measures in this chapter. Red arrows indicate the relevant failure modes (inner slope instability, piping erosion and overtopping), pink dotted arrows indicate if a measure has resolved this deficit. For each type an illustrative reinforcement measure is shown (dashed black lines). From top left, clockwise: a full renewal using a diaphragm wall, a partial renewal using a Vertical Sandtight Geotextile, a partial renovation through crest heightening and a full renovation with combined berm widening and crest heightening.*

some of these structural measures are much more expensive than enlarging the dike profile with additional soil.

Within the context of optimally allocating resources over different failure modes, we classify the considered measures in two main dimensions, extent and type, as is illustrated in Figure 3.2. For the extent we distinguish between full measures that impact all relevant failure modes, and partial measures aimed at improving reliability for only 1 failure mode. A renewal type measure adapts the structural behaviour of the dike, whereas a renovation measure maintains the general structural behaviour but increases the dimensions. For instance, Figure 3.2 (top left) shows a full renewal using a diaphragm wall, which affects all mechanisms and completely adapts the structural behaviour. The bottom right figure shows a crest heightening: the structural behaviour remains the same and only the failure probability of overtopping is reduced. In terms of life-cycle performance each type of measure will have different behaviour in terms of degradation of performance, and for renewal measures the impact of different uncertainties on the performance might shift (i.e., it is a substitution measure, see Section 2.6). It is important to include these effects

in planning decisions, as this might affect the reliability over time significantly.

3.3. Optimizing flood defence system investments

The previous section has outlined the requirements for potential optimization methods for the optimization of dike reinforcements on a system level. Research on optimization of flood defence systems has generally focused on two aspects: either determining optimal safety targets for dike segments [e.g. 18, 90, 91], or optimal design alternatives for dike cross sections [e.g. 19, 86]. van Dantzig [34] first elaborated the derivation of optimal flood defence system safety targets. Later this was advanced in the approach used to derive the new safety standards in the Netherlands [5, 16]. As the Linear Programming approach that was used could not deal with non-homogeneous segments it was improved by Brekelmans et al. [92], and later extended to an integer programming approach by Zwaneveld et al. [18], and a graph-based approach by Dupuits et al. [15] and Dupuits [90]. Both can handle non-homogeneous segments and more complex systems. Nevertheless, in order to prevent state space explosion, typically one has to significantly simplify the optimization problem, which is a major issue when analyzing large dike segments. Optimization of cross-sections using a target reliability typically considers multiple failure modes as well as influence of deterioration such that design alternatives can be optimized in time (e.g. Bischiniotis et al. [19], Voortman & Vrijling [91] and Chen & Mehrabani [86]). Nevertheless, optimization of reinforcement planning of dike segments in connection to the optimization at a cross section level has not been addressed explicitly.

Especially for long term planning of flood protection systems decision makers also have to deal with significant uncertainty in future economic value and climate change effects on extreme flood conditions [e.g. 93]. To deal with this, various approaches have been used to incorporate this uncertainty in strategic planning decisions. For instance, real options analysis such as considered by Woodward et al. [83] explicitly accounts for the flexibility of investment decisions, in order to assure that investments are robust under a wide variety of possible future conditions. Kwakkel et al. [75] and Woodward et al. [94] use a model-driven approach using a Multi-objective evolutionary algorithm (MOEA) that can find robust optimal solutions under a wide variety of uncertainties. Such approaches can be valuable if decisions are sensitive to large (future) uncertainties, or for cases with a high degree of complexity [e.g. 95, 96]. A common class of MOEA are genetic algorithms such as the NSGA-II algorithm [97]. An advantage is that these algorithms do not suffer as much from state space explosion as e.g. integer programming, although they do not provide a guaranteed optimal solution.

In other fields the issue of optimal planning of interventions in complex systems has been addressed in the past for both single and multi-objective problems. For instance Bocchini & Frangopol [98] addressed planning of bridge maintenance in a road network aimed at balancing cost and network performance, by using an adapted version of the aforementioned NSGA-II algorithm. Barone & Frangopol [99] evaluated optimal maintenance planning for a bridge of several components using

different performance indicators, and encountered differences in both computational time as well as performance for each performance indicator. For instance, risk-based maintenance was more cost-effective than reliability-based maintenance. Cavdaroglu et al. [100] determined schedules for network restoration measures after a non-routine disruption of an interdependent network using a heuristic solution method. The accuracy of the developed heuristic solution method was high (in comparison with some commercial solution methods), and had a much lower computational cost thus making it more accessible to decision makers. Bagloee et al. [101] solved a very practical prioritization question for road improvement projects using a combination of supervised learning and integer programming in order to optimally schedule projects within a given budget. Also in the field of project scheduling for offshore asset construction [102] and concrete bridges [103] amongst others, there are several (heuristic) approaches that deal with solving scheduling tasks with multiple objectives and uncertainty in present and future conditions. In summary, there is ample work in other fields that we can utilize for optimization of planning of dike segment reinforcements.

In general, both for flood defences and other applications, once problems become interdependent and the number of variables increases, some type of heuristic is applied in order to reach a near-optimal solution. As the problem considered is of the same complexity, we develop a greedy algorithm that employs heuristics based on the engineering problem of improving a dike segment consisting of many non-homogeneous sections (see Sections 2.4 and 3.1). The advantage of this approach is that the heuristics are relatively easy to understand for dike managers, and that it can reach a (near-) optimal solution quickly on commonly available hardware. This is an advantage compared to existing solutions which are typically less transparent and require a simplification of the problem, making them less useful for application in design and planning of dike reinforcement projects.

3.4. Definition of the optimization problem

In this chapter we consider a dike segment of N sections for a period of T years. The current safety level of the segment is significantly below the safety standard, and measures for dike reinforcement have to be determined. A variety of measures for different dike sections is available. The goal is to determine the optimal combination of measures that ensures that the reliability requirement in year t_{req} is met, in a cost optimal way and considering different failure modes. Table 3.1 presents the used notation. In this section we describe the general problem, in the following subsections we describe the solution methods that are used.

Our general objective can be written as follows:

$$\min TC = TR + TLCC \tag{3.2}$$

where TR is the total flood risk cost over the considered period T and $TLCC$ is the

Table 3.1.: *Descriptions of used symbols*

Symbol	Description
N	Set of all dike sections at which investments are possible
n	A dike section index $n \in \|N\|$
T	Set of all years in the analysed period T
t	Considered point in time $t \in \|T\|$
t_{req}	Point in time $t_{\mathrm{req}} \in \|T\|$ where the reliability requirement is to be met
S_{h}	Set of investment options for overtopping
s_{h}	Index of investment option for overtopping $s_{\mathrm{h}} \in \|S_{\mathrm{h}}\|$
S_{g}	Set of investment options for geotechnical failure modes
s_{g}	Index of investment option for geotechnical failure $s_{\mathrm{g}} \in \|S_{\mathrm{g}}\|$
$P_{\mathrm{f,req}}$	Failure probability requirement for all dike sections N
$P_{\mathrm{f,over}}(n, s_{\mathrm{h}}, t)$	Overtopping failure probability of dike section n, given investment option s_{h}, at time t
$P_{\mathrm{f,pip}}(n, s_{\mathrm{g}}, t)$	Piping failure probability of dike section n, given investment option s_{g}, at time t
$P_{\mathrm{f,inst}}(n, s_{\mathrm{g}}, t)$	Instability failure probability of dike section n, given investment option s_{g}, at time t
$LCC(n, s_{\mathrm{h}}, s_{\mathrm{g}})$	Total life-cycle cost (in €) for the combination of investment options $s_{\mathrm{h}}$and s_{g}at dike section n
$D(t)$	Discounted flood damage at year t in €
$C_{\mathrm{int}}(n, s_{\mathrm{h}}, s_{\mathrm{g}})$	Binary value that takes value 1 or 0 and indicates whether measure $s_{\mathrm{h}}$and s_{g}at section n have been taken (1) or not (0)
$D_{\mathrm{int}}(t, n, s_{\mathrm{h}})$	Binary value that indicates if section n with measure s_{h}implemented is the weakest section (1) for overtopping in year t or not (0)
TC	Value of the objective function of all costs over the period $\|T\|$
TR	Total flood risk cost over the period $\|T\|$
$TLCC$	Total Life Cycle Cost over the period $\|T\|$
$BC(n, s_{\mathrm{h}}, s_{\mathrm{g}}, C_{\mathrm{int}})$	Benefit-cost ratio of investment s_{h}, s_{g} at section n with initial situation C_{int} at the beginning of a greedy search iteration

total life cycle cost of all measures. TR is defined as:

$$TR = \sum_{t \in T} P_{\text{f,total}}(t) \cdot D(t) \tag{3.3}$$

with:

$$P_{\text{f,total}}(t) = 1 - \prod_{n \in N} \Bigg(\sum_{s_{\text{h}} \in S_{\text{h}}} \sum_{s_{\text{g}} \in S_{\text{g}}} \left(\widehat{P_{\text{f,pip}}} \cdot \widehat{P_{\text{f,inst}}} \cdot C_{\text{int}}(n, s_{\text{h}}, s_{\text{g}}) \right) \cdot \left(1 - \left(P_{\text{f,over}} \cdot D_{\text{int}}(t, n, s_{\text{h}}) \right) \right) \Bigg) \quad \text{for } t \in T, \tag{3.4}$$

where $\widehat{P_{\text{f,pip}}} = 1 - P_{\text{f,pip}}(n, s_{\text{g}}, t)$ and $\widehat{P_{\text{f,inst}}} = 1 - P_{\text{f,inst}}(n, s_{\text{g}}, t)$.

$TLCC$ is defined as:

$$TLCC = \sum_{n \in N} \sum_{s_{\text{h}} \in S_{\text{h}}} \sum_{s_{\text{g}} \in S_{\text{g}}} LCC(n, s_{\text{h}}, s_{\text{g}}) \cdot C_{\text{int}}(n, s_{\text{h}}, s_{\text{g}}). \tag{3.5}$$

This is subject to the following constraints:

$$\sum_{s_{\text{g}} \in S_{\text{g}}} \sum_{s_{\text{h}} \in S_{\text{h}}} C_{\text{int}}(n, s_{\text{h}}, s_{\text{g}}) = 1 \text{ for } n \in N \tag{3.6}$$

$$\sum_{n \in N} \sum_{s_{\text{h}} \in S_{\text{h}}} D_{\text{int}}(t, n, s_{\text{h}}) = 1 \text{ for } t \in T \tag{3.7}$$

$$\sum_{\substack{s_{\text{h}} \in S_{\text{h}} \\ P_{\text{f,over}}(n, s_{\text{h}}, t) > \\ P_{\text{f,over}}(n^*, s_{\text{h}}*, t)}} C_{\text{int}}(n, s_{\text{h}}, s_{\text{g}}) + \sum_{n_{\text{h}}} \sum_{\substack{s_{\text{h}} \in S_{\text{h}} \\ P_{\text{f,over}}(n, s_{\text{h}}, t) \leq \\ P_{\text{f,over}}(n^*, s_{\text{h}}^*, t)}} D_{\text{int}}(t, n, s_{\text{h}}) \leq 1 \quad \text{for } t \in T, n \in N, n^* \in N, s_{\text{h}}^* \in S_{\text{h}} \tag{3.8}$$

$$P_{\text{f,total}}(t) < P_{\text{f,req}} \text{ for } t \in T, t_{\text{req}} \leq t \leq t_{\text{horizon}} \tag{3.9}$$

$$C_{\text{int}}(n, s_{\text{h}}, s_{\text{g}}) \in 0, 1 \text{ for } n \in N, s_{\text{h}} \in S_{\text{h}}, s_{\text{g}} \in S_{\text{g}} \tag{3.10}$$

$$D_{\text{int}}(t, n, s_{\text{h}}) \in 0, 1 \text{ for } t \in T, n \in N, s_{\text{g}} \in S_{\text{g}} \tag{3.11}$$

Equation (3.2) describes the objective of our approach, namely to minimize the total cost consisting of flood risk and investment costs over the considered time period T, which are both written more explicitly in Equations (3.3) and (3.5). Equations (3.6) and (3.7) describe relatively simple constraints that ensure that there is only 1 combination of s_{h} and s_{g} chosen per dike section, and that for each time t only 1 section n is the weakest for overtopping. Equation (3.8) is a bit more complicated, but it ensures that if there is an investment s_{h} at a section n, it cannot be the weakest section n^*, s_{h}^* at the same time, so investments in reducing $P_{\mathrm{f,over}}(n, s_{\mathrm{h}}, t)$ are always done at the section with the highest $P_{\mathrm{f,over}}(n, s_{\mathrm{h}}, t)$. Equation (3.9) is a constraint that ensures that after some year t_{req} the system reliability requirement is satisfied. This is optionally limited by t_{horizon}, which is the horizon for which this is to be satisfied. Equations (3.10) and (3.11) ensure that the variables $C_{\mathrm{int}}(n, s_{\mathrm{h}}, s_{\mathrm{g}})$ and $D_{\mathrm{int}}(t, n, s_{\mathrm{h}})$ are binary.

Before the optimization is started $P_{\mathrm{f,inst}}(n, s_{\mathrm{g}}, t)$, $P_{\mathrm{f,over}}(n, s_{\mathrm{h}}, t)$, $P_{\mathrm{f,pip}}(n, s_{\mathrm{g}}, t)$, $LCC(n, s_{\mathrm{h}}, s_{\mathrm{g}})$ and $D(t)$ are precalculated. These are input for both approaches used to minimize the objective function.

3.5. Solving a flood defence reinforcement planning problem

The problem as described above is implemented as a Mixed Integer Programming (MIP) problem in CPLEX 12.9 [104] and solved using branch-and-cut. An advantage of branch-and-cut is that it can be used to exactly solve integer programs with optimality guarantee [105]. However, with the number of investment options that is relevant for a typical dike segment, this is only feasible for relatively small segments up to about 13 dike sections with 16 GB available RAM, which is much smaller than our real world problem. As an illustration — in the case study in Section 3.6 we consider $\approx 10^{100}$ possible combinations of reinforcement measures.

To overcome issues with computational speed and hardware we develop a greedy search algorithm. Greedy algorithms are a class of algorithms that use the locally optimal choice at each stage in order to obtain or approach the global optimum [106]. This means that these algorithms can handle much larger state spaces, which is useful in the context of the large dike segments that we consider. An important property of greedy algorithms is that it never reverses choices but always continues with the next optimal choice until it finds a solution or is stopped.

A potential advantage of greedy algorithms is that the heuristic rules that are used for determining the optimal steps are often easy to understand and can be adapted to the problem at hand. In our case we can use the formulations of segment reliability as well as the principle of total cost optimization as basis for the heuristics.

3.5.1. A greedy algorithm for planning of flood defence reinforcements

In this section we introduce the greedy search algorithm. In the implementation of heuristics, we need to ensure two main points: Firstly, that the search method is in

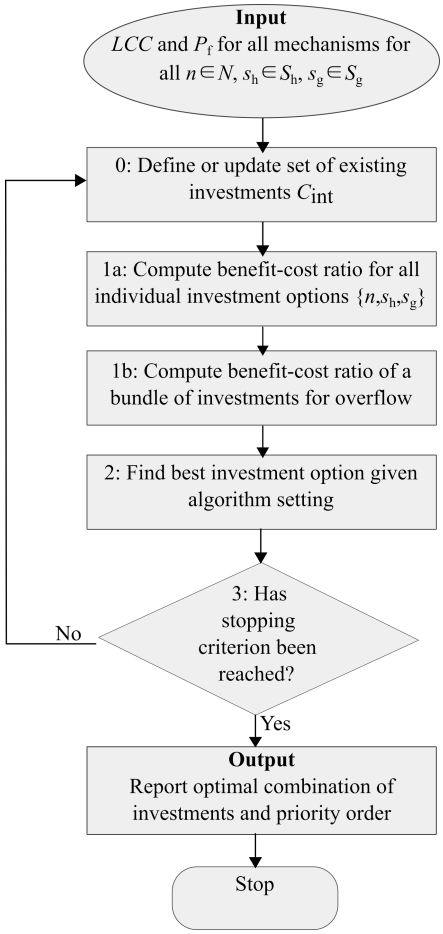


Figure 3.3.: *Outline of the steps taken in the greedy search algorithm.*

line with the objective of finding minimal Total Cost, and secondly that the relation between element and system reliability is properly dealt with for all failure modes.

For the search method the work by Špačková & Straub [107] can be used, who demonstrated that for a case without budget limitation the optimal solution is found if $-\delta C < \alpha \cdot \delta R$, with δR the risk reduction and δC the cost increment. α is an arbitrary factor indicating risk averseness, so how much risk has to be reduced for a cost increment δC (we assume risk neutrality so $\alpha = 1$). This criterion guarantees that Pareto optimal solutions are found in all cases (although not all Pareto optimal solutions are found). If deterministic costs are assumed, such as in our case, this is equal to the benefit-cost ratio $BC = E[(-\Delta R)/\Delta C]$. Špačková & Straub [107] considered single investment strategies to achieve a certain protection level. In our case we aim to determine the combination of measures that is on (or very close) to the Pareto optimal solution. If the utility of our investment is the risk reduction, we would expect that for each subsequent measure that we take, our marginal risk reduction will reduce in accordance with the law of diminishing marginal returns. Thus if we find a path where we continuously maximize the BC until $E[(-\Delta R)/\Delta C] = 1$ we should obtain at least a local optimal solution. Hence, the main parameter used for evaluating the steps in the greedy algorithm is the benefit-cost ratio, which is defined as:

$$BC(n^*, s_{\mathrm{h}}^*, s_{\mathrm{g}}^*, C_{\mathrm{int}}(N, S_{\mathrm{h}}, S_{\mathrm{g}})) = \frac{TR - TR^*}{\sum_{n^* \in N} \sum_{s_{\mathrm{h}}^* \in S_{\mathrm{h}}} \sum_{s_{\mathrm{g}}^* \in S_{\mathrm{g}}} LCC(n^*, s_{\mathrm{h}}^*, s_{\mathrm{g}}^*)} \tag{3.12}$$

where TR denotes the Total Risk of flooding over a period T for a set of measures defined by $C_{\mathrm{int}}(n, s_{\mathrm{h}}, s_{\mathrm{g}})$, and TR^* denotes the same for a case where measures s_{h}^*,s_{g}^* have been taken at sections n^*. $LCC(n^*, s_{\mathrm{h}}^*, s_{\mathrm{g}}^*)$ denotes the cost of these measures. If $BC(n^*, s_{\mathrm{h}}^*, s_{\mathrm{g}}^*, C_{\mathrm{int}}(N, S_{\mathrm{h}}, S_{\mathrm{g}})) < 1$ for all combinations of $n^*, s_{\mathrm{h}}^*, s_{\mathrm{g}}^*$ the optimum with minimal total cost TC has been reached, as the marginal total costs are smaller than 0. This can be either a local or global optimum, depending on the performance of the search routine. In our case this search routine is based on the definitions of system reliability in Equations (2.4) and (2.5).

As was outlined in Section 3.1 the relation between element and system reliability for overtopping and geotechnical failure modes (piping erosion and slope instability) differs. In order to prevent missteps in deriving the local optimal solution in the greedy search algorithm, this has to be dealt with properly in the implemented heuristics. Therefore we need to implement different rules for deciding on investments to improve geotechnical reliability and investments to improve overtopping reliability. The main steps of the algorithm are listed below and displayed in the flowchart in Figure 3.3.

Input for the search algorithm are precalculated arrays of failure probabilities and $LCC(n, s_{\mathrm{h}}, s_{\mathrm{g}})$ for all measures $n \in N$, $s_{\mathrm{h}} \in S_{\mathrm{h}}$ and $s_{\mathrm{g}} \in S_{\mathrm{g}}$. Initial failure probabilities are given per dike section.

Step 0: at the beginning of each iteration the set of existing measures $C_{\mathrm{int}}(n, s_{\mathrm{h}}, s_{\mathrm{g}})$ is updated with all measures that have been prioritized in previous iterations.

Step 1a: In the first part of the second step $BC(n^*, s_{\mathrm{h}}^*, s_{\mathrm{g}}^*, C_{\mathrm{int}}(N, S_{\mathrm{h}}, S_{\mathrm{g}}))$ is computed for all individual measures, based on the existing situation $C_{\mathrm{int}}(n, s_{\mathrm{h}}, s_{\mathrm{g}})$. Due to the formulation of system reliability for geotechnical failure modes (see Equation (2.4)) this approach works well for determining optimal priority orders for segments dominated by these failure modes. The reason is that each dike section contributes to the overall risk, so improving any individual dike section will have a direct influence on the total risk.

Step 1b: While considering individual sections is adequate for geotechnical measures, overtopping system reliability is governed by the weakest section (see Equation (2.5)). Therefore we introduce a second heuristic where we compute the BC-ratio for a bundle of measures that reduce $P_{\mathrm{f,over}}(n, s_{\mathrm{h}}, t)$. By only considering individual dike sections as in Step 1a, situations can occur where improving overtopping reliability at a single section has a low BC-ratio. Considering a bundle of different reinforcement measures at different sections can have a much higher BC-ratio as all weak sections are improved simultaneously. This can result in a much larger marginal increase in system reliability, and thus a larger BC-ratio. This step consists of the following substeps:

1. Sort the investment options based on $LCC(n, s_{\mathrm{h}}, s_{\mathrm{g}})$.
2. Filter the options such that only options $s_{\mathrm{h}} \in S_{\mathrm{h}}$ combined with the existing investment option s_{g} at section n are considered.
3. Determine a priority order of measures, where each time the weakest section is improved with the smallest investment step available from $s_{\mathrm{h}} \in S_{\mathrm{h}}$.
4. Take the set of investment options with the highest BC-ratio.

Step 2 is to find the optimal investment option based on the BC-ratios obtained from steps 1a and 1b. If the BC-ratio from step 1b is higher than the maximum BC-ratio from step 1a, the bundle of investments obtained from step 1b is implemented. Otherwise the algorithm determines the next investment step based on a greediness factor f_{c}, where it holds that the investment step has to have a BC-ratio that is f_{c} larger than the best measure at all other dike sections. This is a factor to set the *greediness* of the algorithm, with a larger factor implying larger steps, but also a less cautious and more error-prone search routine. Additionally, multiple runs with different settings for f_{c} might yield different solutions, meaning that the overall accuracy of the approach increases as the best of those different solutions can be chosen.

Step 3: if the BC-ratio of the best available investment option is smaller than the stopping criterion (default setting: $BC\text{-ratio} < 0.1$) the optimization is stopped. Note that in most cases a stopping criterion $BC\text{-ratio} < 1$ should yield the optimal solution, but due to the properties of a greedy search in some cases a measure with $BC\text{-ratio} < 1$ is followed by one with a $BC\text{-ratio} > 1$ (due to the dependence of BC-ratio on preceding measures). After reaching the stopping criterion the optimal solution is obtained based on the minimum

total cost of all steps in the search path, or the point where the solution meets the target reliability requirement in the years for which this is required. Note that the latter could also be used as a stopping criterion.

Output: the output of the greedy search is a sequence of arrays $C_{\text{int}}(n, s_{\text{h}}, s_{\text{g}})$ which describe the obtained optimal solution as well as the priority order of investments leading to that solution. This is an advantage compared to for instance a MIP solution, where only the array $C_{\text{int}}(n, s_{\text{h}}, s_{\text{g}})$ for the optimum is obtained.

3.5.2. Validation of the greedy algorithm

In many problems greedy algorithms have been found to achieve (near-)optimal solutions, but in general it is hard to prove that a solution is optimal. In order to show that the heuristics yield (near-)optimal solutions, we evaluate the performance of the greedy algorithm and compare it with a Mixed Integer Programming (MIP) implementation in CPLEX 12.9 [104] using branch-and-cut, for a large number of different dike segments and for different combinations of available measures.

A typical run of the greedy search algorithm yields a stepwise prioritization of dike reinforcement measures that eventually ends at or very close to the global optimal solution, which consists of the minimum sum of $TLCC$ and TR. Figure 3.4 shows results for a system with 5 dike sections. We see that the greedy search (red) reaches the global optimum, and follows the Pareto front for $TLCC$ and TR (black) computed using the MIP implementation with variable budget limits (i.e., where $TLCC$ is constrained). This shows that, especially closer to the optimal solution (blue diamond) the investment path of the greedy search not only finds the optimal solution, but the intermediate steps are also (near-)optimal for that budget.

However, this is just 1 case, and in order to validate for a large number of different system configurations, 2800 different configurations were evaluated for different sets of intervention measures, and different system sizes. A full description of the validation can be found in Appendix A, but the main finding is that average difference in Total Cost at the optimal solution is 0.04%, which is negligible, especially in the context of other large uncertainties in dike reinforcement projects. Thus we can conclude that this greedy search algorithm is reliable for the envisaged application in the next section.

3.6. Case study application

This section presents an application of the greedy search algorithm to the planning of a reinforcement project for a dike segment along the Lek river in the Netherlands. The main aim of this section is to illustrate the practical applicability of the developed approach, and show the main advantages of using an optimization algorithm in the process of planning and design compared to the commonly used design approach based on cross-sectional target reliability.

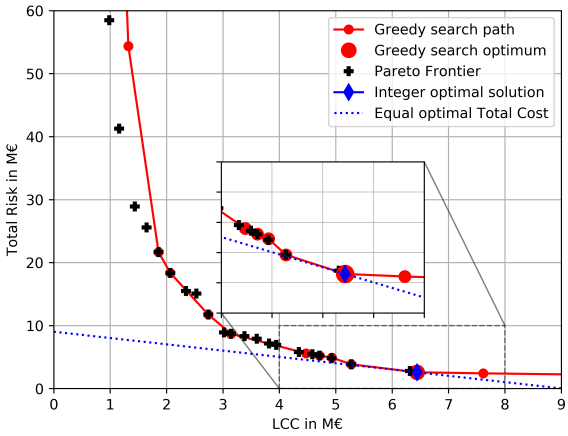


Figure 3.4.: *Example result of a system with 5 dike sections. Red line shows the path of the greedy search, where the large dot denotes the optimal solution. Blue diamond indicates the optimum found using branch-and-cut in CPLEX 12.9. The black pluses denote the Pareto Frontier derived from several branch-and-cut computations with a budget limit. Note that here the TR and TLCC are displayed as the conflicting objectives, whereas in the optimization routine these are summed and considered as a single objective.*

3.6.1. Case description

We consider dike segment 16-4 along the Lek river in the Netherlands, located between the towns of Everdingen and Ameide. The length of this segment is approximately 20 km, and it consists of 45 sections of which 41 are considered in the calculations. The three main failure modes are overtopping, piping erosion and inner slope instability, for which calculations have been made using the statutory safety assessment tools [6, 52]. As such, we have estimates of the predicted reliability over the considered period of 100 years. Further details on the approach for deriving the reliability are given in Appendix B. If a flood occurs along this segment the estimated damage is 23 billion € [108]. Since the introduction of the new safety standards in 2017 the segment failure probability has to be less than 1/10,000 per year ($\beta_{\text{req}} > 3.72$). Due to the new standards (amongst other reasons) many flood defence reinforcement projects are initiated, and the improvement of dike segment 16-4 is one of them. The goal is to meet the safety standards nationwide by 2050. In this study we aim to find an optimal strategy to achieve this for segment 16-4, such that the safety standard is met in 2050 and until at least 2075. This means that Equation (3.9) is now also used, with $P_{\text{f,req}} = 1/10,000$, $t_{\text{req}} = 2050$ and $t_{\text{horizon}} = 2075$. We assume that the reinforcement can start in 2025.

In the case study we consider a specific set of measures as specified by the local water authority. For all sections we consider the following reinforcement measures:

- Soil reinforcement (berm widening and crest heightening) in 2025 and 2045,

- Diaphragm wall
- Vertical Sandtight Geotextile (VSG)
- Stability screen

From these measures the VSG and stability screen, which are both partial reinforcement measures, can be combined with a soil reinforcement. Other assumptions are described in Appendices A and B.

In the analysis we compare the total cost optimization using the greedy algorithm introduced in Section 3.5 with the target reliability-based investment based upon cross-sectional reliability targets. We assume that these targets have to be met in 2075 (i.e., 50 year design horizon).

3.6.2. Case study results

In order to assess the reliability deficit for the segment under consideration we first compute the current and predicted reliability without taking any measures, which is displayed in pane (a) of Figure 3.5 for the year 2075. Here we see that for most sections either piping erosion or inner slope instability has a reliability deficit, and that overall (bar chart on the right) the system reliability (black line) $\beta \approx 1$, which is extremely low, and much lower than the required value. Another thing that quite clearly emerges from the figure is that neighbouring sections often have similar issues: sections 34b-40 have a large deficit for inner slope instability, whereas sections 16-22 have a large deficit for piping erosion.

The reliability information from pane (a) of Figure 3.5, together with information on different available measures, their effects on reliability and their respective costs can be used to generate design alternatives for the entire segment, both with a greedy search optimization and a reliability-based design approach based on cross-sectional requirements.

Panes (b) and (c) of Figure 3.5 show the selected measures and resulting reliability in 2075, for an investment based on the greedy search algorithm. Panes (d) and (e) show the same for a target reliability-based investment. Pane (f) shows the life cycle costs of both approaches. There are a few distinct differences between the two methods. First of all, the target reliability-based investment results in a higher system reliability due to some conservatism in the cross-sectional target reliability values. Also, as for the target reliability-based approach each section has to satisfy a target reliability for each failure mode, there are many sections where expensive diaphragm walls are needed to meet the requirements. When using the greedy search algorithm such investments are avoided by increasing reliability at other dike sections, or by using partial renewal measures (i.e., a Vertical Sandtight Geotextile (VSG, inversed triangle) or stability screen (SS, circle), see Appendix B for specifications). This is not feasible for the target reliability-based investment as it becomes impossible or extremely expensive to meet requirements for other failure modes. For instance: dike section 38 is improved using a diaphragm wall when using the target reliability-based approach at a cost of $\approx$ 17 M€. In order to prevent these large expenses in the optimized approach only a much cheaper

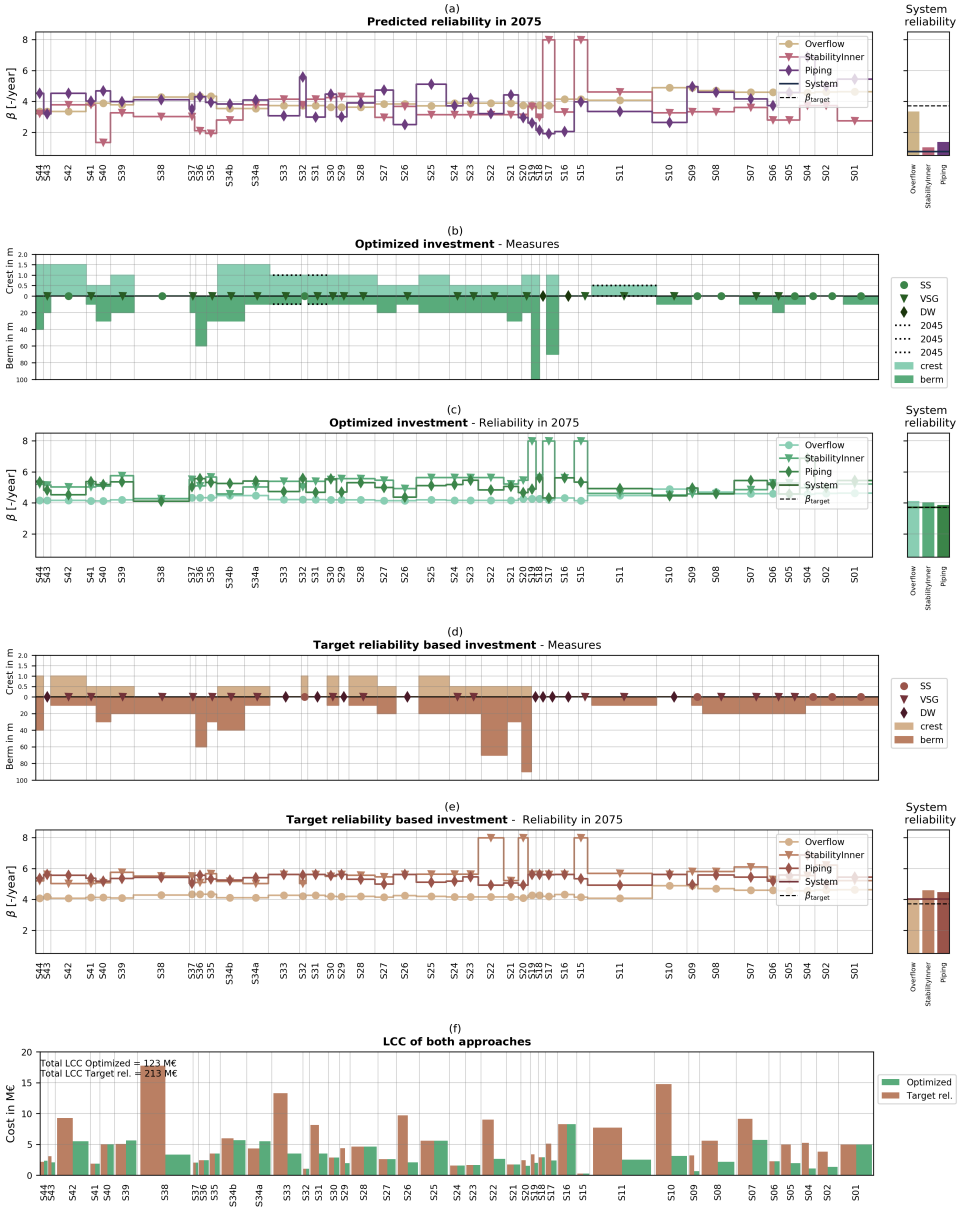


Figure 3.5.: *Input and results for both planning approaches. Pane a shows the reliability in 2075 without investments, panes b and d show measures taken using the greedy optimization and target reliability approaches; c and e show resulting reliability in 2075. Pane f shows the resulting life-cycle cost of both approaches. Vertical grid lines indicate boundaries between sections, and the width of sections indicates their relative length (total length of the segment is 20 kilometres).*

stability screen (cost ≈ 3.5 M€) is constructed, resulting in lower reliability, but also only a fraction of the cost. This is then compensated by using more extensive measures at other dike sections (e.g., section 34a), where the benefit-cost ratio of extra investments is larger. This is a degree-of-freedom ignored by the cross-sectional target reliability approach. When looking at resulting system failure probabilities (bars on the right) there are two main differences between the two methods:

- For the optimized approach the resulting system reliability for overtopping is highest, whereas for a target reliability-based approach it is lowest. There are clearly more investments in crest height for the optimized approach (see Figure 3.5b). Examples are sections S34a, S34b, where higher crest increases are planned.

- The overall system reliability in 2075 is higher for the target reliability-based investment, most notably due to the fact that the reliability for inner slope stability and piping erosion is higher. The main reason is that for many sections a diaphragm wall or Vertical Sandtight Geotextile is applied. These measures have a reliability that is much higher than the required reliability, yet there is no cheaper alternative that also meets the requirement. This explains the higher system reliability for these failure modes.

Figure 3.5f shows the LCC for both approaches. Here we observe that for the target reliability-based approach some sections have very high costs (e.g., 10, 33 and 38), whereas the costs for the optimized approach are much more evenly spread across the different sections. The difference in total LCC is very large: using the greedy search optimization as a basis for planning reduces the total investment from 213 M€ to 123 M€, a reduction of about 42%.

One of the major advantages of the greedy search approach that was mentioned is that it also yields a priority order of measures based on the search path. This order can give decision makers insight into the priorities for improving a dike segment, and can help making a risk-informed selection of parts of the project, if not all budget is readily available. The priority order is given in Figure 3.6, which shows the LCC relative to the reliability index β in the year 2075. The green line denotes the search path for the optimized investments, whereas the brown line denotes the investments based on target reliability, ordered by the initial reliability of the sections (so the weakest is displayed first). For the greedy search path it can be observed that many small increments are taken, and that especially in the beginning the line is very steep, meaning that a large increase in reliability is achievable for a limited amount of money, but as the overall reliability increases, the marginal risk reduction for additional investments decreases.

One particularly important challenge in flood defence reinforcement projects is to properly deal with long term uncertainty in for instance economic growth and increase in hydraulic loads. Figure 3.7 shows the resulting measures for a case where the increase in water level has been multiplied by a factor 3 for all sections. By comparing with pane (b) of Figure 3.5 we can observe that investments in crest height increase are larger, in order to cope with the higher water levels. However,

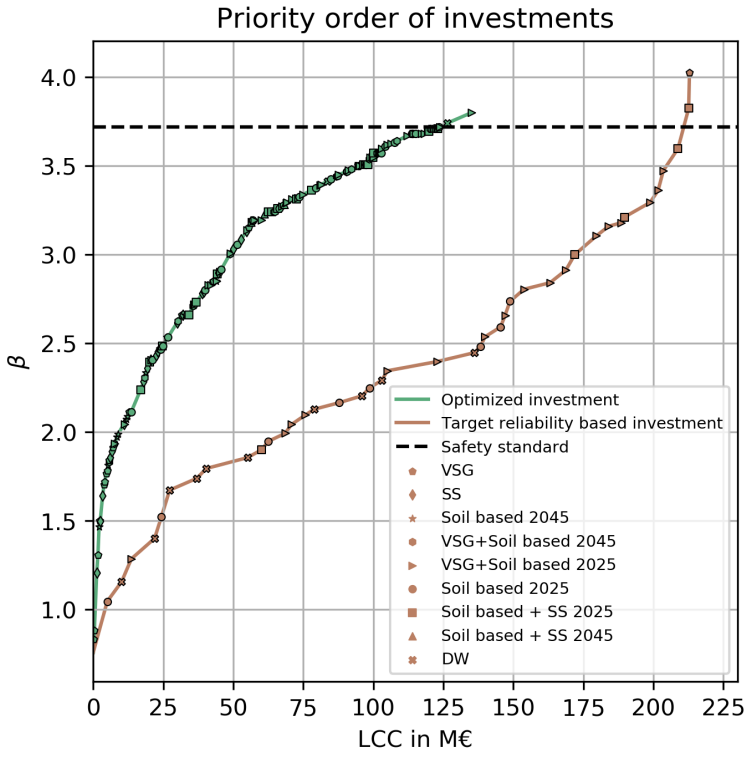


Figure 3.6.: *Priority order for optimized investments (green) and investments based on target reliability (brown). Markers denote different types of measures at different dike sections, β is the reliability index in 2075.*

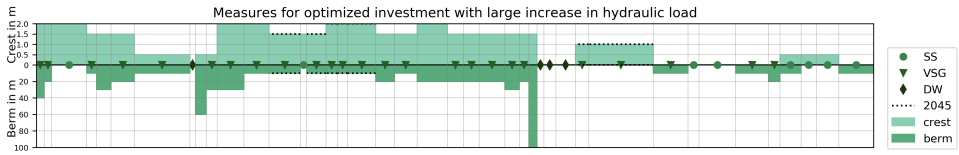


Figure 3.7.: *Measures for optimized investment with 3 times higher water level increase rates.*

the influence on both the priority order and the investments in geotechnical measures is very similar, which demonstrates that for this case study the added value of including multiple scenarios for hydraulic load increase is limited, both in terms of the prioritized measures as well as the potential for future extension.

3.7. Discussion

In this chapter we demonstrate how the cost effectiveness of dike reinforcements can be improved using a system optimization. The advantage of this approach is that we do not constrain the solution space, which is a limitation of the commonly used target reliability-based approach. We use a greedy search algorithm with heuristics based on system reliability rules and benefit-cost ratio of measures that yields near-optimal plans for reinforcement of dike segments, at a much lower cost than the commonly applied approach based on cross-sectional target reliability. From a comparison with a Mixed Integer Programming (MIP) approach it is shown that in the majority of cases the algorithm finds the optimal solution, and in the other cases it is close (i.e., very small differences in Total Cost (TC)). Even for the systems where the algorithm performs relatively bad (difference in TC $> 1\%$), this inaccuracy is minor in comparison to other major uncertainties in the design, such as estimates of reinforcement cost and geotechnical strength parameters.

It has to be noted that in this study we only consider dike sections as part of the segment, but in practice there can also be hydraulic structures (e.g. inlet sluices) that are part of the flood defence. As typically the reliability of these structures can be computed, these can also be included in the analysis. Thus, the approach is not solely applicable to dikes, but to flood defences in general.

The method can also be used with more advanced methods such as probabilistic reliability calculations and more advanced methods for computing system reliability including correlation between sections and mechanisms, such as the Equivalent Planes method [50]. In our schematization of the dike system we use semi-probabilistic estimates for reliability and relatively simple approaches to model the correlation between different dike sections. As our case study concerns the early planning phase of a reinforcement project it is not yet sensible to use more advanced methods due to uncertainties in for instance geotechnical strength parameters. The method is useable with more advanced computation methods, the only requirement is that for each failure mode per section the reliability (in time) can be computed, and that this can be translated to a system reliability estimate.

Of course, when changing major underlying assumptions a re-evaluation of the accuracy of the greedy algorithm is advisable.

In the considered case study we do not explicitly deal with the wide variety of potential future scenarios for increasing hydraulic load or socio-economic conditions as has been done in other studies (e.g., [75, 83, 109]). For the case study it is shown that, mainly due to the dominance of geotechnical failure modes, different scenarios for hydraulic loads have little bearing on the priority order and type of investments. However, this might not always be the case, and in such cases including these future uncertainties is recommended. This can be done either by using probabilistic estimates of future uncertainty [e.g. 110], or by considering multiple scenarios for which the different dike segment designs can be evaluated (in line with e.g. [75]). The latter is quite feasible as the current computation time is still relatively limited (approximately 10 minutes for 1 evaluation of the case study). It has to be noted however, that such an extension and the subsequent increase in computation time makes the approach less useable in the design process, where it is often desired to have a practical tool that can be used to quickly evaluate various design considerations.

In our approach we also assumed that the reinforcement costs for different dike sections are independent. In reality this is not necessarily the case: especially if a large project is cut up in different fragmented small projects overhead costs may rise, resulting in higher overall costs. In our case, as there are only 2 moments of investment this is not relevant, but it might be for other cases. Correlations between the cost of different otherwise independent sections can not yet be dealt with in the algorithm. It has to be noted that accounting for such correlations is not common practice in cost estimates for dike reinforcements [111].

The approach has clear benefits compared to the commonly applied reliability based design approach using cross sectional requirements: the overall life-cycle costs are about 42 % lower for an optimized design based on the greedy search algorithm, while the reliability requirement is met in both cases. It has to be noted that in practice the target reliability-based approach is also refined throughout a reinforcement project, for instance by slightly altering the different cross-sectional requirements between mechanisms or sections. Then the difference would become slightly smaller, albeit in an ad hoc manner. It has to be noted that the cost savings depend strongly on the accuracy of the input, specifically the reliability estimates. For this and other methods for planning flood defence reinforcements it is therefore advised to ensure that trustworthy reliability estimates are available.

An added advantage of the optimized approach is that it helps focusing attention on the reinforcement of the most important parts of a large and complex dike segment. This structures the technical challenge which aids in adding risk-informed information to the multi-objective task of improving dike systems, where also other aspects of spatial planning have to be dealt with. For instance, in our case study (Section 3.6) we selected measures such that in both moments of investment considered (now and in 20 years), only one type of investment with a major impact could be done in order to limit nuisance to inhabitants. Other secondary objectives could also be included in the choice of measures, such that appropriate risk informa-

tion is obtained, also for other stakeholders with different objectives. The approach facilitates such considerations, first of all by enabling quick evaluation of the influence of such restrictions on the overall solution, and secondly by providing insight into the importance of different measures through analysis of the priority order of investments. Additionally, the impact of the reinforcement on the environment is also limited due to the more optimized system configuration.

While in this study the approach was applied to an investment decision for a dike segment, it is also useable for other types of decisions, such as optimizing investments over multiple dike segments. For instance, in the Netherlands about 1500 kilometres of flood defences have to be improved, resulting in (currently) a list of over 50 projects. The greedy search algorithm could also be used to find a total cost optimal prioritization that balances national flood risk and reinforcement costs.

Compared to other optimization approaches such as Mixed Integer Programming, the greedy search routine encounters less issues with state space explosion, making it more suitable for systems with many components. Of course one could apply for instance a Mixed Integer Programming approach, whilst expanding the available hardware. However, the dimensions of especially the constraint in Equation (3.8) increase almost quadratically with the number of sections (and solutions), so this would still be challenging for segments of over 30 independent sections. An added advantage of the greedy search routine is that it is very explainable, also to stakeholders with different expertise. This is an advantage compared to Mixed Integer Programming, and for instance genetic algorithms.

3.8. Conclusions

In this section we have considered a decision problem where a large-scale reinforcement of a dike segment is optimized. We have proposed a design approach to optimize reinforcements of a dike system. To make this computationally possible we used a greedy search algorithm, for which we derived heuristic rules that can be used for planning dike reinforcement projects in flood defence systems with a large number of independent elements. It was demonstrated that for a real world dike system the approach results in a 42 % reduction of investment costs compared to the method that is typically applied in the same phase of a reinforcement project. An additional advantage is that a priority order of measures is determined, which is useful in making and explaining risk-informed decisions during the planning and design of dike reinforcement projects.

The greedy search algorithm employs two main heuristic rules: use of the benefit-cost ratio to select the local optimal investment, and relations between system and element reliability to translate investments at a dike section to risk reduction for the system as a whole. The approach is very useful in dike reinforcement projects as it offers good accuracy compared to a Mixed Integer Programming implementation, while it hardly suffers from state space explosion. From an analysis of 2800 different realistic dike segments the average difference in objective value was only 0.04 %, which is negligible compared to other uncertainties in dike reinforcement projects.

The key advantage of the proposed approach compared to a target reliability-based approach based on cross-sectional reliability targets, is that the entire solution space of reinforcement measures can be used, rather than only those reinforcement measures for which the cross-sectional targets for all failure modes are met. The differences found in terms of life-cycle cost are a clear demonstration of the potential loss of utility that is invoked by restricting decisions through (in)formal choices in a decision analysis, such as imposing standardized cross-sectional reliability targets.

In this chapter we did restrict the solution space to some extent, as we did not include investments in uncertainty reduction measures, for instance through monitoring, proof load tests or additional site investigation. The next chapter will consider this in the context of flood defence reinforcements.

4

Uncertainty reduction for dike reinforcement projects

Efforts to solve a problem must be preceded by efforts to understand it.

Herbert Simon

In the previous chapter we have successfully applied a 'system thinking' approach towards design of flood defence systems, and as such optimized reinforcement measures at a system level. However, investment in uncertainty reduction of time-invariant uncertainty was not included in the analysis. As the previous chapter demonstrated that uncertainties in soil parameters are an important cost driver, reduction of these uncertainties is also a potentially promising solution. Therefore in this chapter we consider uncertainty reduction in the context of dike reinforcement projects. We apply a decision tree framework using Bayesian pre-posterior analysis in order to quantify the joint benefits (Value of Information (VoI)) of various methods for reducing uncertainty. The context chosen is that of a dike reinforcement, for which we consider uncertainty through proof load tests and pore pressure monitoring at a section, and pore pressure monitoring at a segment level.

The majority of the first case study has been published in van der Krogt, M. G. et al. Value of information of combinations of proof loading and pore pressure monitoring for flood defences. *Structure and Infrastructure Engineering,* 1–16. doi:10.1080/15732479.2020.1857794 (2020).
Parts of the second case study have been published in Klerk, W. J. et al. *Influence of monitoring on investment planning of flood defence systems* in *Proceedings of the 7th International Symposium on Geotechnical Safety and Risk (ISGSR 2019): State-of-the-Practice in Geotechnical Safety and Risk* (eds Ching, J. et al.) (Research Publishing, Singapore, Taipei, Taiwan, 2019), 792–797. doi:10.3850/978-981-11-2725-0_IS4-10-cd.

IN the previous chapter we considered a system optimization for dike reinforcements. However, within the context of such reinforcement projects, many efforts can be undertaken to gather data through for instance site investigation [8] or pore pressure monitoring [21]. This can aid in reducing the influence of uncertainty in for instance soil strength parameters important for slope instability, or uncertainty in permeability of the subsoil which is important for piping erosion (and/or slope instability). In Chapter 3 we did not explicitly consider uncertainty reduction in the context of the considered dike reinforcement, but this will be the topic of this chapter.

For the case study considered in Chapter 3, especially the reliability estimates for piping erosion and slope instability were very low for several dike sections, which is typically caused by large uncertainty in strength parameters dominant for these mechanisms (e.g., spatial variability, measurement and transformation uncertainty [48, 49]). As many of these uncertainties are epistemic and reducible [114], this might justify investments in uncertainty reduction. However, it is not always clear-cut which uncertainty reduction method will provide the largest value, and this is also something which can be very case-specific. In this chapter we will consider two decision problems in a pre-posterior decision analysis.

In Section 4.2 we consider a reinforcement decision for a single dike section, represented by a cross-section. This dike section is sensitive to slope instability and is to be reinforced using a stability berm. In the time available before the reinforcement (a period of 5 years) we consider a sequential investment in reduction of uncertainty through a proof load test and through pore pressure monitoring.

However, as we do only consider a single dike section in this case study, the decision problem considered in this case does not provide insight into some of the relevant aspects when reducing uncertainty in the context of a large scale dike reinforcement project. To that end, in Section 4.3 we consider a slightly different case, based upon Chapter 3. Here we consider a dike segment with 5 sections where a variety of reinforcement measures is available, and where the uncertainty on the permeability of the piping-sensitive aquifer can be reduced through pore pressure monitoring. Through a pre-posterior analysis we consider at which dike section pore pressure monitoring provides the largest value. This gives insight into some of the aspects involved when considering uncertainty reduction in a larger dike system.

Together, both cases provide practical insight into the most relevant factors of reducing uncertainty in the context of dike reinforcement projects and advance both key topics 1 and 2. These general findings are further discussed in Section 4.4.

4.1. Reducing time-invariant uncertainty in flood defence reliability

Many of the uncertainties that are related to relatively low reliability estimates for geotechnical failure modes are time-invariant uncertainties. This means that a single investment in uncertainty reduction can yield benefits for a long period of

time, for instance for the entire design period, and even beyond. There is a variety of methods available for this, of which amongst others additional site investigation [8] and pore pressure monitoring [21, 29] have been considered in literature. Another method to demonstrate the safety of structures is proof loading, for example proof pile load tests to verify the reliability of foundations [115]. All these methodes are aimed at reducing uncertainties that are predominantly time-invariant. In this chapter we consider two uncertainty reduction methods for dikes: pore pressure monitoring and proof loading. Pore pressure monitoring will be considered in both cases, proof loading only in the first case in this chapter.

Pore pressure monitoring is aimed at reducing uncertainty on the position of the phreatic line [65], representing the response of pore water pressures in a conductive soil layer to hydraulic loads. Such responses are typically dependent on the hydraulic conductivity (or permeability) of the dike material, or of the aquifer underneath, both of which are often heterogeneous. Hence, given the typically limited amount of available measurements, this is often a very relevant uncertainty in reliability estimates for slope stability and piping erosion [49]. An important aspect of pore pressure monitoring is that the information obtained (resulting in uncertainty reduction) is dependent on the water levels observed during the monitoring period [8, 29]. In some cases, for instance at locations with a large tidal range, frequently occurring situations are similar to design conditions, resulting in significant uncertainty reduction[1]. At other locations, such as river dikes, conditions leading to large uncertainty reduction occur less often. Consequently, the longer the monitoring period, the higher the probability of obtaining useful information that can be used to reduce uncertainty, as was shown in Frangopol et al. [116], and will also be demonstrated in Chapter 5.

Proof loading involves imposing a design load in order to prove the resistance of a structure. Since a rise of the phreatic line in the dike is one of the main factors causing instability (as it leads to higher pore pressures and hence lower effective stress), it is considered to artificially impose a high phreatic line by means of infiltration on the dike crest as proof load. The observed performance, being survival information (i.e. a stable dike under the imposed loading condition), is used to improve the reliability estimate, using Bayesian Updating [see 117–119]. Note that proof loading only reduces uncertainty in the (variables relating to the) overall resistance, conditional on the imposed proof load. It typically does not lead to additional knowledge about the actual response of the phreatic line to flood conditions. Thus, pore pressure monitoring and proof loading are complementary.

The decision whether, where, and which type of uncertainty reduction method to use is typically difficult for decision makers, as this can vary strongly per location, and this also depends on the context of their decision [120, 121] (e.g., within what time a dike reinforcement has to be carried out). It was identified by Klerk et al. [122] that a short time horizon until a dike reinforcement is often unfavourable for pore pressure monitoring as the amount of information obtained depends on the

[1] It should be noted that although the pore pressures itself vary in time due to variations in water levels, the uncertain parameters governing the translation of water level to pore pressures are typically time-invariant.

duration of monitoring, and the probability of not obtaining useful information is relatively large. In such cases, proof loading or additional site investigation might be more promising, as the information is time-independent for these methods.

The aim of the first case study in Section 4.2 is to answer the question under what conditions to invest in uncertainty reduction for a dike section (proof loading and/or pore pressure measurements), and which strategy (combination of proof loading and pore pressure monitoring) yields the highest Value of Information (VoI). This is illustrated by a case study of a typical river dike section that currently does not meet the required reliability for the failure due to slope instability. All uncertainty reduction efforts have to be carried out before the reinforcement project will start (a period of 5 years is assumed here). In the next section we will further outline the framework used for evaluating such sequential decisions for uncertainty reduction.

4.2. Case study 1: Pre-posterior analysis of proof loading and monitoring for reinforcement of a dike section

4.2.1. Calculating slope reliability

In this case study we will focus on a flood defence that is sensitive to slope instability. Generally, the safety against slope instability is assessed using limit equilibrium methods, (e.g. Bishop, Spencer, Uplift-Van), which calculate the factor of safety against instability (F_S) considering driving forces (e.g. weight) and resisting forces (e.g. shear stress) acting on a slip plane. The probability of failure is defined as $P(F) = P(g < 0)$, where F is the failure event of instability, and g the performance function. It holds that $g = F_S(\mathbf{X}) - 1$ with F_S the factor of safety against instability with input variables $\mathbf{X}$, being the soil parameters, (hydraulic) loads and model uncertainty. For convenience, it is written $P(F) = P(g(\mathbf{X}) < 0)$.

This section considers a dike which has insufficient safety against instability, i.e. the failure probability ($P(F)$) is larger than the (risk-based [see e.g. 7]) economically optimal target failure probability (P_T): $P(F) > P_T$. Or, in terms of reliability index: $\beta < \beta_T$, where $\beta = -\Phi^{-1}(P(F))$ and Φ^{-1} is the inverse standard normal cumulative distribution function. This section considers only slope stability reliability in the flood risk analysis, as it is the most prominent failure mode. Although other failure modes also contribute to flood risk depending on the local conditions, here it is assumed these contributions are minor. Besides, extension of the approach with more failure modes is straightforward and especially useful if the reliability of other failure modes is also mainly determined by reducible uncertainties.

Here, fragility curves are used to calculate the slope reliability of dikes, [20]. Fragility curves describe the conditional failure probability given a (load) variable. Here the failure probability conditional to the water level h: $P(F|h) = P(g(\mathbf{X}) < 0|h)$ is considered, where $\mathbf{X}$ is the vector of all random variables except h. The annual probability of failure is obtained by combining $P(F|h)$ with the PDF of annual

maxima of the load h as follows:

$$P(F) = \int P(F|h)f(h)dh. \tag{4.1}$$

Note that fragility curves can in principle be made for any (load) variable. This will be used in Section 4.2.3.

4.2.2. Formulation of the decision problem

Insufficient safety against slope instability is typically remedied by decreasing the slope angle or by constructing a stability berm at the inner toe of the dike to increase the resisting weight at the passive side of the slip plane. When space is available, these measures are relatively cheap as the construction and material costs of soil are low. However, when space is scarce (e.g. in densely built areas), reinforcement can become extremely expensive, for example because adjacent home owners have to be moved and compensated, or because other design options are applied such as expensive sheet pile walls and diaphragm walls (see for examples Chapter 3). In such cases, methods for reducing uncertainty that might result in lower reinforcement costs can be very valuable as the required reliability can be achieved at much lower cost. For example, a less costly reinforcement method becomes feasible, or reinforcement projects become obviated.

To evaluate the benefits of pore pressure monitoring and proof loading Bayesian pre-posterior analysis is used. The basic idea of pre-posterior analysis is that, based on a priori available information, one can determine the best decision based on an evaluation of all possible outcomes. Decision trees are the most common approach to visualise and structure pre-posterior decision analysis [e.g. 22, 47, 123, 124]. A decision tree shows a sequence of decision (choice) nodes and outcomes (chance). For more details see Section 2.7.

A disadvantage of a decision tree is that it can become cumbersome to visualise and solve if many sequential decisions are considered, in such cases other approaches such as influence diagrams (i.e., an extension of Bayesian networks) are more adequate [26], possibly combined with heuristic decision rules. This case study considers three decision options (proof loading, monitoring and dike reinforcement), hence a decision tree is well suited. Figure 4.1 presents the decision tree for the sequential decision strategy of proof loading, pore pressure monitoring and dike reinforcement, denoted with p, m, and a, respectively. Note that a specific sequence for proof loading and pore pressure monitoring is assumed, the effect of reversing this is discussed in Chapter 5.

In the evaluation of choices on proof loading and pore pressure monitoring it is desired to evaluate what is the optimal strategy, given the prior belief $f_{\mathbf{X}}(\mathbf{x})$ of the random variables $\mathbf{X}$. Here $f_{\mathbf{X}}(\mathbf{x})$ is the joint probability density, where $\mathbf{x}$ is the realization of $\mathbf{X}$. The failure probability is then given by $P(F) = \int_{g(\mathbf{x})<0} f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$, the integral over the prior belief for all values where the limit state function evaluates to a value smaller than 0.

The first decision considered is whether to execute a proof load test of a certain magnitude $p \in \mathbf{P}$. The outcome z_{p} (a survived or failed proof load test) depends

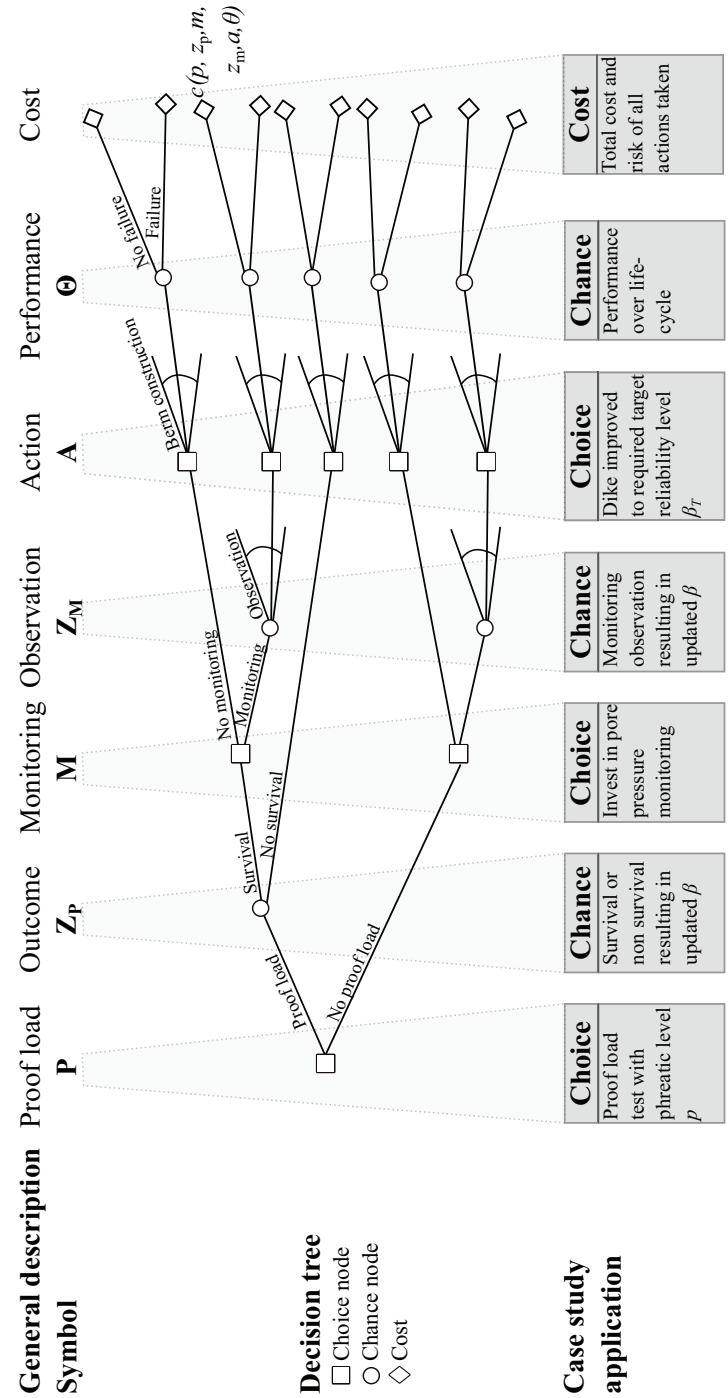


Figure 4.1.: *Decision tree for a sequential decision on proof loading, monitoring, and reinforcement of a dike section. The decision tree is a graphical presentation of the choices $p \in \mathbf{P}$, $m \in \mathbf{M}$, and $a \in \mathbf{A}$, and chances $z_p \in \mathbf{Z}_p$, and $z_m \in \mathbf{Z}_m$.*

on the magnitude of the proof load and the prior belief $f_{\mathbf{X}}(\mathbf{x})$. The higher the magnitude of the survived test load (i.e. the artificially induced phreatic line), the larger the uncertainty reduction, and the higher the updated reliability. On the other hand, the higher the magnitude of the survived test load, the higher the probability that the test is not survived. In that case the dike is damaged and needs to be reinforced immediately and the part of the section that was proof loaded has to be repaired such that extra costs are incurred.

After deciding whether to do a proof load test (and on the magnitude of the test load), it can be decided to invest in pore pressure monitoring ($m \in \mathbf{M}$) in order to reduce uncertainty on the response of the phreatic line to outside water levels. Again, two outcomes are possible: either an observation is made or not. The observation z_{m} depends on the belief after proof loading $f_{\mathbf{X}|z_{\mathrm{p}}}(\mathbf{x})$. Whether an observation is made in the considered time period depends on whether the water level exceeds a certain threshold required to obtain useful measurements [29, 116]. Note that this is time-dependent: the longer the monitoring period, the higher the probability of a useful observation, opposite to a proof load test which is time-independent.

After the outcome of the monitoring, the dike is improved to the required target reliability level. This is done using decision rules that translate combinations of outcomes of proof loading and monitoring to actions: $\mathbf{d}(\mathbf{Z}_{\mathrm{p}}, \mathbf{Z}_{\mathrm{m}}) = \mathbf{A}$, where for an individual decision rule d it holds that $d \in \mathbf{d}(\mathbf{Z}_{\mathrm{p}}, \mathbf{Z}_{\mathrm{m}}) = \mathbf{A}$. Note that through an action $a \in \mathbf{A}$ also some variables in $\mathbf{X}$ can be adapted (e.g., the length of the stability berm).

The cost of a branch in the decision tree is determined by the costs of every individual step and the expected damage given the performance θ (failure/no failure): $c(p, z_{\mathrm{p}}, m, z_{\mathrm{m}}, d, \theta)$. The cost of the optimal strategy $c(p^*, m^*, d^*)$ can be computed by combining the cost of different branches over the possible outcomes:

$$c(p^*, m^*, d^*) = \min_{p \in \mathbf{P}, m \in \mathbf{M}, d} E_{\Theta|\mathbf{X}}\left[E_{z_{\mathrm{m}}|\mathbf{X}}\left[E_{z_{\mathrm{p}}|\mathbf{X}}\left[c\left(p, z_{\mathrm{p}}, m, z_{\mathrm{m}}, d, \theta\right)\right]\right]\right]. \tag{4.2}$$

Specifically in this chapter, the cost of a strategy $c(p, m, d)$ is defined by the sum of costs of each step in the decision tree (decision and outcome):

$$\begin{aligned} c(p, m, d) = {} & I_{\mathrm{p}} \cdot C_{\mathrm{proofload}} + I_{\mathrm{m}} \cdot C_{\mathrm{monitoring}} + \\ & \int_{f_{\mathbf{X}}(\mathbf{x})} P(\bar{S}|p \cap \mathbf{x}) \cdot C_{\mathrm{repair}} + C_{\mathrm{reinforcement}}(d, \mathbf{x}) + C_{\mathrm{failure}}(d, \mathbf{x}) \mathrm{d}\mathbf{x}, \end{aligned} \tag{4.3}$$

where I_{p} and I_{m} are indicator random variables (value 0 or 1) that indicate whether proof loading or monitoring is done. C parameters indicate different cost components. The cost components of proof loading and monitoring are independent of the prior belief $f_{\mathbf{X}}(\mathbf{x})$. There are three cost components that depend on the prior belief: first of all, the cost of failure after a failed proof load test, for which $P(\bar{S}|p \cap \mathbf{x})$ is the probability of not surviving a proof load with magnitude p, and C_{repair} are the repair costs. Secondly, the costs of reinforcement $C_{\mathrm{reinforcement}}$, which depend

on decision d and the realization $\mathbf{x}$. Finally, as the annual failure probability is assumed to be constant in time, for the Present Value of the failure costs $C_{\text{failure}}(d, \mathbf{x})$ an infinite time horizon can be considered, such that:

$$C_{\text{failure}}(d, \mathbf{x}) = \frac{P(F|d, \mathbf{x}) \cdot D}{r}. \tag{4.4}$$

Here $C_{\text{failure}}(d, \mathbf{x})$ is the cost of failure in € for an infinite time horizon, D is the expected damage in case of a flood (in €), r is the annual discount rate, and $P(F|d, \mathbf{x})$ is the annual failure probability given an action following from decision rule d and a realization of the set of random variables $\mathbf{x}$. A reference period of 1 year is assumed, in line with common practice for flood defence structures. It should be noted that in some cases for geotechnical structures use of other reference periods might be more adequate [59], and for instance the time factors provided in Diamantidis et al. [120] may be applied (see also Section 2.5). The cost of the reference strategy without monitoring and proof loading is defined as c_0. The VoI of a strategy p, m, d can be computed by:

$$\text{VoI} = c_0(d) - c(p, m, d) \tag{4.5}$$

The next subsections go further into the choices the decision maker is faced with (summarized in Table 4.1) in more detail.

Table 4.1.: *Methods, goals and activities considered in this chapter.*

Method	Goal	Activity
P: Proof load test	Reduce uncertainty in geotechnical parameters	Artificially raise the phreatic line by infiltrating water in the dike
M: Monitoring	Reduce uncertainty in response of the phreatic line to floods	Monitor the response of pore water pressures during floods using piezometers
A: Dike reinforcement	Increase the reliability of the dike	Increase stability by construction of a stability berm

Step 1: doing a proof load test (P)

Proof loading involves imposing a representative design load on the dike body, for example a high phreatic line (see Figure 4.4). If such a proof load is survived, it proves that there is a minimum resistance along a slip plane. Conversely, when the dike fails under the conditions of the proof load test, it reveals that the structure was not safe enough. Note that a higher proof load yields more information, but also results in a higher risk of failure during the test. The outcome of the proof load test is used to update the failure probability based on the outcome z_{p} of the proof load test, and hence the updated probability of failure is written as:

$$P(F|z_{\text{p}}) = \int_{g(\mathbf{x})<0} f_{\mathbf{x}|z_{\text{p}}}(\mathbf{x}) d\mathbf{x}, \tag{4.6}$$

with:

$$f_{\mathbf{X}|z_\mathrm{p}}(\mathbf{x}) = \frac{P(\mathbf{x} \cap z_\mathrm{p})}{P(z_\mathrm{p})} f_{\mathbf{X}}(\mathbf{x}) = \frac{P(z_\mathrm{p}|\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{\int_{\mathbf{X}} P(z_\mathrm{p}|\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}} \tag{4.7}$$

Here, z_p is the observation of no instability at a proof load level p, for which the performance function for stability at a proof load level p, $g(\mathbf{X}, p) \geq 0$. Instead of updating the probability density $f_{\mathbf{X}|z_\mathrm{p}}(\mathbf{x})$, the updated failure probability is directly calculated by applying Bayes' rule:

$$P(F|z_\mathrm{p}) = \frac{P(F \cap z_\mathrm{p})}{P(z_\mathrm{p})} = \frac{P(g(\mathbf{X}) < 0 \cap g(\mathbf{X}, p) \geq 0)}{P(g(\mathbf{X}, p) \geq 0)} \tag{4.8}$$

This formulation in terms of conditional probability avoids the explicit calculation of the updated joint probability distribution $f_{\mathbf{X}|z_\mathrm{p}}(\mathbf{x})$. Note that a proof load test does not update all parameters, for example for those related to response of the phreatic line to an extreme flood water level (e.g., the head level in the aquifer below the soft soil blanket or pore pressures in the dike body in flood conditions) no additional information is obtained.

The proof load test considered in this section consists of a controlled experiment to artificially raise the phreatic line, assuming a successful test in the sense that it always succeeds in increasing the water pressures to the desired level, throughout the dike body. The cost of such a proof load test involve the set-up of a test, equipment, analysis, and monitoring to substantiate the observations of a survived proof load, such as deformation monitoring to indicate that a rotational shear failure was not initiated under the observed loading conditions [e.g. 125]. Subjecting a structure to a proof load also involves the possibility that instability occurs during the test, with additional repair costs involved.

Step 2: setting up a pore pressure monitoring campaign (M)

After or instead of proof loading, uncertainty can be reduced by setting up a pore pressure monitoring campaign. Pore pressure monitoring aims to reduce uncertainty about the response of the phreatic line in the dike. The parameters characterizing this response are part of the belief $f_{\mathbf{X}|z_\mathrm{p}}(\mathbf{x})$, where conditioning on z_p is not needed if proof loading was not done beforehand. If a pore pressure monitoring campaign yields an observation z_m, $f_{\mathbf{X}|z_\mathrm{p}}(\mathbf{x})$ can be updated to a posterior estimate including z_m:

$$f_{\mathbf{X}|z_\mathrm{p} \cap z_\mathrm{m}}(\mathbf{x}) = \frac{P(\mathbf{x} \cap z_\mathrm{m}) | f_{\mathbf{X}|z_\mathrm{p}}(\mathbf{x})}{P(z_\mathrm{m})} = \frac{P(z_\mathrm{m}|\mathbf{x}) | f_{\mathbf{X}|z_\mathrm{p}}(\mathbf{x})}{\int_{\mathbf{X}} P(z_\mathrm{m}|\mathbf{x}) f_{\mathbf{X}|z_\mathrm{p}}(\mathbf{x}) d\mathbf{x}}. \tag{4.9}$$

Note that the likelihood $P(z_\mathrm{m}|\mathbf{x})$ (sometimes referred to as $L(\mathbf{x})$) is calculated with the updated probability distribution $f_{\mathbf{X}|z_\mathrm{p}}(\mathbf{x})$. Then it holds for the posterior probability of failure with monitoring and proof load:

$$P(F|z_{\mathrm{m}} \cap z_{\mathrm{p}}) = \int_{g(\mathbf{x}<0)} f_{\mathbf{X}|z_{\mathrm{p}} \cap z_{\mathrm{m}}}(\mathbf{x}) d\mathbf{x}. \tag{4.10}$$

Note that the parameters in $\mathbf{X}$ related to the response of the phreatic line are now directly updated, as there are direct observations of input parameters, contrary to proof loading.

An important parameter for pore pressure monitoring is the probability that a useful observation is obtained. Often discontinuities in a dike body (e.g., an older clay dike), or different permeability values in general can result in different responses of the phreatic line for different outside water levels, and therefore, an observation z_{m} is to give more useful information if measurement conditions are closer to design conditions. To incorporate this, it is assumed that a valuable measurement (i.e. uncertainty reduction) is only obtained if the annual maximum water level h exceeds a predefined threshold water level h_{thresh}. Thus the probability of obtaining a valuable measurement z_{m} can be computed using the following formula:

$$P(z_{\mathrm{m}}) = 1 - F(h < h_{\mathrm{thresh}})^t \tag{4.11}$$

where $F(h < h_{\mathrm{thresh}})$ is the cumulative probability per year that the outside water level does not exceed the threshold water level h_{thresh}, and t is the duration of monitoring in years.

Step 3: dike reinforcement (A)

In practice, numerous reinforcement methods are available to increase the stability of dikes, for example: stability berms, sheet pile or diaphragm walls, or soil anchoring techniques. For this case, only the most common (and often cheapest) method of stability berm construction is considered. Adding a stability berm at the inner toe of the dike increases the weight on the passive side of the slip plane and increases the resisting shear stress.

The target reliability that has to be satisfied after a dike reinforcement is often predetermined, and typically based on an optimization of various risk indicators and costs of reinforcement [e.g. 7, 16, 17]. If the reliability of dikes is changing significantly in time, one also has to consider reinvestments. However, due to the dependence of slope stability reliability on time-invariant ground-related uncertainty, slope stability of dikes (and other geotechnical structures, [e.g. 59]) is typically rather time-invariant. Therefore, in an economic optimization one can estimate the annual target reliability by considering an infinite time horizon, such that the optimal level of protection β_{T}, follows from the following minimization:

$$\beta_{\mathrm{T}} = \arg\min_{\beta} \left(C(\beta) + \frac{\Phi^{-1}(\beta) \cdot D}{r} \right) \tag{4.12}$$

where D is the annual expected damage in case of flooding, r is the annual discount rate and $C(\beta)$ is the cost of achieving a certain reliability index. It has to be noted that in practice reliability targets are typically specified in standards and are

not based on a case-specific optimization. In Chapter 3 we already saw an example of how this can lead to suboptimal choices on a cross-sectional level. For this case study this will be further addressed in the sensitivity analysis in Section 4.2.5.

4.2.3. Case description

The reference case is a dike section of 1 km in length, inspired by an actual dike section currently being reinforced. It is slightly simplified such that it contemplates a typical dike section in the Dutch riverine area. The dike cross section, displayed in Figure 4.2, consists of a traditional clay dike which has been reinforced with sand in the past. It is assumed that the dike is scheduled for reinforcement in 5 years as it currently does not meet the safety standard. Until that time there is opportunity to do a proof load test and pore pressure monitoring to reduce uncertainty on the resistance parameters and the position of the phreatic line in the dike body, respectively. The goal is to determine the optimal course of action for the coming 5 years.

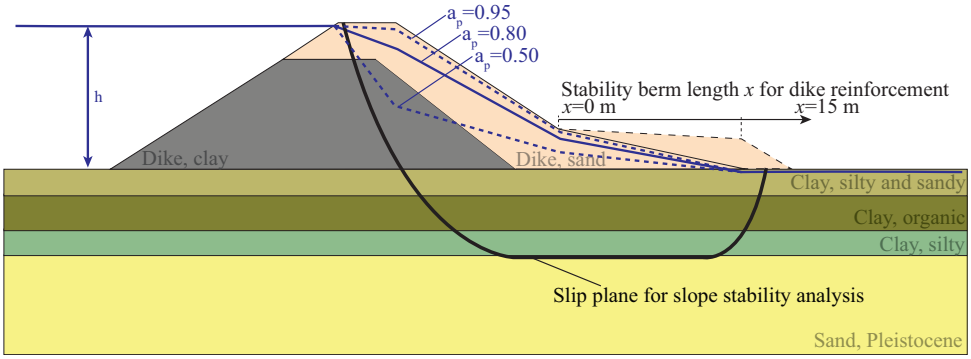


Figure 4.2.: *Cross section of the considered case study. Blue lines indicate the simplified schematization of the phreatic line for different response factors a_p (at an extreme water level).*

The dike consists partly of clay and partly of sand, has a crest level at 14.0 m+ref. (reference level), a landside elevation of 6.0 m+ref., an inner slope of 1:3 (v:h) and is situated on (Holocene) clay layers on top of a (Pleistocene) aquifer. A cross-section of the considered dike is shown in Figure 4.2. The strength of the soil is modelled according to the Critical State Soil Mechanics framework [126] with a critical state friction angle (ϕ) or undrained shear strength (s_u) calculated using the SHANSEP formulation [127], see Equation 4.13.

$$s_u = \sigma'_v \cdot S \cdot OCR^m \tag{4.13}$$

Here, S is the undrained strength ratio for normally consolidated soil, m is the strength increase exponent, and OCR is the over-consolidation ratio: the ratio of in situ vertical effective stress (σ'_v) and pre-consolidation stress $\sigma'_p = \sigma'_v + POP$, where POP = pre-overburden pressure. The vertical effective stress is the total vertical

soil stress σ_v minus the pore pressure σ_p. Note that a high phreatic line leads to higher pore pressures, thus lower s_u, and lower stability (F_s). Additionally, the stability decreases because a higher phreatic line corresponds to a higher weight of the dike body.

In the case study, only monitoring of the phreatic line in the dike body is considered, not of the pore water pressures in other soil layers. The position of the phreatic line in the dike at flood conditions typically depends on the permeability of the dike material which is often heterogeneous and uncertain. Especially when a dike has a long history of reinforcements with various materials, the phreatic line is uncertain. For example a traditional clay dike reinforced with sand as considered in this case. Therefore, the position of the phreatic line in steady state seepage conditions is parametrised, using an uncertain response factor (a_p). The response factor represents the degree of saturation of the dike body at the inner crest line, in response to an extreme water level. Values can range between $a_p = 0$ (phreatic level at the landside elevation level) and $a_p = 1$ (phreatic level equal to the outside water level). For intermediate values of a_p, the phreatic line is interpolated accordingly, see Figure 4.2. Because the dike body will always saturate to some degree, and in case of a fully saturated dike ($a_p = 1$) other mechanisms such as micro-instability become dominant, the value of $a_p$is limited between 0.5 and 0.95. The bounds represent realistic values based on physical considerations. Furthermore, the lower bound has a limited influence on the reliability, indicated by the results in the next paragraph.

To facilitate the probability updating outlined in Section 4.2.2, 3-dimensional fragility surfaces are derived, where the failure probability is conditional to response factor a_p and the water level h. These surfaces are derived both for the prior situation, and the situation posterior to surviving a certain proof load level p using the approach outlined in Schweckendiek et al. [20]. Figure 4.3 shows this fragility surface, plotted in terms of reliability index for convenience. The reliability is calculated at discrete intervals of h and a_p, and linearly interpolated to obtain intermediate values. The fragility surface directly shows the influence of the response factor a_p(mainly at high water levels), and clearly illustrates the potential benefit of reducing uncertainty herein. Separate fragility surfaces $\beta(h, a_p)$ are derived for berm lengths of 5, 10, 15, and 20 m. For other values fragility surfaces are interpolated or extrapolated. Table 4.2 lists the input probability distributions for parameters in the reference case. The probability distributions for these spatially averaged soil parameters are derived from regional data for typical geological deposits of the Dutch situation, see Rijkswaterstaat [128]. Integration of the fragility surface with the prior probability distribution of $a_p$and h along the lines of Equation (4.1), results in a prior failure probability of $2.7e-4$ ($\beta = 3.46$).

4.2.4. Implementation of decision problem

Proof loading

Proof loading is done by artificially raising the phreatic line in the dike by infiltrating water into the dike from the crest (similar to van Hoven & Noordam [129]), see

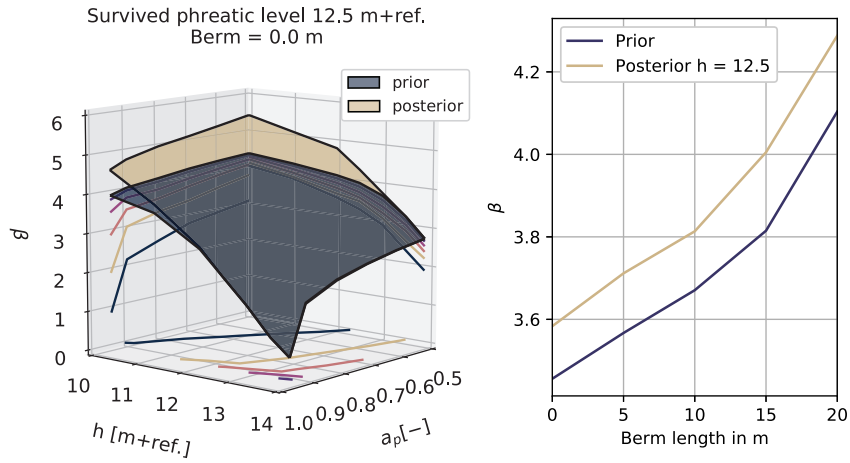


Figure 4.3.: *(left) Prior and posterior fragility surface (in terms of reliability index β) of the water level h and response factor a_p, for the considered case study without berm. The overall reliability index (integrated with the prior probability density of a_p and h) is 3.46. (right) Relationship between berm length and overall reliability index β for the prior situation and posterior after a proof load level of 12.5 m+ref.*

Figure 4.4. Survival of the situation with an imposed phreatic level leads to a higher reliability because of an implicit update of the probability density of soil parameters involved (which are a subset of $\mathbf{X}$). The higher the phreatic level, the larger the uncertainty reduction, and hence, the larger the reliability update; but also the higher the probability the proof load is not survived.

Contrary to the phreatic level in flood conditions (dependent on amongst others the flood water level, the duration of the flood wave, and permeability of the outer slope cover layer), the phreatic level during proof loading is induced/imposed by infiltrating water into the dike using e.g. infiltration wells or an irrigation system (see van Hoven & Noordam [129] for pictures). Therefore, the outcome of the proof load test (and hence the updated reliability) is independent of the response factor a_p, and the posterior reliability β conditional to a_p: $\beta|a_p = -\Phi^{-1}(P(F|z_p, a_p))$ can be computed, using the formulation from Equation 4.6 to calculate $P(F|z_p, a_p)$.

Figure 4.3 shows that a significantly updated reliability for water levels lower than the survived proof load of 12.5 m +ref. is to be expected. The reliability update is relatively larger for lower values of a_p. This is in line with expectations because the survived proof load becomes more valuable if a high phreatic line is less likely. Note that the failure probability for water levels lower than the survived proof load level is not reduced to 0 (infinite beta) because of irreducible uncertainty (see Schweckendiek et al. [20] for a consideration of reducible and irreducible uncertainty). In this case this mostly concerns uncertainty in time-variant parameters, such as the rainfall intensity.

It is assumed that the proof load is applied over a stretch of 100 m length. This is considered representative for the 1 km dike section because of a limited variation

Table 4.2.: *Random variables in the reference case.*

Property	Symbol	Unit	Soil type	Distribution[1]
Normally consolidated undrained shear strength ratio	S	-	Clay, silt and sand	Lognormal(μ=0.36, CoV=0.15)
			Clay, organic	Lognormal(μ=0.29, CoV=0.15)
			Clay, silt	Lognormal(μ=0.32, CoV=0.25)
Strength increase exponent	m	-	Clay, silt and sand	Lognormal(μ=0.84, CoV=0.05)
			Clay, organic	Lognormal(μ=0.93, CoV=0.05)
			Clay, silt	Lognormal(μ=0.83, CoV=0.05)
Pre-overburden Pressure at daily stress conditions (no flood)	POP	kPa	Clay, silt and sand	Lognormal(μ=27.0, CoV=0.45)
			Clay, organic	Lognormal(μ=27.0, CoV=0.45)
			Clay, silt	Lognormal(μ=27.0, CoV=0.45)
Critical state friction angle	ϕ_{cs}	°	Dike, sand	Lognormal(μ=32.6, CoV=0.05)
			Dike, clay	Lognormal(μ=35.0, CoV=0.05)
			Clay, silt and sand	Lognormal(μ=32.3, CoV=0.05)
			Sand, Pleistocene	Lognormal(μ=35.0, CoV=0.05)
Model factor stability model	d	-		Lognormal(μ=0.995, CoV=0.033)
Parameter for phreatic line	a_{p}	-		Uniform(a=0.5, b=0.95)
Water level	h	m+ref.		Gumbel(loc=11.9, scale=0.2)

[1] Note that μ is the mean value, not the lognormal distribution parameter, CoV is the Coefficient of Variation.

in the dike body in longitudinal direction as a result of the quite recent reconstruction with sand, see Figure 4.2. The total cost of a proof load test is assumed to be $500000\,€$ consisting of costs for equipment required for infiltration, monitoring during the test, emergency measures to mitigate slope failures induced by the test and analysis of the test results. It is assumed that the test is carried out in a period where a potential failure does not cause flooding. Therefore, the costs of not surviving a proof load only consist of repairing the damaged slope. These costs are estimated to be $2000000\,€$, based on the costs of full reconstruction of the existing dike over a length of $100\,\mathrm{m}$. Additional costs such as follow-up damage to buildings, transportation infrastructure, agricultural areas etc. are disregarded in this case study. For damage during a proof load test (C_{repair}) occurrence of each slip circle (also very shallow) is considered as failure, contrary to flooding. For flooding damage (D) only larger slip circles which will lead to flooding of the hinterland are considered, as is depicted in Figure 4.4. After a proof load test failure, no pore pressure monitoring is done.

Pore pressure monitoring

Pore pressure monitoring is carried out by measuring the phreatic line in the dike body (see Figure 4.4 for location of sensors). The measurement will lead to an update of the probability distribution of a_{p} ($a_{\mathrm{p}} \in \mathbf{X}$). Because of the chosen limits of the prior distribution of a_{p}, it is assumed that the posterior distribution of a_{p} is a truncated normal distribution with μ the observed value (i.e., based on possible state), standard deviation $\sigma = 0.05$ and upper and lower bound equal to the upper

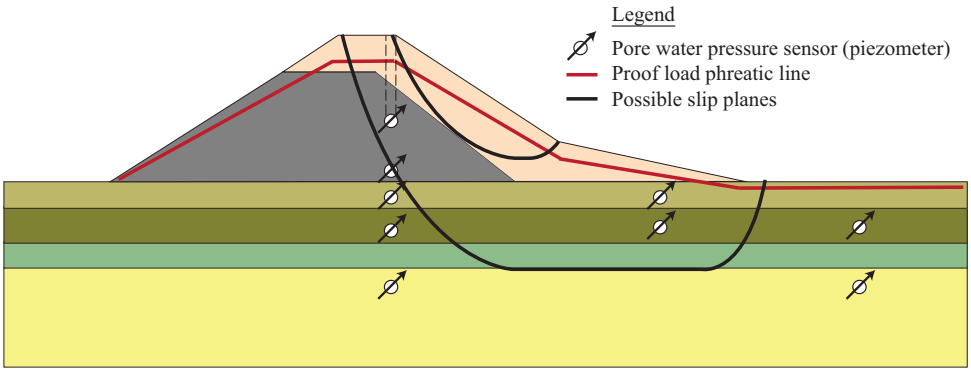


Figure 4.4.: *Overview of the positioning of sensors installed for pore water pressure monitoring, and the imposed phreatic level during a proof load test. The larger black line indicates the slip plane relevant for flooding, the smaller slip plane is relevant for failure of the proof load test but does not cause flooding.*

and lower bound of the uniform prior. The value of σ accounts for measurement errors and transformation errors, and corresponds with a standard deviation of 0.3 m in the position of the phreatic line. This value is in accordance with commonly found values in the Dutch practice [130].

Due to the old clay dike located in the cross section the sensors will only yield relevant results if the water level is somewhat above the crest of the old clay dike (see Figure 4.2, it is assumed that this threshold is 12.2 m +ref. (0.2 m above the top of the clay). Given the local probability distribution for water levels and 5 years of monitoring, the probability that a relevant observation is obtained is 67% (using Equation (4.11)).

While not explicitly modelled, the costs are based on plans for measuring the entire section including redundancy in measurements and multiple cross-sections with sensors. The cost of pore pressure monitoring is estimated at 100000 € for 5 years and include cost for installation, maintenance, decommissioning and analysis of the obtained data, based on the number of sensors in Figure 4.4, installed at two cross-sections.

Dike reinforcement

The reliability requirement for the dike section is determined based on the level of protection with minimal total cost (see Equation (4.12)). This value is derived based on the prior $f_{\mathbf{X}}(\mathbf{x})$. The costs for reinforcement are shown in Figure 4.5, both for the reference case and some alternatives that will be used in a sensitivity analysis. Except for alternative 2, these curves have been derived using KOSWAT, a software program used for cost calculations for dike reinforcements in the Netherlands [111]. Only reinforcement through a stability berm is considered (see Figure 4.2 for dimensions). The costs are calculated using Equation (4.3). Note that the risk in the 5 years before reinforcement is not considered, as this is the same for each

Table 4.3.: *Parameters for cost, risk and monitoring.*

Parameter	Description	Unit	Value
r	Annual discount rate	-	0.035
D	Damage in case of flooding	million €	5000
C_{repair}	Cost of repair after failed proof load test	million €	2.0
$C_{\text{monitoring}}$	Cost of 5 years of pore pressure monitoring	million €	0.1
$C_{\text{proofload}}$	Cost of proof load test	million €	0.5
σ	Uncertainty in observation of a_{p}	-	0.05
h_{thresh}	Minimum water level for a useful observation	m +ref	12.2

strategy (and thus does not lead to differences in VoI).

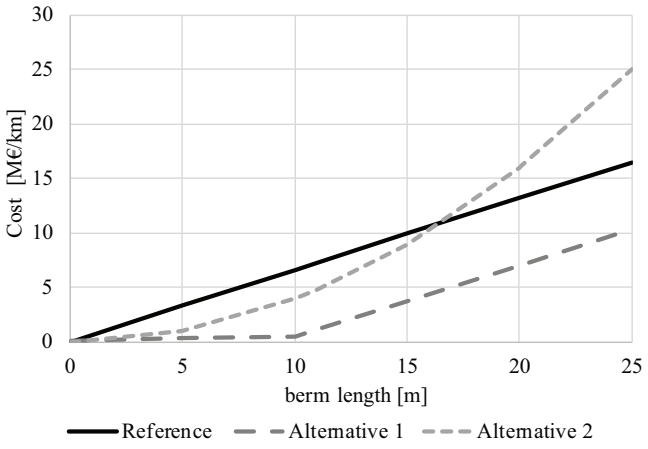


Figure 4.5.: *Different cost functions for dike reinforcements.*

4.2.5. Case study results

Reference case

First, it is evaluated whether proof loading and/or pore pressure monitoring reduces overall total cost for a reference case. Here, a proof load test where the phreatic line is artificially increased to 12.5 m+ref. is considered. The parameters used for the cost benefit analysis are shown in Table 4.3. Figure 4.6 shows Total Cost (TC) and Value of Information (VoI) (see Equation (4.5)) for all combinations of proof loading and monitoring, compared to a conventional strategy without monitoring and proof loading. Both monitoring and proof loading reduce TC, with the optimal strategy being a combination of proof loading and monitoring (VoI ≈4.6 M€). For the optimal strategy the reduction in total cost is approximately 30% compared to a conventional reinforcement, strategies with only proof loading or monitoring have a lower but also positive Value of Information. The most important component for

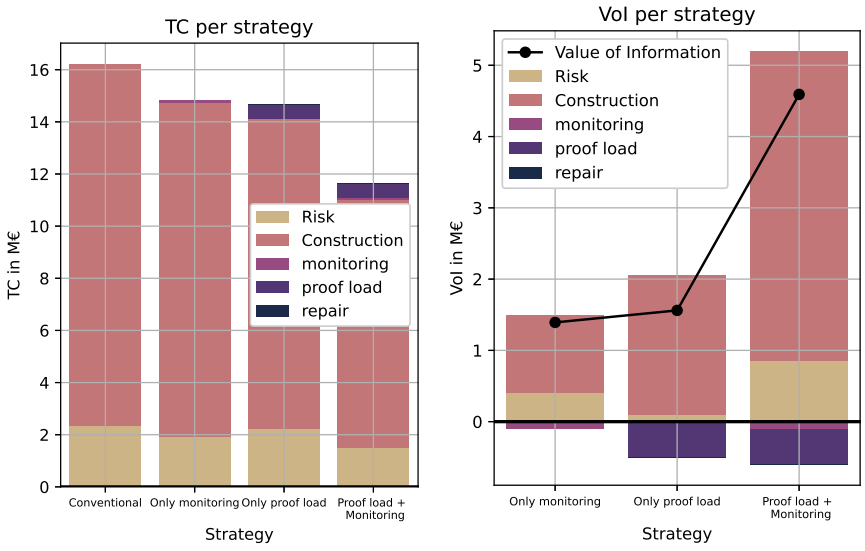


Figure 4.6.: *Total Cost (TC) (left) and Value of Information (VoI) (right) per strategy for the reference case. Colors indicate what the contribution is of different components to the TC (left) and VoI (right). The VoI for each strategy (the sum of the components) is calculated relative to the conventional strategy.*

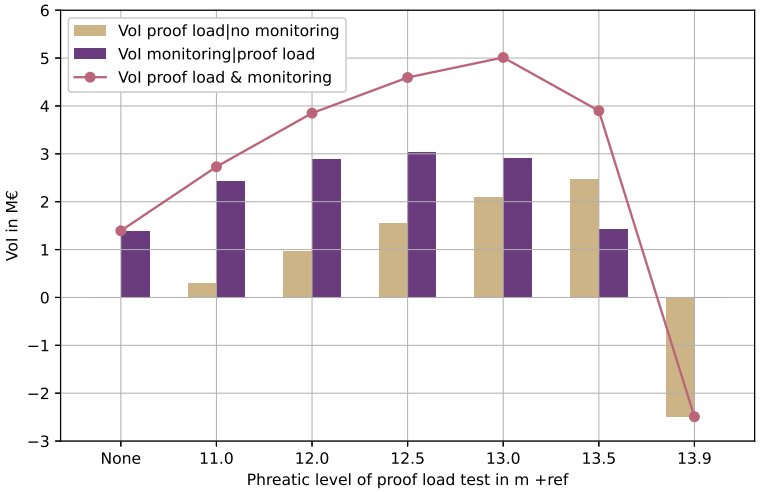


Figure 4.7.: *Value of Information (VoI) related to the level of the proof load test. Red line indicates the VoI for a combination of proof loading and monitoring, yellow bars indicate the value of a proof load test without monitoring. Purple bars indicate the added value of monitoring after a proof load test.*

the VoI is the reduction in construction cost, which significantly outweighs the costs of monitoring and proof loading.

Optimization of proof load level

Although Figure 4.6 clearly shows that a combination of monitoring and proof loading is an effective approach to reduce TC, another important choice is the phreatic level that is to be tested. While lower levels will result in a smaller reduction of uncertainty, higher levels have higher uncertainty reduction but also the added risk that the dike section fails during the test and has to be repaired. Figure 4.7 shows the relation between phreatic level in the proof load test and the VoI. The red line indicates the VoI for different combinations of proof loading and monitoring, for which the optimum is at a proof load phreatic level $13.0\,\mathrm{m}+\mathrm{ref}.$. If no monitoring is done, the optimal proof load level is 13.5 m +ref. (see yellow bars). However, combined with monitoring, the VoI is highest with a lower proof load level (e.g., $12.0\,\mathrm{m}+\mathrm{ref}$). For proof load levels above $13.5\,\mathrm{m}+\mathrm{ref}.$, the VoI becomes negative because of the high risk of failure during the test (i.e., there is a critical proof load level where the VoI = 0). Another interesting observation is that in this case the VoI of monitoring after a proof load test (purple bars) is higher than the VoI without a preceding proof load test (left purple bar). Thus, the monitoring becomes more valuable after reducing uncertainty through proof loading. Obviously, this can differ per case, and it is also dependent on for instance the shape of the relationship between construction cost and berm length. In the next sections we explore the sensitivity of the VoI to three differences often encountered in practice:

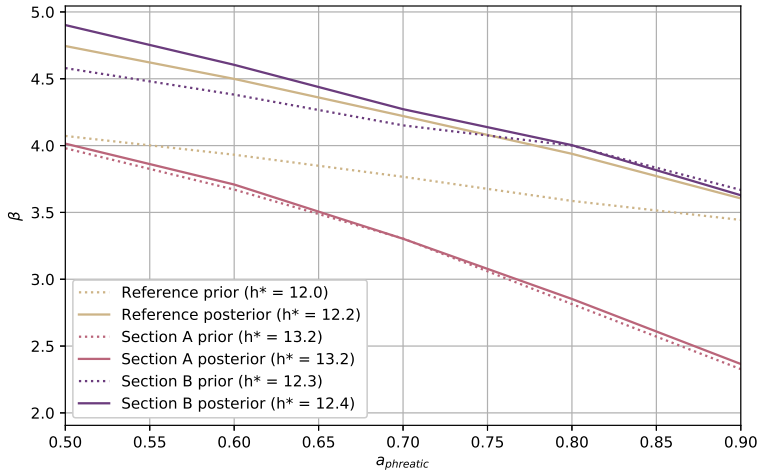


Figure 4.8.: *Fragility curves at the design point water level, showing an increasing dependency for the response factor $\alpha_{\mathrm{p(hreatic)}}$ after proof loading (steeper curve).*

- Influence of the reliability requirement: in many practical cases reliability requirements are not based on an economic optimization, such that the VoI might be different.
- Influence of different soil parameters: different locations can have significantly different mean values and variance of soil parameters, such that he benefits of different types of uncertainty reduction might shift.
- Influence of different cost functions: due to local circumstances (e.g., density of adjacent buildings) costs of reinforcement can vary, which can influence the VoI.

A proof load level of $13.0\,\mathrm{m}+\mathrm{ref}$ is assumed in all cases of the sensitivity analysis, which is (close to) optimal in all cases and strategies (see also Figure 4.7).

Influence of the reliability requirement

In the reference case an optimal target reliability level is determined based on a Total Cost minimization using prior information. In reality, the section studied is part of a larger flood defence system where other safety requirements (e.g., loss-of-life) might be dominant, or requirements are based on general codes. It would therefore be unlikely that the safety standard is exactly economically optimal for this specific dike section, with its specific characteristics. Figure 4.9a and Figure 4.9b show a comparison of Total Cost and Value of Information for 4 cases: the reference case with optimized target reliability based on the prior information ($\beta_{\mathrm{T}} = 4.13$), a case with 10 times higher requirement ($\beta_{\mathrm{T}} = 4.63$), 10 times lower requirement ($\beta_{\mathrm{T}} = 3.56$), and a case where the optimal target reliability is determined based on the posterior information after a proof loading and/or monitoring.

Without monitoring, the cases with lower and higher reliability requirements are significantly more expensive in terms of Total Cost. For the case with a higher requirement this is mainly caused by higher reinforcement costs, whereas for the case with a lower requirement this is due to higher risk costs. As reinforcement costs for the case with a higher requirement are still high after monitoring, the VoI is limited for this case. For the case with a lower requirement, the VoI of a combination of a proof loading and pore pressure monitoring is very high. The reason is that in case of very unfavourable values of a_{p} (and therefore high risk costs), observations are very valuable. In addition, it is prevented that an insufficiently safe dike is constructed as a result of an already too low reliability requirement. The most efficient strategy in terms of Total Cost is if proof loading and monitoring are combined with a posterior optimal reliability requirement. Concretely, the optimal target reliability to be met after the dike reinforcement is determined based on the posterior information after monitoring and/or the proof load test (using Equation (4.12)), rather than the prior information. Consequently, the optimized target reliability depends on the obtained information z_{p} and z_{m}, and the determination of β_{T} becomes a part of the decision rules $\mathbf{d}(\mathbf{Z}_{\mathrm{p}}, \mathbf{Z}_{\mathrm{m}})$ in the decision tree. Hence, each branch in the decision tree can have a different β_{T}, dependent on the observations. This is slightly more efficient than having a requirement based on prior information, especially in case of a very favourable or unfavourable outcome after monitoring, because the change in expected reinforcement cost can be adjusted in the posterior optimization of the requirement. It has to be noted that the differences with the reference case with (prior) optimized β_{T} are limited, but it demonstrates that using a suboptimal target reliability has a large influence on the results of a Value of Information analysis.

Influence of different soil parameters

The reference dike section is characterised by relatively large uncertainties in soil parameters, and therefore the Value of Information of both proof loading and monitoring is found to be relatively large. However, not all dikes might have such large uncertainties, and therefore the VoI is assessed for two other cases: dike section A with lower uncertainty in soil parameters and a prior reliability index of 3.99, with a target reliability of 4.07. Hence, there is only a small reliability deficit, that would in practice likely be accepted as is. Section B also has relatively low uncertainty in soil parameters but lower mean values, so the prior reliability index is 3.61, with a target reliability of 4.02. Figures 4.9c and 4.9d show the Total Cost and VoI for each dike section for 4 different strategies.

Compared to the reference case, section A has considerably lower Total Cost as it is much closer to the target reliability (so the construction costs are much lower). At the same time, the VoI for proof loading is negative, which is due to the fact that the initial reliability is relatively high and the influence of soil parameter uncertainty is limited. Therefore a high proof load has to be applied to learn anything, which results in a higher probability of failure during the test. Thus, for this section proof loading adds very limited value. Although the uncertainty of soil parameters for section B is similar to that of section A, the fact that the initial reliability is lower results in a small but positive VoI for proof loading.

For pore pressure monitoring the VoI is positive in all cases. While the absolute VoI for section A is quite low compared to the other cases, relatively speaking monitoring reduces Total Cost by 22%. One thing that is quite apparent for the reference case is that the relevance of monitoring increases significantly once it is combined with a proof loading, which is not the case for the other cases with lower uncertainty in soil parameters. This can be explained as follows: a priori, the reliability in the reference case is hardly influenced by the response factor a_{p}, whereas, a posteriori, the reliability is dependent on the response factor. This is shown by the less steep fragility curve for the reference case in Figure 4.8. These curves are plotted conditional to the design point (i.e. most probable failure point) of the water level such that it best illustrates the contribution to the failure probability. So the results show that when geotechnical uncertainty is the dominant uncertainty in the prior failure probability (as it is only in the reference case), pore pressure monitoring is less effective than proof loading. After proof loading, geotechnical uncertainty is reduced, and pore pressure monitoring becomes much more effective.

Influence of different cost functions

Local differences in density of buildings, land prices, and available space for reconstruction, can significantly influence the costs for reinforcing dikes using stability berms. The reference dike section is considered for three different cost functions (see Figure 4.5). Figures 4.9e and 4.9f show Total Cost and Value of Information for the three different functions. Alternative 2 has relatively large benefits for proof loading, compared to the reference case (relative to Total Cost). This is caused by the lower marginal cost of the berm in €/m′ after proof loading, due to the fact that part of the cost function is less steep than the reference case.

However, for alternative 2 the benefits of monitoring are much larger if the reliability requirement is optimized based on the posterior information after monitoring, rather than the prior information. Note that the same holds for alternative 1, but results are not shown. The reason is that the marginal costs of reinforcement differ per berm length. Henceforth, if the posterior reliability estimate differs significantly from the prior estimate, the marginal costs of reinforcement might change significantly as well. As a consequence the initial reliability requirement might be suboptimal given the obtained information. Thus, especially if a cost function is highly non-linear, such a posterior optimization of the target reliability might yield significant benefits.

4.2.6. Discussion & conclusions

In this case study we analysed the Value of Information (VoI) of pore pressure monitoring, combined with proof load tests. It was found that both methods for uncertainty reduction have a positive VoI in most, but not all cases, as was demonstrated by the sensitivity analysis. The generic decision tree framework that was used can be easily extended such that also other decisions and sensitivities can be included. In this case for instance different uncertainty reduction methods such as site investigation were not considered, while these also might have a positive VoI.

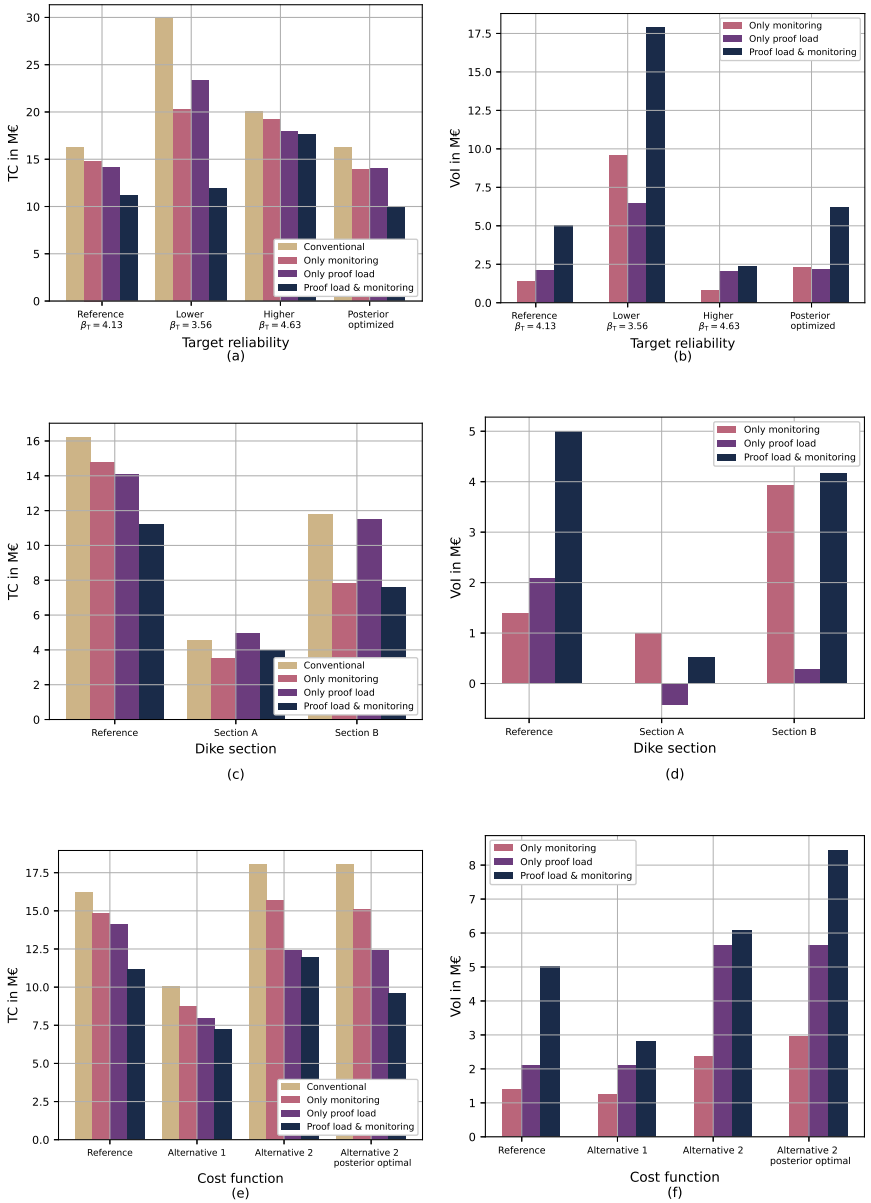


Figure 4.9.: *Total Cost and Value of Information for different target reliability values (panes a and b), for different dike sections (panes c and d) and for different reinforcement cost functions (panes e and f). Proof load test level for all strategies is 13.0 m +ref.. Conventional strategy has no proof load test and no monitoring.*

The same holds for inclusion of other failure modes, non-linear cost functions and optimization based on total system cost rather than reliability-based design (all to be considered in the next case).

It has to be noted that in this case only slope stability failures are considered, whereas in practical situations there are often multiple failure modes that can be of relevance. This will change the VoI for reduction of uncertainties in slope stability, for instance if an increase in crest level is also required to mitigate risks from overtopping failure. However, the presented framework facilitates such a straightforward extension. It is also assumed that failure probabilities in different years are uncorrelated. While this is in line with common practice in flood defence reliability analysis, knowledge uncertainties on soil parameters are typically correlated in time. Consequently the future failure probability might be overestimated in cases with large knowledge uncertainty, most notably the case without uncertainty reduction. However, based upon the findings by Roubos et al. [59] and the examples in Section 2.5 the overall effect is expected to be small here.

In the first case study it is assumed that proof loading is executed first, and after that pore pressure monitoring. However, in practice it might also make sense to alter the sequence of testing, for instance if it is expected that the outcome of pore pressure monitoring is already sufficient to ensure that the target reliability is met, or if the insights obtained from pore pressure monitoring can aid in optimizing the proof load level. This can be incorporated within the framework of sequential decisions.

Practical applications of pore pressure monitoring might or might not concern cross-sections with a threshold, such as the old clay dike in the cross section considered in this study. If there is not such a clear threshold, including monitoring can be done in a similar manner, although additional (less binary) monitoring outcomes might have to be considered. Another point of attention is that in this case a useful observation is obtained at a water level that occurs approximately once per 5 years. There might be situations where useful observations are less (or more) frequent, which obviously has an influence on the VoI of pore pressure monitoring. These considerations will be elaborated further in Chapter 5. The presented framework using a decision tree approach does facilitate adding additional outcomes or changing the threshold level. Additionally it is assumed that findings of monitoring and proof load tests at a single cross section can be translated to the entire dike section. Due to spatial variability in the longitudinal direction this might not always be the case, although such factors can be accounted for in the spacing of pore pressure sensors [131], provided that an adequate amount of site investigation is available.

In conclusion, this case has demonstrated that a decision tree framework with sequential decisions can answer the question under what conditions to invest in different measures to reduce uncertainty for a dike section. It is found that a strategy consisting of a proof load test and/or pore pressure monitoring has a positive VoI. The effectiveness of both methods depends greatly on the specific case. The relative reduction in Total Cost for all cases considered in this section ranges between 11 and 60 % (on average 35%), of which the main contribution is a reduction in construction costs. However, the optimal strategy is not the same in all cases.

Proof loading is most beneficial for cases where the uncertainty in soil properties is dominant and where the initial reliability is relatively low. Obviously the potential benefit must outweigh the additional risk of a failing proof load test and its costs. Pore pressure monitoring is most beneficial for cases where the uncertainty in the phreatic response is dominant.

While this case only considered a reliability-based design for a single dike section, the next case study will consider uncertainty reduction in the context of a Total Cost optimization at the level of a dike segment.

4.3. Case study 2: Reducing uncertainty in a dike segment reinforcement

4.3.1. Formulation of the decision problem at a segment level

The previous case considered uncertainty reduction for a single dike section. It was found that this can yield significant benefits. However, as was demonstrated in the sensitivity analysis, the Value of Information (VoI) depends strongly on the reliability requirement considered in the design. From the perspective of Chapter 3, this means that if a Total Cost optimization for a dike segment of multiple sections is considered, it is likely that the VoI will also be influenced by taking the perspective of a dike segment rather than a single dike section. In this case study we consider a dike segment that consists of 5 sections, each of which is susceptible to overtopping, inner slope instability and piping erosion.

For each dike section susceptible to piping erosion we consider the possibility to do pore pressure monitoring to reduce uncertainty on the permeability of the piping-sensitive aquifer. A decision tree with all possible decisions is given in Figure 4.10.

For monitoring actions $\mathbf{m} \in \mathbf{M}$, we consider the option to execute a pore pressure monitoring campaign to reduce the uncertainty in aquifer permeability (k) for the piping assessment. Note that, contrary to the case in the preceding section we do not consider any benefits of pore pressure monitoring for slope instability. Additionally we assume that the probability of having a useful observation (z_{m}) is 1, and does not depend on the water levels during the monitoring period.

For the actions $\mathbf{a} \in \mathbf{A}$, we consider three possible methods for reinforcing the flood defence, which is to be done in 2025, next to the monitoring:

- Reinforcement with soil consisting of either or both berm widening and an increase in crest level.
- A diaphragm wall that eliminates both piping and inner slope instability (reliability index $\beta > 6$).
- A Vertical Sandtight Geotextile (VSG) [65] that largely eliminates the probability of piping failures. This measure is only available at dike section 4. This is a typical example of an uncertainty substitution measure.

For details on the modelling of these measures see Appendix B. After the reinforcement in 2025, the performance in terms of life-cycle reliability of the dike segment

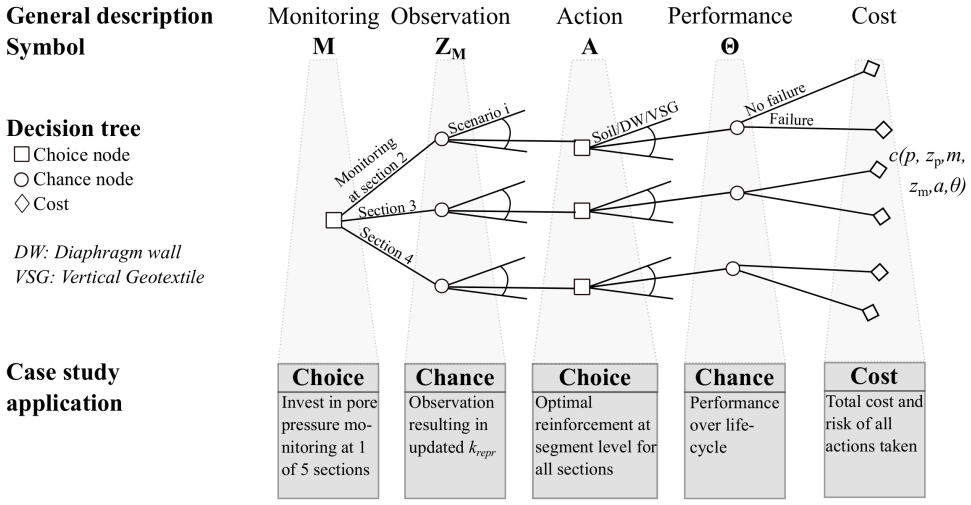


Figure 4.10.: *Decision tree for the second case. Note that monitoring can only be done at 1 section, and that the set of actions* **A** *denotes all reinforcement measures for the entire dike segment.*

is considered for a period of 50 years in the Total Cost (TC) computation.

4.3.2. Case study description

Figure 4.11 shows the reliability indices for each failure mode and each section at year 2025, without reinforcement. It can be seen that several sections have a reliability deficit for piping erosion, and section 5 has a large deficit for inner slope instability. Input values are of the same format as in Chapter 3, underlying assumptions in the modelling of reliability of the different mechanisms and measures are outlined in Appendix B. The target reliability of the dike segment is 1/10,000 years ($\beta = 3.72$), for which we compute the VoI, however we do optimize the entire segment based on Total Cost (which might yield a slightly different resulting reliability index β).

Pore pressure monitoring is included in the analysis by first determining the prior distribution of k and subsequently determining the influence of possible monitoring outcomes on the posterior distribution of k after pore pressure monitoring. The prior distribution is a lognormal distribution of which the 95% representative value (k_{repr}) is known, and a Coefficient of Variation (CoV) of 50% is assumed (default value in the Netherlands). The posterior distribution is governed by a monitored mean k_{i} and a CoV of 10%, caused by measurement inaccuracy of the monitoring method. This can then be translated to a posterior value for k_{repr}. Possible values for k_{i} are discretized into 8 scenarios, weighted by their a priori probability. Figure 4.12 shows the prior and posterior distribution of the representative value of

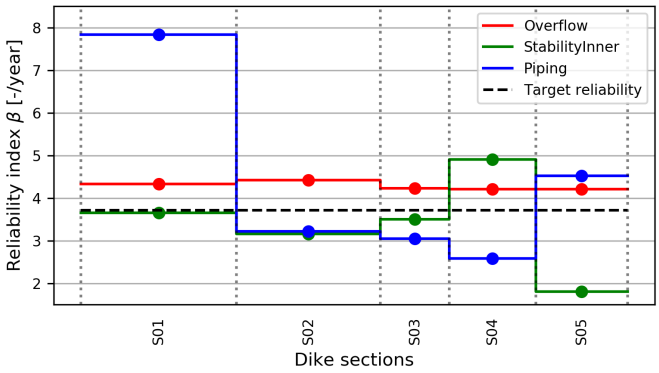


Figure 4.11.: *Reliability indices β for each section per failure mode in 2025. The dashed black line denotes the target reliability for the segment as a whole.*

k, including the discretization into 8 scenarios of representative values. The left pane shows the results for a prior k_{repr} of $1.74e-4\,\text{m}\,\text{s}^{-1}$ (sections 2,3 and 5), the right pane for a k_{repr} of $3.47e-4\,\text{m}\,\text{s}^{-1}$ (sections 1 and 4). It has to be noted that these k-values also translate to different values for damping factor r_{exit} which is considered in the limit state functions for uplift and heave.

4.3.3. Case study results

We consider whether reducing the permeability uncertainty by monitoring is useful at sections 2, 3 and 4, where there is insufficient reliability for piping erosion. First we consider the conditional VoI, which is the VoI conditional on a certain monitoring outcome. Figure 4.13 shows the resulting conditional VoI for the segment. It is observed that the expected VoI ($E(\text{VoI})$) is positive for sections 2 and 3, and 0.0 for section 4. In line with expectation it is observed that more favourable monitoring outcomes (smaller non-exceedence probability $P(k_{\text{repr}})$ of the posterior k_{repr}) result in a higher VoI. An exception is section 4, where there are no direct benefits of monitoring: due to the availability of the relatively cheap Vertical Sandtight Geotextile (VSG) for reinforcement, there is no cost reduction by monitoring and in any case the section will be reinforced using the same method. A peculiar thing is that for monitoring at section 2 there is a negative VoI around $P(k_{\text{repr}}) = 0.8$, caused by the relatively large discrete steps that are used for the available measures (e.g. berm widening in steps of 10 meters) — here the obtained information causes a slightly different sequence of measures which results in a small negative VoI at the target reliability. At other reliability levels there is no difference in cost and reliability.

Each of the possible scenarios in Figure 4.13 has an underlying investment pattern that consists of a prioritized sequence of measures based on the optimal BC ratio of investment steps (see section 2.2). Figure 4.14 shows such a sequence for monitoring at section 2 with a favourable monitoring outcome ($P(k_{\text{repr}}) = 0.025$).

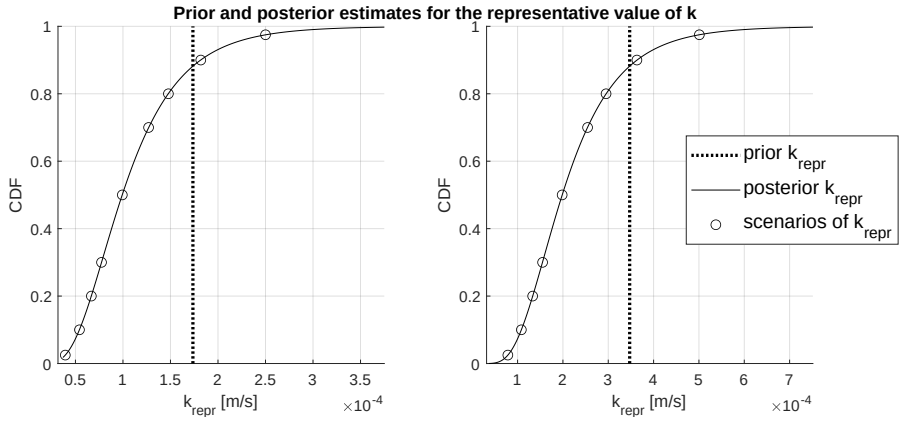


Figure 4.12.: *A priori estimate of the representative k (dotted vertical line) versus the posteriori distribution of k, after monitoring. o indicates a combination of a scenario considered in the calculations. Left and right pane show results for different a priori values for k_{repr}.*

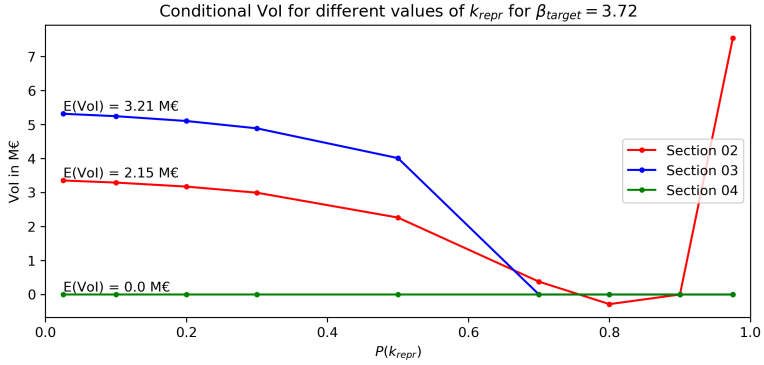


Figure 4.13.: *Conditional Value of Information for various values of k_{repr} for monitoring at different sections. $P(k_{\mathrm{repr}})$ denotes the non-exceedance probability of k_{repr} after monitoring. A small $P(k_{\mathrm{repr}})$ denotes a very favorable monitoring outcome. $E(VoI)$ denotes the VoI weighted by the scenario probabilities. The VoI is here the difference in investment cost to segment a certain target reliability β_{target} compared to the investment cost without monitoring.*

This shows that the measures taken at section 2 in the case without monitoring (red dashed line), do not increase the reliability, whereas in the case with monitoring (blue line), these measures are deemed unnecessary and not taken at all. This shows that monitoring improves the efficiency of investments, although it strongly depends on a variety of factors, such as the available reinforcement measures and the relative strength of a section, compared to other sections in a segment.

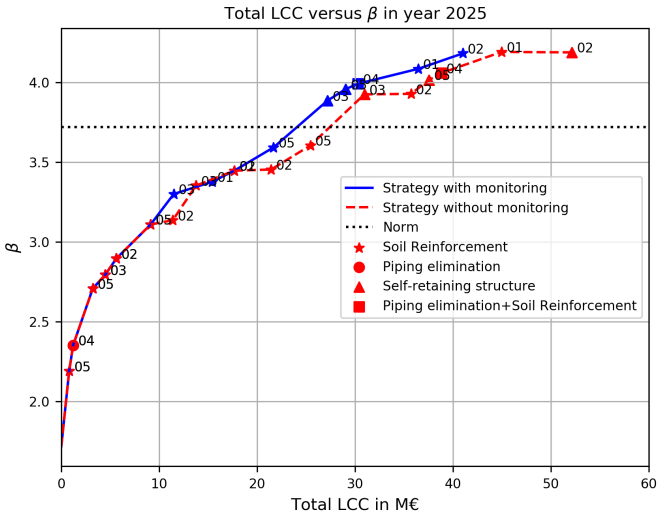


Figure 4.14.: *Total investment costs (LCC) versus the reliability for the entire segment for prioritized investments with and without monitoring for a scenario with* $P(k_{\text{repr}}) = 0.025$. *Symbols indicate different measures, numbers indicate the section. Where lines overlap identical measures are taken in both strategies.*

4.3.4. Discussion and conclusions

In this section we have considered a relatively simple example of pore pressure monitoring for a dike segment. The main goal of this section was to explore some of the effects of taking a system perspective rather than looking at a single dike section. Also on the level of a dike segment, the VoI of pore pressure monitoring to reduce uncertainty in aquifer permeability is typically positive. However, this changes if there are cheap reinforcement alternatives that substitute the reduced uncertainty: this can be a competitive alternative from a Total Cost perspective, as is observed for dike section 4 where the VoI was 0 for all scenarios, due to the availability of Vertical Sandtight Geotextile (VSG) as an intervention. This also demonstrates that in such cases investments in monitoring do not yield a reduction of Total Cost. Obviously, for such measures other (ir)reducible uncertainties might be of importance.

Reducing uncertainty on a dike segment level not only improves decisions at a dike section itself, but can also result in a reordering of the priority of interventions

for other sections, and possible additional investments for a Total Cost optimal solution as in general the reinforcement costs are lower due to the obtained information.

4.4. Practical implementation of uncertainty reduction of time-invariant parameters

The two case studies in this chapter have demonstrated that in many cases some investment in (a combination of) pore pressure monitoring and/or proof loading yields a positive VoI for dikes that are sensitive to slope stability and/or piping failures. However, the first case study showed that proof loading is not economically efficient in all the case variations considered, and in some cases also pore pressure monitoring has very limited benefits. We have identified two main reasons for this: first of all, not in all cases the behaviour of the phreatic surface is dominant for slope instability. In a practical situation, decision makers therefore have to carefully consider what are the uncertainties that dominate the reliability, and from that determine measures to reduce these uncertainties (if available). For example, the design point (see Figure 4.8) can provide indications to estimate the relative influence of different uncertain parameters, after which (sequential) decisions can be formulated based on the proposed decision tree framework. Secondly, it was found from the second case study that the reduction of uncertainty does not always contribute to a better solution of the decision problem: at dike section 4 in the second case study, reinforcement using a Vertical Sandtight Geotextile (VSG) largely eliminates the relative influence of uncertainty in aquifer permeability from the reliability estimate after reinforcement. Thus, in decisions on reducing uncertainty one should not only consider the uncertainty contribution in the current situation, but also the (expected) uncertainty contribution after different actions are taken. This is something that is facilitated in the pre-posterior analysis of the proposed decision tree framework as applied in both cases. Obviously, while a single failure mode for uncertainty reduction is considered in both cases, extension towards uncertainty reduction for multiple failure modes is straightforward and advisable towards practical implementation.

Chapter 3 demonstrated the potential influence of suboptimal reliability requirements on the cost-effectiveness of dike reinforcements. For the pre-posterior analysis at a section level, the sensitivity analysis in the first case study (Section 4.2.5) demonstrated the influence of (economically) suboptimal requirements on the VoI for uncertainty reduction. There it was clearly shown that using suboptimal requirements can have a large influence on the VoI and TC. Aside from different target reliability levels that are optimal for prior information, a case where the target reliability level is optimized based on the posterior information after reducing uncertainty is also presented. It is found that this increases the VoI, in particular if the marginal cost of a dike reinforcement varies for different dimensions of the reinforcement (i.e., different increases in berm length). Specifically for cases with highly non-linear cost functions or jumps in cost functions, a local optimization based on posterior information after uncertainty reduction efforts can increase the

effectiveness of uncertainty reduction, and flood risk management in general. As was shown in Chapter 3 and Section 4.3 this is often the case in practice.

If we consider optimal allocation of investments at a dike segment, it can be observed from the investments in the second case (see Figure 4.14) that also the order of investments at other (non-monitored) dike sections changes due to a change in marginal utility. Thus, if we aim to fully optimize investments in both dike reinforcement, and reduction of time-invariant uncertainty it is preferable to take into account the dike segment as a whole. A summary of the different findings on factors that influence the VoI from uncertainty reduction has been given in Table 4.4.

Especially in the first case study we simplified the decision problem to be more or less time-invariant, with constant reliability in time. In practice, as is accounted for in the second case, this is rarely the case and reliability estimates will change in time. However, in both cases we considered a single investment, and over longer time periods multiple investments in reinforcement might be necessary to keep dike segments up to the required safety level. In such cases, the benefits of reducing time-invariant uncertainties, for instance through pore pressure monitoring and proof load testing, but also site investigation, reach beyond the scope considered in this chapter. In Chapter 5 we extend the pre-posterior decision framework to include multiple future investments as well.

Table 4.4.: *A summary of findings on the impact of different influential factors on decisions on uncertainty reduction from the different cases in this chapter.*

Influential factors for decisions on uncertainty reduction	**Positive impact**	**Remark**
Proof load level	Higher proof load, more uncertainty reduction.	The increased risk of failure during a proof load test does not always outweigh the potential benefits, especially if consequential damage is high.
Optimization of target reliability before uncertainty reduction	Can lead to significant reduction of total cost	In practice only possible if economic risk is the governing risk indicator rather than e.g., individual risk.
Optimization of target reliability after obtaining information	Reduction of total cost through inclusion of obtained information in target reliability	If target reliability was already optimized this will only be beneficial in very specific cases where information results in a posterior that strongly differs from the prior.
Optimization based on total cost	Allows for optimization on the scale of a larger system	Including different (combinations of) uncertainty reduction might be computationally challenging
Larger geotechnical uncertainty for slope stability	Proof loading is more effective	Pore pressure monitoring might become attractive only after reducing geotechnical uncertainty. It is recommended to determine the sequence of measures based on their relative uncertainty contribution and consider other methods as well(e.g., site investigation).
Higher construction cost of stability berms	Uncertainty reduction methods are more attractive as the benefits are larger.	Other methods for reinforcement might be more effective.
Availability of uncertainty substitution measures	Can reduce cost of reinforcement	Value of additional information can become very small. Other uncertainties can become more dominant.

5

Uncertainty reduction for asset management of flood defences

Uncertainty is the friend of the buyer of long term values.

Warren Buffett

The preceding chapter has explored the influence that uncertainty (and reducing uncertainty) can have on reinforcement decisions. However, while Chapter 4 considers a relatively stationary situation with 1 action in the decision tree, reality is all but stationary. In this chapter, we assess pore pressure monitoring in the context of multiple cycles of reinforcement, in order to relate the Value of Information to long-term asset management decisions. This is placed in the context of significant uncertainty in future developments (specifically tied to sea level rise), an approach that allows us to evaluate the benefits of investments in uncertainty reduction for the distant future. As decision trees are a less adequate approach for recurring decisions, we use influence diagrams to set up a decision model to quantify the benefits of monitoring for flood defence asset management.

Parts of this chapter have been published in Klerk, W. J. et al. Value of Information of Structural Health Monitoring in Asset Management of Flood Defences. *Infrastructures* **4,** 56. doi:10.3390/infrastructures4030056 (2019).

THE previous chapter demonstrated that uncertainty reduction through amongst others pore pressure monitoring typically yields a significant Value of Information (VoI) for the next large investment in dike reinforcement. As uncertainties reduced through pore pressure monitoring are predominantly time-invariant, and as reliability of flood defences generally varies in time, the benefits of uncertainty reduction might extend beyond the next large investment and also hold value for investments in the more distant future. This is the topic of this chapter, where we primarily look at the long-term benefits of pore pressure monitoring.

As was outlined in Section 2.1, asset management principles can be applied in different contexts, which each have relevant temporal and spatial scales. In the context of strategic (asset management) decisions on flood defence systems, typically longer time scales of 100-200 years are considered, which also means that such analysis contain multiple reinforcements. With large uncertainties in for instance the future development of sea level rise — for which estimates are in the range of 0.5 to 3 meters for the Netherlands [93, 132] — it is likely that the investment pattern and frequency will change if circumstances change rapidly in the future. However, if we consider uncertainties nowadays to be time-invariant, they will likely still be time-invariant in 50 to 100 years from now. And while the relative influence of different uncertainties on parameters might change, they will still influence reliability estimates. Hence, reducing time-invariant uncertainty might hold value for multiple future reinforcements as well.

It should be noted that in this chapter we speak of Structural Health Monitoring (SHM), which is a term used for a variety of techniques used for improving the prediction of the health of structures. This can concern both detection of damage, as well as model identification and extraction of structural features [22]. Plenty of examples are available, for instance for monitoring of bridges [133], aircraft and wind turbines [134]. Also for flood defences several techniques for SHM are available, but the one most frequently applied is pore pressure monitoring [e.g. 21]. In the case of pore pressure monitoring we deal with time-invariant uncertainties: the behaviour itself does not change, but our estimates of behaviour are obscured by dominant time-invariant uncertainties. It is thus an application aimed at identification of the model that best describes the reliability of the structure.

The specific focus of monitoring for a flood defence is case-specific: for flood defences with relatively low reliability requirements the aim is typically to provide insight into anomalies and system responses during (frequently occurring) crisis situations [135]. For flood defences with relatively stringent reliability requirements such observations will be rare, and SHM systems will be mostly aimed at reducing uncertainty in reliability estimates [21, 122]. In this chapter we will focus on the latter.

The aim of this chapter is to identify the characteristics of cases where SHM systems are expected to yield the largest benefits for a long term horizon. We use a set of case studies that have representative characteristics in terms of uncertainty contributions of load and strength, based on previous flood risk analysis [e.g. 136]. Using these cases we aim to identify key characteristics for which a pore pressure monitoring system is particularly beneficial in terms of the VoI over a period of 200

years.

The decision analysis in this chapter is based upon the main principles outlined in Chapter 2, primarily in Section 2.7. However, as we consider recurring investments we use influence diagrams rather than decision trees to visualize the model structure. The basic principles and used approach are further outlined in Section 5.1. Section 5.2 introduces the parametrized case, while Sections 5.3 to 5.5 provide results, discussion and conclusions.

5.1. Methodology

5.1.1. Structuring decision problems using influence diagrams

In various studies on inspection and Structural Health Monitoring decision trees such as the one in Figure 2.8 have been extended with additional levels (see e.g., [22, 73] and Chapter 4). While decision trees are insightful, they have as a disadvantage that they grow exponentially when multiple sequential decisions are considered [47]. As an analysis of an asset management strategy for 100-200 years consists of multiple sequential and dependent decisions (i.e., a decision in $t+1$ should account for findings at t) simply repeating a decision tree is not practical for this type of problem. Sequential decision problems can be solved by defining policies or strategies, which are sets of decision rules that indicate what action has to be taken under which circumstances at which time step t [22, 26, 137]. An example is an inspection strategy where an inspection is carried out at interval δt, possibly followed by a repair given that some parameter $X < x$. The disadvantage of using policies is that it might not yield an optimal solution, although it was found by Luque & Straub [80] that the solution can be close to that obtained with other methods that are capable of finding an optimal solution, provided that the heuristics are well formulated. Of course, there are also examples [e.g. 27] where optimal policies have been derived. A heuristic approach can be combined with any type of temporal (Bayesian) decision model [26]. In this study we use a Bayesian decision model based on First Order Reliability Method calculations for the flood defence reliability in each year. In Chapter 7 a similar approach is used using a Dynamic Bayesian Network.

Influence diagrams are an extension of Bayesian Networks that also include decision and utility nodes [79]. Influence diagrams can aid in visualizing decision problems with multiple observations and decisions, for instance in the context of Dynamic Bayesian Networks [26]. As such, influence diagrams are well suited for the decision problem considered in this chapter.

The pre-posterior Bayesian decision model used in this chapter is based on the general decision tree in Figure 2.8 in Chapter 2. We evaluate the model for different monitoring strategies for different possible posterior states of the world θ. These states are a description of the system should all epistemic uncertainties be reduced. Strategies are defined as sets of heuristic policy rules, and contain all actions during the considered time period. Figure 5.1 gives an overview of the approach in this

chapter, including the subsections in which each part is discussed in more detail. The figure consists of 2 main parts: the top part concerns how input is derived for runs of the Bayesian decision model. An influence diagram that represents the Bayesian decision model is shown in the bottom part. For most blocks in the diagram parameters are given that relate to the parameters in the generic decision tree in Figure 2.8. Note that in this model observations z are translated to posterior failure probabilities ($P(Z < 0|\Theta, z)$, while in Figure 2.8 z was directly translated to actions through decision rule $d(z)$. In our modelling framework, several strategies are evaluated using a set of sampled possible posterior states that are consistent with the prior belief.

The remainder of this section will further outline the building blocks of the model. As the failure model is of importance for all other sections its general set-up is discussed first in Section 5.1.2. Next, Section 5.1.3 describes how prior distributions Θ are translated to samples of posterior states of the system θ and observations z, this concerns the left part of the figure (left of the dotted line). Section 5.1.4 deals with the right part of the figure, most notably how heuristic decision rules for monitoring and reinforcement (e and $d(P_{\mathrm{f}})$) are used to update $P_{\mathrm{f,state}}$ and $P_{\mathrm{f,bel}}$. In Section 5.1.5 it is discussed how the utility for each run can be translated to estimates for the cost and Value of Information.

5.1.2. Temporal failure probability model

While in the decision tree in Figure 2.8 the state of the observed variable could be translated to utility directly, in our case the utility depends on the failure probability of the flood defence. Hence we have to translate observations (z in the decision tree) to a failure probability. In the decision model we assess the performance of the flood defence by means of a fragility curve [138]. A fragility curve denotes the probability of failure given a realization of some parameter; for flood defences typically the water level h is used, so $P_{\mathrm{f|h}}$. For the simple limit state function $Z = h_{\mathrm{c}} - h$, with h_{c} and h the critical water level and water level in m +ref, if $Z < 0$ denotes failure it holds that:

$$P_{\mathrm{f}} = P(Z < 0) = P(h > h_{\mathrm{c}}) = \int_h P(h > h_{\mathrm{c}}|h) \cdot f_{\mathrm{h}}(h)\mathrm{d}h = \int_h F_{h_{\mathrm{c}}}(h) \cdot f_{\mathrm{h}}(h)\mathrm{d}h, \quad (5.1)$$

where P_{f} is the annual failure probability for the flood defence section considered. We assume that the uncertainty in h is fully aleatory and irreducible meaning that no measure is available to reduce this uncertainty.

Over time both h_{c} and h will change due to respectively deterioration (e.g., settlement) and climate change induced increase in extreme water levels. The time-variant limit state function is then described by:

$$Z(t) = [h_{\mathrm{c}}(t) - \Delta h_{\mathrm{c}}(t)] - [h + \Delta h(t)], \quad (5.2)$$

where $\Delta h_{\mathrm{c}}(t)$ and $\Delta h(t)$ denote the deterioration respectively water level increase in meters. In our decision problem we consider the Value of Information from pore pressure monitoring. Pore pressure monitoring would reduce a part of the

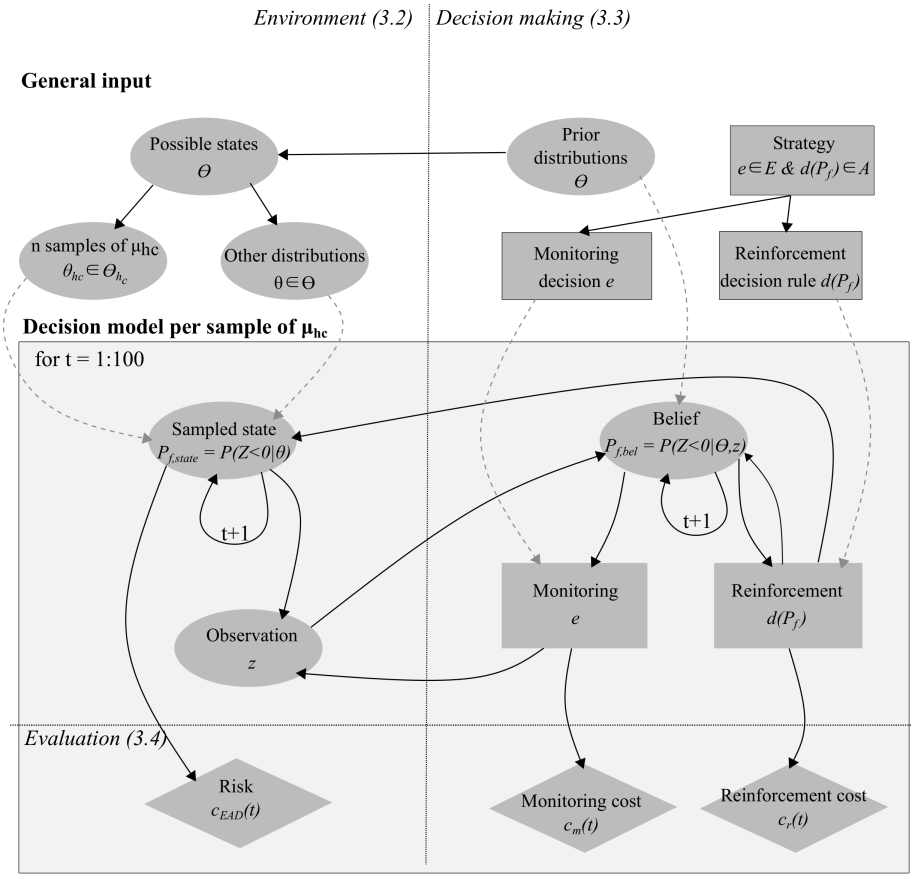


Figure 5.1.: *Overview of the methodology. The top part shows in general how the input values are related to each other. The bottom part shows an influence diagram for the Bayesian decision model which is run for each sampled state. Dashed arrows indicate how general input is transferred to the model per sample. The dotted line in the middle indicates in which section (Section 5.1.3 or Section 5.1.4) the various parts are discussed. Parameters in blocks relate to Figure 2.8.*

uncertainty in h_c(t). Hence, within the constraints of our decision problem, we consider all other uncertainties to be irreducible. It has to be noted that in Klerk et al. [137] uncertainty reduction in $\Delta h_c(t)$ and $\Delta h(t)$ was also considered. For $\Delta h_c(t)$ this was found to have very limited effect, whereas $\Delta h(t)$ is better dealt with using Value of Information conditional on different scenarios for $\Delta h(t)$. Both are considered to be deterministic in evaluations of this Bayesian decision model in Section 5.3.1, in Section 5.3.2 we will consider multiple scenarios for $\Delta h(t)$.

While pore pressure monitoring can reduce a part of the uncertainty in h_c(t), there is also an irreducible part. It should be noted that this part is irreducible by pore pressure monitoring, but there might be other methods to reduce this uncertainty. However, we do not consider these here. We can describe the uncertainty in h_c as follows:

$$h_c \sim N(\mu_{h_c}, \sigma_{irr}), \tag{5.3}$$

$$\mu_{h_c} \sim N(\mu, \sigma), \tag{5.4}$$

where μ_{h_c} denotes the mean of a possible state of the flood defence. σ_{irr} denotes the part of the uncertainty that is irreducible in the decision problem. How this definition is translated into values for $P_{f,state}$ and $P_{f,bel}$ is discussed in the following sections.

5.1.3. Environment

General input

The environment describes the state space of the flood defence which consists of all sampled posterior states ($\theta \in \Theta$ in the decision tree). Equation (5.2) shows that there are 4 random variables in the Limit State Function that determine this state. In our analysis we only consider observations of parameter h_c. Based on our prior belief (Θ_{h_c}) there are many possible posterior states for parameter h_c in Figure 2.8) which is reflected by the fact that μ_{h_c} is normally distributed with mean μ and standard deviation σ (see Equation (5.4)). This distribution of μ_{h_c} reflects the epistemic uncertainty of our state space [47]. Thus it holds that:

$$\Theta_{h_c} \sim N(N(\mu, \sigma), \sigma_{irr}), \tag{5.5}$$

$$\theta_{h_{c,j}} \sim N(\mu_j, \sigma_{irr}) \text{ with } j = 1, 2, \dots, n, \tag{5.6}$$

where n is the number of samples drawn from μ_{h_c} (Equation (5.4)) and $\theta_{h_{c,j}}$ is the j^{th} sample of a possible posterior state.

For the variables governed by a temporal process ($\Delta h_c(t)$ and $\Delta h(t)$) we use deterministic variables for the rate. In reality both variables will contain at least some uncertainty, but it was shown by Klerk et al. [137] that reducing uncertainty in $\Delta h_c(t)$ has little influence on decisions. For $\Delta h(t)$, some recent scenarios show that sea level rise in 2100 might range between 0.5 and 3 meters [132], although this has been nuanced by [139, 140]. In the pre-posterior analysis we do not include this uncertainty such that the effect of monitoring of h_c is more clear from the

results. Thus the prior distribution is a deterministic distribution with a deterministic annual rate $\Delta h(t)$. To assess the effect of different future rates of $\Delta h(t)$ we analyse the Value of Information conditional upon $\Delta h(t)$. Thus, for each possible value a separate pre-posterior analysis will be carried out. This will be discussed further in Section 5.3.2.

Decision model

For every time step Equation (5.2) is re-evaluated in order to account for the temporally changing variables. Thus for every time step a failure probability $P_{\text{f,state}} = P(Z < 0|\theta)$ is computed. A reinforcement decision might incrementally change the belief of h_{c} for the next time step. For strategies where monitoring equipment is installed an observation of h_{c} is sampled from the state.

5.1.4. Decision making

General input

Decisions are defined using strategies consisting of heuristic decision rules S_{j}. A heuristic rule typically has the form: if *some_variable* is larger than *some_threshold* we take *some_action*. The model contains decision rules for monitoring and reinforcement. We consider three different sets of decision rules for monitoring ($i \in I$ in Figure 2.8):

- Strategy a: no monitoring.
- Strategy b: monitoring is started if the failure probability $P_{\text{f,bel}} > 0.5 \cdot P_{\text{f,req}}$, where $P_{\text{f,req}}$ is the reliability requirement. Monitoring is stopped after 25 years.
- Strategy c: continuous monitoring starting at $t = 1$.

For reinforcement decisions the same rules are used in all calculations: if $P_{\text{f,bel}}(t) < P_{\text{f,req}}$ a flood defence is reinforced such that $P_{\text{f,bel}}(t + t_{\text{design}}) = P_{\text{f,req}}$. Here t_{design} is the design period of the flood defence. In terms of the decision tree in Figure 2.8 this means that we have only one set of $d(P_{\text{f}}) \in A$ which is identical for all t.

Decision model

In order to obtain pre-posterior estimates for the cost of each strategy, we evaluate a set of possible posterior states j. For each sampled posterior state j (see Equation (5.6)), the belief estimate of the reliability ($P_{\text{f,bel}}(t) = P(Z < 0|\Theta, z(t))$ in the decision model is recalculated for every time step in order to account for the temporally changing variables and reinforcement decisions. Additionally observations from monitoring can result in an updated belief using the observations from monitoring up to time t, $z(t)$. The initial belief of h_{c} is defined as: $\Theta_{\text{h}_{\text{c}}} \sim N(N(\mu, \sigma), \sigma_{\text{irr}})$.

If monitoring equipment is present, an observation of the state might be made, depending on the extremity of the observed circumstances. While in Klerk et al. [122] it was found that more extreme circumstances gradually increase the VoI,

here we use a discrete threshold value to distinguish between years with and without useful observation. Whether an observation at $t = i$ is useful is determined by the following condition:

$$P(h_{\mathrm{i}} > H) < P_{\mathrm{thresh}}, \tag{5.7}$$

where $P(h_{\mathrm{i}} > H)$ is the annual exceedance probability of a randomly sampled water level h_{i}, and P_{thresh} is a predefined threshold value that has to be exceeded for an observation to be useful for updating the belief of h_{c}. Analogous to a Probability of Detection (PoD) for inspections this can be interpreted as a Probability of Observation [22, 73]. As both prior and posterior of h_{c} are normally distributed, the conjugate distributions can be used to obtain the posterior distribution at time t, $h_{\mathrm{c}}|z(t)$ [47]:

$$h_{\mathrm{c}}|z(t) \sim N\left(\frac{n\bar{z}}{n' + n}, \sqrt{\frac{\sigma_{\mathrm{irr}}^2}{n' + n} + \sigma_{\mathrm{irr}}^2}\right), \tag{5.8}$$

where the prior weight is given by $n' = \sigma_{\mathrm{irr}}^2/\sigma'$ with $\sigma' = \sqrt{\sigma^2 - \sigma_{\mathrm{irr}}^2}$, n is the number of observations obtained until t, and $\bar{z} = \sum(z(t))/n$ is the mean of the observations up to t. The observations are random samples from the considered state $\theta_{\mathrm{h_{c,j}}}$ (see Equation (5.6)). The effect of this approach is that for longer periods of monitoring the expected number of useful observations will increase. Also, for a short period of monitoring the number of useful observations can vary, and there might not be any information obtained at all.

A reinforcement is carried out if the flood defence no longer meets the required minimum annual failure probability requirement $P_{\mathrm{f,req}}$. In such a case the mean μ of h_{c} is iteratively increased such that it holds for the belief of h_{c} that:

$$P(Z(t + t_{\mathrm{design}}) < 0) > P_{\mathrm{f,req}}, \tag{5.9}$$

where t_{design} is the design period in years. In case of a reinforcement the Coefficient of Variation of the belief of h_{c} is assumed to be the same before and after reinforcement.

5.1.5. Evaluation

For each evaluation of a strategy for a sampled possible state θ we compute 3 discounted cost components for the evaluated period of n years: the Expected Annual Damage (EAD) due to flooding (C_{EAD}), the cost of monitoring equipment (C_{m}) and the cost of reinforcement (C_{r}). The overall discounted cost of a strategy for a sample $\theta \in \Theta$ over a period of n years can be written as:

$$c_{\mathrm{s}|\theta} = C_{\mathrm{EAD}} + C_{\mathrm{m}} + C_{\mathrm{r}} = \sum_{t}^{n} \frac{c_{\mathrm{EAD}}(t) + c_{\mathrm{m}}(t) + c_{\mathrm{r}}(t)}{(1 + r)^t}, \tag{5.10}$$

where $c_{\mathrm{EAD}}(t)$ is the cost component of the EAD at time t and $c_{\mathrm{m}}(t)$ and $c_{\mathrm{r}}(t)$ are the costs for monitoring and reinforcement at time t. Note that these are equal to 0

if no reinforcement or monitoring is done at a specific time step t. r is the discount rate, here we assume a value of 3%, although discount rates in the Netherlands have been revised recently [141].[1]

We study the Value of Information (VoI) of different asset management strategies based on their performance over a time span of 200 years. If we consider a strategy s_j where information is acquired the VoI can be calculated as the difference in expected costs based on all samples θ with the expected costs of baseline strategy a ($E(c_a)$):

$$\text{VoI}(s_j) = E(c_a) - E(c_{s_j}) \text{ with:} \tag{5.11}$$

$$E(c_{s_j}) = \frac{\sum_N c_{s_j|\theta}}{N}, \tag{5.12}$$

where $\text{VoI}(s_j)$ is the Value of Information for strategy s_j and N is the number of samples of θ that are considered.

If the VoI is *conditional* upon a value for an observed parameter (in our case h_c), it is defined as a conditional Value of Information (cVoI) or option value [47, 142]. This can be used to assess the VoI given the possible outcomes of monitoring (e.g., the VoI given θ or z). To assess the influence of future uncertainty on the Value of Information we use a slightly modified formulation of this concept in Section 5.3.2. Here we do not formulate the VoI conditional on the monitored parameter itself (h_c), but rather on a different model parameter, namely the value for the change in water level $\Delta h = h_j$. Thus the conditional VoI for a strategy s_j can be obtained using the following equation:

$$\text{cVoI}(s_j|\Delta h = h_j) = E(c_A(\Delta h = h_j)) - E(c_{s_j}(\Delta h = h_j)) \tag{5.13}$$

It has to be noted that this conditional VoI is formulated differently than the conditional Value of Information concept given by Raiffa & Schlaifer [47].

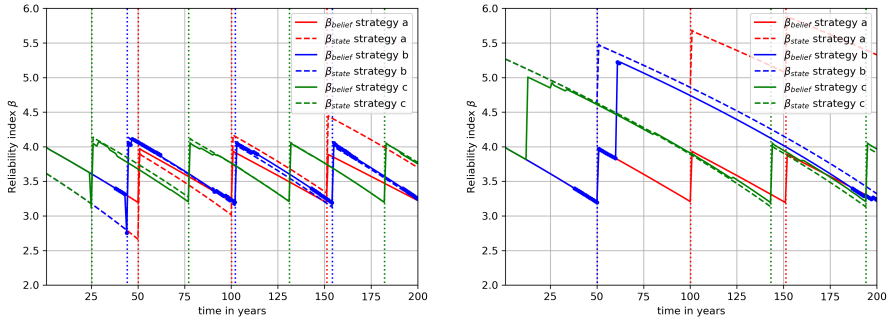
5.2. Case study

In order to determine the VoI of pore pressure monitoring we consider a set of parametrized cases, based on values obtained in actual probabilistic assessments such as VNK2 [136]. Each case is obtained by modifying input distributions such that the initial reliability index β is around 4 ($P_f \approx 1/30000$). We parametrize in two ways: First we consider different sets of the α influence coefficients in the design point obtained from calculations using the First Order Reliability Method (FORM). A high influence coefficient α indicates that the uncertainty in a parameter has a major influence on the reliability index β [143]. From probabilistic calculations throughout the Netherlands it is found that for instance for the failure mode piping erosion the α^2 of the strength can vary between 0.1 and 0.9. Next to that, we vary the extent to which the uncertainty in the strength is epistemic and reducible. The

[1] In Werkgroep Discontovoet [141] lower discount rates have been advised. Lower discount rates result in higher Life Cycle Cost, as such the Value of Information will increase.

Table 5.1.: *All input data for cases A, B, C and D. The top part shows all input distributions, the middle part shows the initial design point values for the influence of the strength uncertainty and epistemic part of the strength uncertainty, as well as the prior reliability index β_{prior}. The bottom part shows the threshold P_{thresh} for the normal case and a lower threshold, safety standard P_{norm}, planning period t_{plan} and discount rate r for each case.*

	Name	Unit	Distribution		Case			
			Type	*Values*	*A*	*B*	*C*	*D*
Input distributions	μ_{h_c}	m +ref	Normal	μ	7.56	6.12	5.76	5.88
				σ	1.03	0.58	0.42	0.50
				σ_{irr}	0.2	0.2	0.2	0.2
	h	m +ref	Gumbel	a	3.2			
				b	0.2			
	Δh	mm yr^{-1}	Determ.		8			
	Δh_c	mm yr^{-1}	Determ.		5			
Initial design point	β_{prior}	-			3.96	4.00	4.01	4.03
	$\alpha^2_{h_c}$	-			0.74	0.57	0.34	0.41
	$\alpha^2_{\mu_{h_c}}$	-			0.69	0.51	0.28	0.39
Other input	P_{thresh}	yr^{-1}			Default: 0.1 Lower: 0.5			
	P_{norm}	yr^{-1}			1/3000			
	t_{plan}	yr			50			
	r	% yr^{-1}			3			



(a) *Unfavourable sample* (b) *Favourable sample*

Figure 5.2.: *Two calculations of β in time for samples $x_i \in \mathbf{x}_n$ for Case B. The left pane shows an unfavourable sample ($\beta_{\text{state}} > \beta_{\text{belief}}$). The right pane a favourable sample. Circled markers on the line for strategy b indicate presence of monitoring equipment. Dotted vertical lines indicate a reinforcement.*

highest VoI is expected to be encountered for cases with a large α of the strength, of which a major part is epistemic.

Additionally we consider 2 different threshold values at which an observation can be made (P_{thresh}). The default threshold represents a flood defence at a lowland river with larger floodplains bounded by summer dikes. Hence one would expect a relatively high threshold ($P_{\text{thresh}} = 1/10$ yr^{-1}) in Equation (5.7), as water has to overflow the summer dike before a measurement can be obtained. We also consider a lower threshold which represents a location without summer dikes in a delta region relatively close to the sea ($P_{\text{thresh}} = 1/2$ yr^{-1}).

For the costs of reinforcement the relation for dike ring 16 given in Eijgenraam et al. [16] is used, with which the exponentially increasing costs of an incremental increase in h_c can be obtained. The costs for installing monitoring equipment are assumed to be €200.000. All other input values as well as prior design point values are shown in Table 5.1.

For each considered case 250 different posterior states are sampled from the state space. In some cases samples are drawn from the state space that have a very low reliability after 50 years ($\beta < 2$) and have a probability smaller than $1/N$, where N is the sample set. Such samples result in extremely high risk costs which dominate the Total Cost computation. In practice, dependence between years in such cases will result in a higher β which we do not account for here (see Section 2.5 for background). Therefore all samples where $\beta(t = 50) < 2$ are removed from the state space, this will be further discussed in Section 5.4.

5.3. Results

5.3.1. Benefits of monitoring for a deterministic future

Our first analysis concerns the benefits of monitoring without uncertainty in future development (temporal changes in load and resistance are deterministic). In order to better understand the way individual samples of posterior states θ are dealt with, we examine the influence of monitoring on the reliability in time for 2 sampled posterior states for Case B. This is shown in Figure 5.2. Figure 5.2a shows a relatively unfavourable sample of $\mu_{\mathrm{h_c}}$, which means that the posterior reliability after obtaining observations is lower than the prior estimate. Figure 5.2b shows a favourable sample, here the posterior reliability is higher than the prior. For each strategy two computations of β are shown: one for the sampled posterior state β_{state} (dashed), this indicates the reliability should all epistemic knowledge be reduced. A second line is shown for the development of the decision makers' belief over time β_{belief} (solid line). The left pane shows that in strategies with monitoring (b and c), reinforcement is done earlier than without monitoring, provided that observations are obtained which confirm the unfavourable $\mu_{\mathrm{h_c}}$. Hence life-cycle costs will be higher with monitoring, but risks costs will be lower. In Figure 5.2b the sampled $\mu_{\mathrm{h_c}}$ is much more favourable and for strategy c, reinforcement is postponed by about 85 years. For strategy b the reinforcement is not postponed, as there is no useful observation in the first 20 years of monitoring (indicated by the line marked with circles). However, in the entire sample set there are realizations with a similar $\theta_{\mathrm{h_c}}$ where there are observations and the reinforcement is also postponed for strategy b. The ratio between cases with and without postponed investment depends on the value of P_{thresh}.

This is shown on a more general level in Figure 5.3. Here the VoI for strategies b and c compared to reference strategy a is represented by a (Gaussian) kernel density estimation (KDE). It can be observed that for strategy b (in red) there are two peaks in density, whereas for strategy c there is only one peak. As the first reinforcement is after about 50 years, for strategy c the probability of not having any useful observation before that time is quite small (about $1/200$). As strategy b only monitors for a limited amount of time, this probability is much larger (about 30 %), thus two modes are found in the KDE estimation for each case. The left mode corresponds to cases where the first reinforcement is not postponed, either because observations have not been obtained, or because the posterior strength is still insufficient. The right mode corresponds to cases where it is postponed. This difference illustrates how important it is to include the uncertainty on whether an observation will be obtained during the envisaged monitoring period. The influence on subsequent reinforcements is not clearly visible from the KDE, as the life-cycle costs are significantly lower due to discounting.

Theoretically it is expected that $E(\mathrm{VoI})$ decreases if the relative contribution of reducible uncertainty in $\mu_{\mathrm{h_c}}$ decreases. In Figure 5.3 we can see that this is indeed the case: cases with a lower influence coefficient of reducible uncertainty in the strength $\alpha_{\mu_{\mathrm{h_c}}}$ have a lower $E(\mathrm{VoI})$. Figure 5.4 shows this more clearly, also for an additional case where $\alpha_{\mu_{\mathrm{h_c}}}$ is even smaller. From this figure we see that the $E(\mathrm{VoI})$

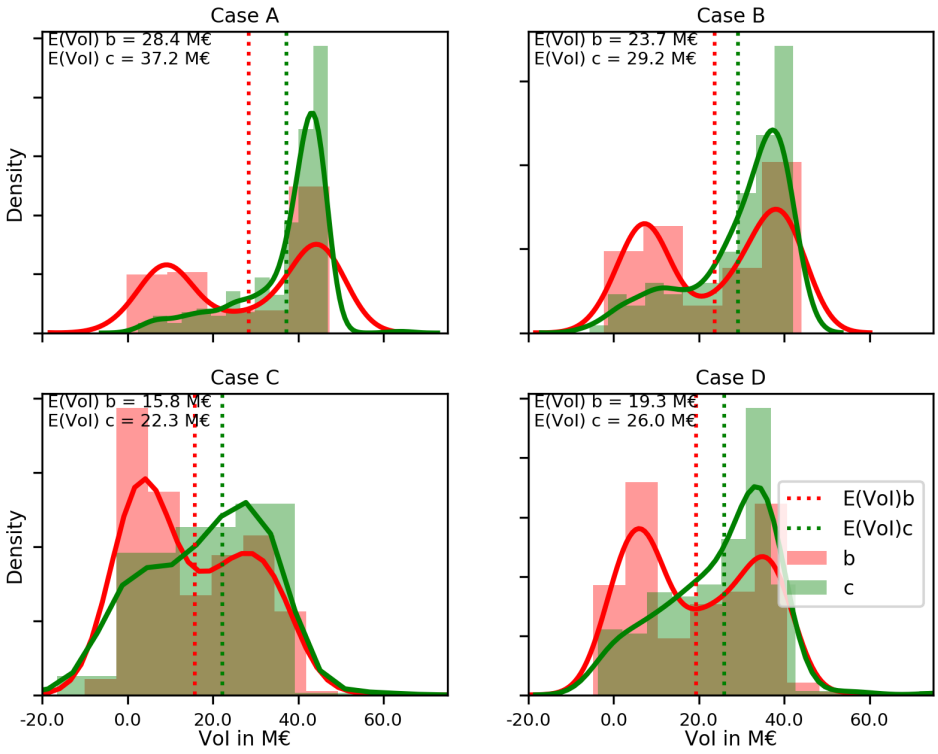


Figure 5.3.: *Value of Information (VoI) for cases A to D. Thick lines represent a Gaussian kernel density estimation (KDE). Bars denote the histograms of the underlying samples. Dotted lines represent computed expected Value of Information for both strategies b and c, for which values are shown in the left top of each figure.*

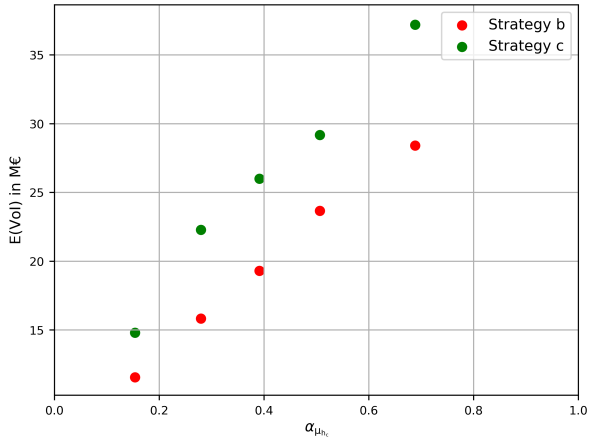


Figure 5.4.: *Relation between the influence coefficient of the strength* $\alpha_{\mu_{h_c}}$ *and the expected Value of Information (VoI).*

clearly increases with larger influence of epistemic uncertainty.

In relative terms it is important to note that the Total Cost for all cases is on average in the order of 50-60 M€, where it is low for most samples but the average is strongly influenced by cases where the flood defence strength is more unfavourable than expected, and the risk costs are above the a priori estimate. It is found that on average the Total Cost for strategy b is between 40 and 70 % of the reference strategy a, and between 25 and 60% for strategy c.

It was shown before that monitoring benefits strongly depend on the duration of monitoring and the probability of having a useful measurement before the next major investment decision. Hence we compare case B with an identical case where it holds that $P_{\text{thresh}} = 1/2$ yr^{-1} rather than the default value $P_{\text{thresh}} = 1/10$ yr^{-1}. We would expect the left peak in the density estimate to be significantly smaller for strategy b in case B with lowered threshold. Say that there are 10 years before the planned reinforcement, the probability of having no useful observation for a lower threshold compared to the default case B is a factor

$$\frac{P(\overline{\text{B}})}{P(\overline{\text{B}_{\text{low threshold}}})} = \frac{(1-P_{\text{thresh}})^t}{(1-P_{\text{thresh,lower}})^t} = \frac{(1-0.1)^{10}}{(1-0.5)^{10}} \approx 357 \tag{5.14}$$

smaller, where $P(\overline{\text{B}})$ is the probability of having no observation for case B, and t is the period up to the next major decision. Figure 5.5 shows KDEs for both cases. Here it can be observed that for case B with lower threshold there are no longer two peaks in density for strategy b (in red): KDEs are very much alike for both strategy b and c. This again illustrates the importance of including the relation between monitoring duration and expected (number of) observations.

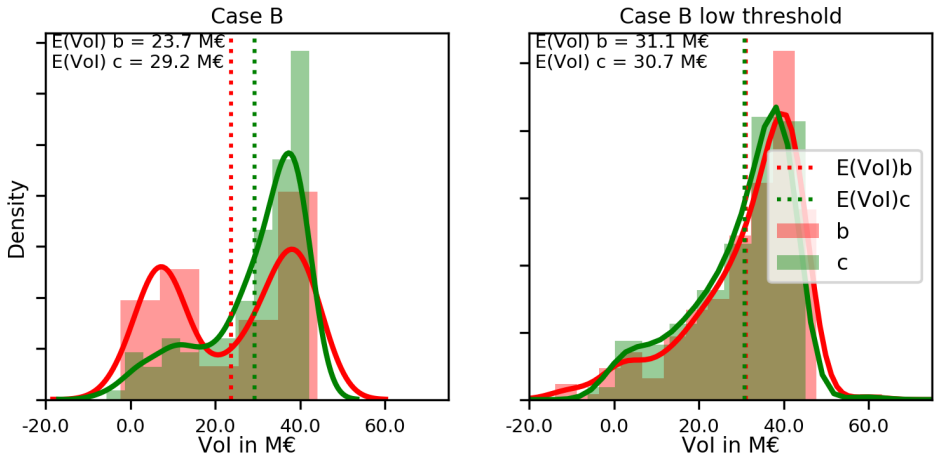


Figure 5.5.: *VoI for case B with default and lower threshold. Thick lines represent a Gaussian kernel density estimation (KDE) and bars the histograms of underlying data. For B it holds that* $P_{\text{thresh}} = 1/10\ \text{yr}^{-1}$ *for B low threshold* $P_{\text{thresh}} = 1/2\ \text{yr}^{-1}$.

5

5.3.2. The effect of future uncertainty on value of information of monitoring

In the previous section we assumed deterministic changes of load and resistance. However, in reality especially the future development of the load is very uncertain, amongst others due to sea level rise (SLR) and changing hydrological conditions. Contributing factors such as Antarctic ice sheet mass loss might or might not result in an increasing acceleration of global sea level rise, for which high-end estimates are in the order of 3 meters by the end of the 21st century [132, 144]. Recent publications have argued that these estimates are likely too high but that there are many remaining uncertainties towards the precise contribution of ice sheet mass loss to sea level rise [139, 140]. Studies on sea level rise observations have found some indication of acceleration from satellite observations [145], whereas others have not found this from tidal gauges [146, 147]. For river discharges the effects of a changing climate vary per region but there are several examples of catchments where discharges, and thus water levels are expected to increase [148]. Given the large uncertainty that exists it is of interest to identify actions that are beneficial in a wide variety of future scenarios.

In the context of this chapter, the exact rate of water level increase is not directly of interest to the question whether SHM should be used to reduce epistemic uncertainties, but rather the question whether the VoI is sensitive to the range of possible future scenarios. Hence, in this section we explore whether an investment in efforts to reduce epistemic uncertainties is robust in the sense that the VoI is positive for the entire range of possible rates of water level increase.

In this section we explore this for Case B from the preceding section, which has a

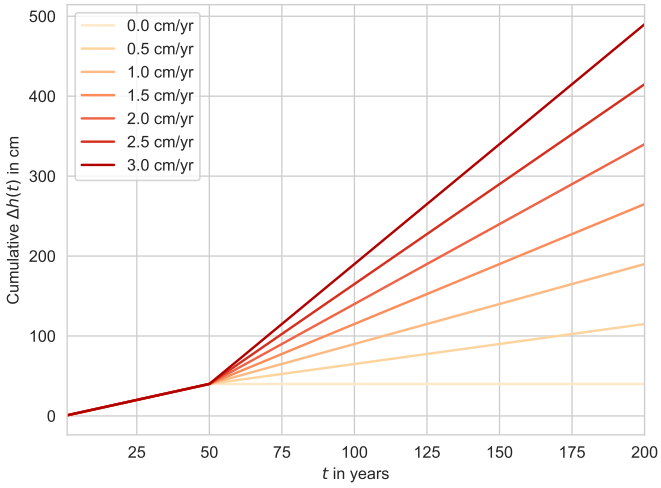


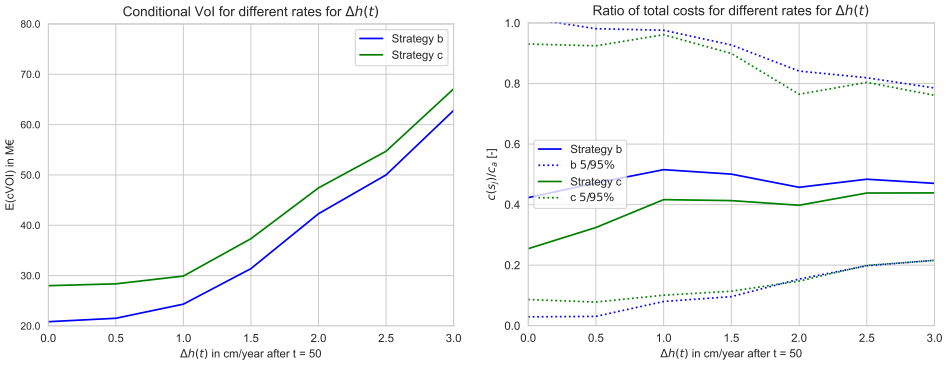
Figure 5.6.: *Considered scenarios for the increase in water level* $\Delta h(t)$.

commonly encountered value for influence coefficient α_{h_c}. We consider a range of scenarios with increases in water levels $\Delta h(t)$ between 0 and 3 cm/yr, in line with the ranges given by [132, 149]. We assume that $\Delta h(t)$ changes from the original value to the scenario rate at t=$50\,$yr, this is shown in Figure 5.6.

Figure 5.7a shows the conditional VoI given different rates of $\Delta h(t)$. Here we see that the VoI increases for both strategy b and c for higher rates of water level increase. In Figure 5.7b the ratio of the Total Cost with and without monitoring is shown for the different scenarios, including 5/95% upper and lower bounds. This ratio is quite stable, as aside from the increase in VoI also the total investment costs increase significantly as there are more reinforcements needed to maintain the reliability requirement. Also the overall uncertainty reduces slightly for more extreme rates of water level increase. This is explained by the fact that the two modes observed in the KDE-estimates disappear for high rates of water level increase. The reason is that the Total Cost is not completely dominated by the first reinforcement, but also by the number of consecutive reinforcements which increases for high rates of water level increase. Hence the KDE-estimates become more smooth for more extreme changes.

5.4. Discussion

In this chapter we have focused on the Value of Information of reducing epistemic uncertainty in flood defence strength, specifically for asset management decisions over a longer time horizon. The two key points considered are first of all whether —compared to Chapter 4 — inclusion of multiple reinvestments over a longer period of time has a significant influence on estimates of the VoI, and secondly whether uncertainty reduction holds its value, even if considered from the perspective of



(a) *conditional Value of Information* (b) *Relative reduction of Total Cost ($c(s_j)$).*

Figure 5.7.: *Conditional and Relative VoI for different rates of water level increase after $t = 50$ for Case B. Relative VoI is the Conditional VoI normalized by the Total Cost.*

5

very uncertain future developments.

Compared to Chapter 4, this chapter further elaborates the influence of the probability of an useful observation — multiple thresholds have been considered and found to influence the VoI. It should be noted that still a relatively discrete threshold is considered: in reality the relation between observed water level and the obtained information might be more gradual: medium-high water levels might yield some information, whereas high water levels yield even more information. Also the duration of a high water situation might influence the VoI. This can also be considered relatively easily by extending the decision trees, both in this chapter and in the decision problem in Chapter 4.

In this section we simplified the reliability model using a simple fragility curve rather than specifying a model for a specific failure mode. This is due to the fact that in this chapter we focus more on the general aspects of uncertainty reduction from a life-cycle perspective, rather than for a more specific case such as in Chapter 4 or in Klerk et al. [122]. Obviously the simple limit state function used here can be easily changed towards a more complex and realistic failure model.

In many cases the VoI is based on the reduction in investment cost, but here we also looked at the risk costs. This is particularly important as the cost of failure increases exponentially with an increase in failure probability when the resistance is lower than expected (i.e., state θ is very unfavourable). Hence monitoring has two potential benefits: reducing failure costs for unfavourable posterior outcomes and reducing investment costs for favourable posterior outcomes. In line with what is experienced in practice, in most of the cases the posterior after monitoring is favourable. Only VoI analysis that include the cost of failure of flood defences yield insight in both these benefits.

A point of attention towards the approach used is that some samples contain posterior states that have very high failure probabilities. In such cases the risk exponentially increases, resulting in single samples that have a very large influence on

the VoI estimates. As we consider uncertainties that are at least partially dependant in time, such reliability estimates are unrealistic (see also Section 2.5). In this chapter we have therefore excluded all samples for which $\beta(t = 50) < 2$. Potentially more elegant solutions to this are including survival observations or explicitly computing conditional failure probabilities, although the general effect is likely the same.

From the results with deterministic temporal changes we see that there is a strong relation between the influence coefficient of reducible uncertainty in critical height $\alpha_{\mu_{h_c}}$ and the VoI that is obtained (Figure 5.3), which is in line with the original hypothesis. Also the P_{thresh} at which valuable observations are obtained has a significant influence on VoI outcomes (Figure 5.5). Design loads for flood defences typically have very small probabilities of exceedence, meaning that deriving meaningful information from observed loads often requires that monitoring equipment is present during a relatively rare event. As it is found that mainly the cost for the first reinforcement dominates the Total Cost (due to discounting), this has to be accounted for in estimating the VoI. A consideration between the duration of a monitoring campaign and the probability that a useful observation is obtained is of importance in the context of methods for reducing time-invariant uncertainty such as pore pressure monitoring.

On the long term, development of the loads on flood defences is highly uncertain. In such cases the investment required to deal with this uncertainty in design can be large. Hence it is of interest to identify investment options that are beneficial in a wide variety of scenarios. In the analysis in this chapter it was shown that the conditional VoI can be a useful measure for this[2]. A particular advantage is that it gives more insightful information than an expected VoI, especially if there are other large uncertainties aside from the epistemic uncertainty that is reduced. It was shown that both the considered monitoring strategies have a positive VoI and yield a significant reduction of Total Cost in all considered future scenarios. Thus investments in reducing epistemic uncertainties are concrete and economically efficient options for preparing for potentially large future changes in load conditions. it should be noted that here we modelled reinforcements through an exponential cost function, i.e., we only considered reinforcements where uncertainty is mitigated. However, within the context of large epistemic uncertainties, consideration of reinforcement methods that substitute uncertainty (see Section 2.6) can be very effective as was demonstrated in Chapter 3 and Section 4.3. Especially in the context of long term asset management with multiple reinvestments, including such measures, as well as valuing the flexibility/expandability (e.g. through real options analysis [84, 150]) of such measures is an important part of the solution space.

5.5. Concluding remarks

This chapter has explored the benefits of uncertainty reduction through monitoring for flood defences for long-term investments, specifically types of monitoring where

[2]It should be noted that de conditional VoI considered here is slightly different from the one proposed by Raiffa & Schlaifer [47].

the amount of information obtained depends on observed loads such as pore pressure monitoring. A Bayesian pre-posterior decision model and the concept of Value of Information (VoI) have been used to quantify the benefits of different monitoring strategies. It has been shown that the VoI is directly related to the relative influence of epistemic uncertainties which can be easily obtained from probabilistic computations through for instance influence coefficients of different random variables. In many cases a monitoring campaign of some duration will be needed to obtain the information needed to reduce epistemic uncertainty. This has been investigated by comparing different threshold values at which information is obtained. The value of this threshold was found to be of major influence on monitoring outcomes and this aspect should therefore be considered in determining the duration of monitoring campaigns. This is of specific importance as often such campaigns are to be executed before a major reinforcement (see e.g. Chapter 4), and here we found that the first reinforcement has the largest impact on the Total Cost, and hence the VoI. Nevertheless, also subsequent flood defence reinforcements become more efficient if time-invariant uncertainty is reduced.

Many investments in flood defences have to consider large uncertainties in future loads due to climate change. Therefore the VoI of different scenarios for future load increase on a flood defence has also been considered. The VoI conditional on different rates of increase was found to be positive for all investigated rates. As Total Cost increases for higher rates of change, the relative cost reduction was similar over all different rates. Thus, monitoring aimed at reducing (time-invariant) epistemic uncertainty is an economically efficient investment in preparation of potentially large future changes. It should be noted that in this chapter we only considered a relatively simple reinforcement by means of uncertainty mitigation, but in the context of long term asset management it is advised to extend the analysis by valuing the flexibility of reinforcement measures and extending the solution space with measures that substitute dominant uncertainties.

6

Accuracy of visual inspections of flood defences

We have a very narrow view of what is going on.

Daniel Kahneman

In the preceding chapters we have looked at the influence of time-invariant uncertainty on performance of flood defence systems. However, also time-variant processes such as degradation have an impact on flood defence performance. Many degradation processes are shock-based (see Section 2.5) and manifest themselves as damage to the flood defence, in particular the revetment. A cornerstone in detection of these damages is visual inspection, which is done on a regular basis for most flood defences. However, the accuracy of such inspections is not known for flood defences, and the influence of the detected (and non-detected) damages is typically not considered in reliability analysis. For other applications the probability of not detecting damages in visual inspections is found to be considerable. In this chapter we describe an experiment where the accuracy of visual condition inspections was determined, which we can then use for including time-variant uncertainty due to degradation in reliability analysis for a flood defence system in Chapter 7.

The majority of this chapter has been published in Klerk, W. J. et al. Accuracy of visual inspection of flood defences. *Structure and Infrastructure Engineering,* 1–15. doi:10.1080/15732479.2021.2001543. https://www.tandfonline.com/doi/full/10.1080/15732479.2021.2001543 (2021).

IN between flood defence reinforcements, asset managers have to maintain flood defences in the required condition. Such maintenance is for a large part aimed at the revetment, the outer protection layer that protects the flood defence from erosion through waves and currents. Examples of maintenance are repair of drought cracks, resowing grass revetments, repair works on pattern-placed revetments, and repair of damage from animal burrowing. Many of such damages are found to have a significant impact on flood defence safety [152].

The most important method for detecting such damage is visual inspection, which is typically carried out at different times throughout the year. In such an inspection flood defence inspectors walk or drive along a flood defence and register all relevant anomalies and defects resulting in a condition report. This condition report is then used as basis for maintenance planning. The International Levee Handbook (ILH) lists a variety of inspection types [3], most notably: general inspections to determine the condition of a flood defence and/or whether maintenance works have been conducted properly, inspections before, during or after flood conditions, and special inspections aimed at detection of a specific type of damage (e.g., drought cracks). In this chapter we focus on general condition inspections, inspections during flood conditions are discussed in a.o. Lendering et al. [153].

Currently most risk-based assessments of flood defence safety assume that the flood defence is in good condition, and do not include the possibility of undetected defects and their potential effect on flood defence safety. While for instance the International Levee Handbook does note that inspections are not perfect, this is not translated to consequences for flood risk assessments [3]. Assuming that existing (visual) inspection policies give a complete overview of defects will thus lead to an underestimation of actual flood risk. For instance, if the erosion resistance of a grass revetment is overestimated due to undetected damage, the actual flood risk might be higher than is estimated in typical flood risk assessments. As, based on the literature, it is likely that at least a part of the damaged areas remains undetected for some time, including such factors in risk assessments will improve risk estimates. Insight in the accuracy of visual flood defence inspection, and identifying factors that cause damages to remain undetected, can aid in defining targeted actions to improve inspection quality. Examples are improving task definitions, targeted training of inspectors, and improvement of inspection guidelines [154]. Such insights can help improve both assessment of existing and prediction of future performance [e.g., 155, 156]. Obtaining estimates of flood defence inspection accuracy can therefore provide a basis for further improvement of such inspections, improve flood risk estimates, and improve flood protection performance.

This chapter presents the results of a field experiment conducted in March 2020. In this experiment 4 different flood defence sections along the Dutch Rhine river were inspected 14 times by 22 different inspectors. The goal was to answer three main questions with regards to the quality of visual condition inspections of flood defences:

1. What percentage of defects is detected in a typical condition inspection?

2. What is the consistency with which defects are classified?

3. Can we identify influential factors that impact inspector performance?

Answers to these questions can be used to identify possible improvements for inspection of flood defences. Sections 6.1 and 6.2 present relevant background on past research on inspection accuracy, and the current practice of flood defence inspections, both internationally and in the Netherlands. Section 6.3 presents the methods and setup of the field test, as well as a description of the field test location. Results of the field test are presented in Section 6.4, after which Section 6.5 provides a discussion on findings. Conclusions are summarised in Section 6.6.

6.1. Past research on inspection accuracy

From past research on inspections in other applications, it is found that the detection rate of visual inspection displays significant variation among different applications and inspection types. In general terms, Drury & Fox [157] report an error rate of 20-30 %, but these values vary significantly among different applications and specific situations. For instance Graybeal et al. [158] carried out a field test to investigate the performance of highway bridge inspectors. This test included 49 inspectors who fulfilled 10 different inspection tasks at 7 different bridges. For general condition assessment it was found that 68% of condition ratings varies within 1 point from the assigned reference rating (10-point scale). Several potentially important variables were identified, such as visual acuity, the extent to which inspectors were rushed, and the perceptions of aspects such as the complexity of the structure and worker safety during inspection. Aside from routine inspections also in-depth visual inspections were carried out. It was found that these inspections were not likely to detect the types of defects such inspections are aimed at. Also a large variation in detection rates was observed, ranging from a detection rate of approximately 4% for some weld cracks to 100 % for defects to the paint system [159]. Again inspectors who spent more time, were more comfortable during inspection, and perceived the structures to be more complex, performed better overall. However, what stands out is the large variability in defect detection rates in such general in-depth inspections.

Spencer [160] and Drury & Spencer [161] performed similar investigations to determine the accuracy of crack detection for airplanes. Here a detection rate of 68 % was found. This is significantly higher than the weld crack detection in the research reported by Graybeal et al. [158]. Although obviously the type of cracks and type of structure differs, this might also be explained by the fact that the complexity of objects correlates with lower inspection performance [162]. Additionally Harris [162] found that giving more time for inspection of a complex structure does not increase the detection rate, suggesting that there is an upper limit for a given inspection type. Another explanatory factor might be that the number of fault types to be considered, and thus the complexity of the inspection task itself, strongly correlates with the error rate in inspections [163, 164]. This is supported by findings from research on the accuracy of visual inspection of sewer systems [165, 166]. Here a clear relation between the number of False Negatives and the complexity of

the used coding systems was found, with more complexity leading to lower inspection performance [166]. These examples emphasise the variety of variables that influence inspector performance, of which See [167] provides a structured overview. Here distinction is made between a.o. variables related to the formulation and scope of the task, the individual characteristics, and the environmental conditions and social circumstances in which inspections have to be carried out. An important example of the latter was identified by Wiener [168], who found that both complex procedures for rejection and peer pressure to accept products, led to an increase in flinching, resulting in defect products being accepted.

6.2. Practice of flood defence inspections

6.2.1. Approaches for flood defence inspections in different countries

Crespo Márquez [169] defines inspection as a 'check for conformity by measuring, observing, testing and gauging the relevant characteristics of an item'. For the process of translating inspection findings to maintenance actions, Bakkenist et al. [170] distinguish 4 steps:

- Observation: observing a defect, anomaly or condition.
- Diagnosis: assessing the nature and type of a defect or the condition, as well as the severity, based on relevant (predefined) characteristics [169].
- Prognosis: assessing whether the severity of the defect or the general condition will change in the future.
- Operationalization: defining appropriate actions to deal with the observed defect, such as repair, overhaul or doing nothing [169].

In many countries the basic principles for flood defence inspection are in line with those outlined in the International Levee Handbook [3]. Generally inspections focus on the first three steps, and always at least combine observations and diagnosis: defects or conditions are classified using different parameters and severity classes. In some cases inspections are aimed at observing and diagnosing defects, in some cases at observing the general condition of a structure. Inspections aimed at observing defects (e.g., animal burrows, rutting and corrosion) are our main focus here. Inspections at assigning condition ratings to flood defence sections are not further considered here, although it should be noted that such ratings are sometimes achieved by translating observed defects into condition ratings [3]. E.g., the US Army Corps of Engineers translates ratings for 125 specific items considered in the inspection to 'acceptable', 'minimally acceptable' and 'unacceptable' ratings.

In the UK a similar approach is used, but additionally the condition grade (scale 1-5) established from visual and other types of inspection is used to predict remaining life using deterioration curves [171]. As such, the condition grades are directly translated into a prognosis of the future condition. However, defects that determine

the condition grade might often be caused by shock-based processes rather than continuous degradation processes [13]. It was demonstrated in Klerk & Adhi [172] that the condition of a certain flood defence section can vary significantly over time, indicating that the use of standardised deterioration curves might not correctly reflect the actual degradation behaviour.

Several factors for high quality flood defence inspections are mentioned in literature. Specific focus is often on training, and in the UK new inspectors first have to gain in-field experience under supervision of a more experienced inspector. Compared to the factors mentioned by See [167], being able to evaluate flood risks based on an understanding of the failure mechanisms, experience with inspections and computer literacy are some of the factors mentioned that ensure consistent, efficient and thorough inspections [3]. Long et al. [173] describe a blueprint of an 'ideal condition indexing' process. This mostly concerns more extensive use of information from other sources, adding other types of measurements, and increasing the range of condition values to enable greater gradation of asset condition. While such efforts can indeed lead to a better overall estimate of flood defence condition, it is doubtful whether the condition estimates from visual inspection itself would improve, as the task complexity will increase, which typically results in lower defect detection and less consistency in classification of defects as was shown by a.o. van der Steen et al. [166], Gallwey & Drury [163] and Dalton & Drury [164].

6.2.2. Routine condition inspections of flood defences in the Netherlands

As the field test reported here was carried out using the general approach used in the Netherlands, this is described in more detail in this section. The main focus of the field test was to mimic a spring inspection, usually carried out in March, after the winter season, i.e. the period between October and March during which most storms and flood waves occur. The goal of the spring inspection is to identify defects and anomalies at all dike sections such that, if required, repair or overhaul works can be carried out before the next winter season. While other inspections are also of importance, the spring inspection is the backbone of maintaining the overall condition, as most of the repair and maintenance works are based in the spring inspection results.

Spring inspections are typically carried out using the Digigids guideline [174]. The Digigids is a comprehensive guideline with many different types of damage for different types of flood defence elements. Inspectors have to classify defects/damages in three variables: the flood defence element (e.g., grass revetment), the damage parameter (e.g., animal burrowing or bare spots) and the severity, which is a classification on a 4 point scale: good, reasonable, mediocre and bad. The Digigids provides descriptions for each category, as well as reference photos of damages. Figure 6.1 illustrates this for bare spots. The definitions for severity are not explicitly related to failure behaviour or failure mechanisms, although these can be related to the sod quality of the grass revetment, which is an important input parameter in reliability assessments [172]. In principle however, the Digigids is aimed

(a)
Good: no bare spots.

(b) *Reasonable: at most 5 spots with a diameter <0.2 m where vegetation is gone.*

(c) *Mediocre: at most 5 spots with 0.2 m < diameter < 0.3 m where vegetation is gone.*

(d) *Bad: > 6 spots with diameter > 0.2 m, or 1 spot with diameter > 0.3 m where vegetation is gone.*

Figure 6.1.: *Example of Digigids classification for bare spots. panes a-d show increasing severity (good, reasonable, mediocre, bad). Captions for subfigures give description of category. All descriptions apply to an area of 25 square meters. All figures originate from Het Waterschapshuis [174].*

at inspecting the condition of the revetment and not at assessing the risk of failure.

Although the Digigids does facilitate registering using severity estimation 'good', in practice this is not done and only points with severity 'reasonable' or worse are registered. For spring inspections some prioritization is made in terms of the parameters to be inspected. For instance, most water authorities do not register flotsam on slopes, and burrowing by moles and mice is typically also not registered as it is dealt with in routine maintenance.

Despite some differences in rating systems, interpretation of results, and specific prioritizations, the approaches towards inspections are fairly similar in other countries. In most countries (e.g., France, UK, USA, and Ireland) also a system of 3 to 5 condition grades is used for diagnosis of the severity [3]. Additionally sim-

ilar types of defects are considered, for instance: unwanted (woody) vegetation, bare spots in the grass cover, deformations, erosion, cracks and animal burrows. Inspections are typically carried out by 2 inspectors, both to ensure worker safety and to ensure completeness of the data. In most countries registrations are made including photographs and GPS coordinates. Data is reported to the flood defence asset manager and stored for future analysis.

6.3. Methods

6.3.1. Quantifying the accuracy of inspection

Results from an inspection can be classified in 4 different categories [175]:

- True Positive (TP): a defect exists and is detected.
- False Positive (FP): a defect does not exist, but is detected.
- True Negative (TN): a defect does not exist, and is not detected.
- False Negative (FN): a defect exists, but is not detected.

The effectiveness of non-destructive evaluation techniques such as visual inspection is typically quantified using the Probability of Detection (PoD). The PoD can be computed using:

$$\text{PoD} = \frac{TP}{TP + FN}. \tag{6.1}$$

It should be noted that in some fields (e.g. pattern recognition) this parameter is named 'recall'. The other way around, the probability that a registered defect does not exist can be quantified using the Probability of False Alarm (PFA):

$$PFA = \frac{FP}{TN + FP}. \tag{6.2}$$

As hardly any false positives are registered in the field test, the PFA will not be considered any further in the analysis of the results.

Another important point is the distinction between classification and detection errors. Practically, any defect that has not been registered can be considered a False Negative. However, in some cases an inspector might detect a damage (with for instance severity 'mediocre'), but classify its severity as good. As points with severity 'good' are not registered in spring inspections, such a case might be incorrectly judged as a detection error, while it is in fact a classification error. Hence, the data would suggest it is a False Negative, whereas it is in fact a True Positive, with an error in the classification of severity. Therefore all PoD-values computed for the field test are lower limits. It has to be noted that the practical effect of such classification errors is that the damage remains unknown to the asset manager.

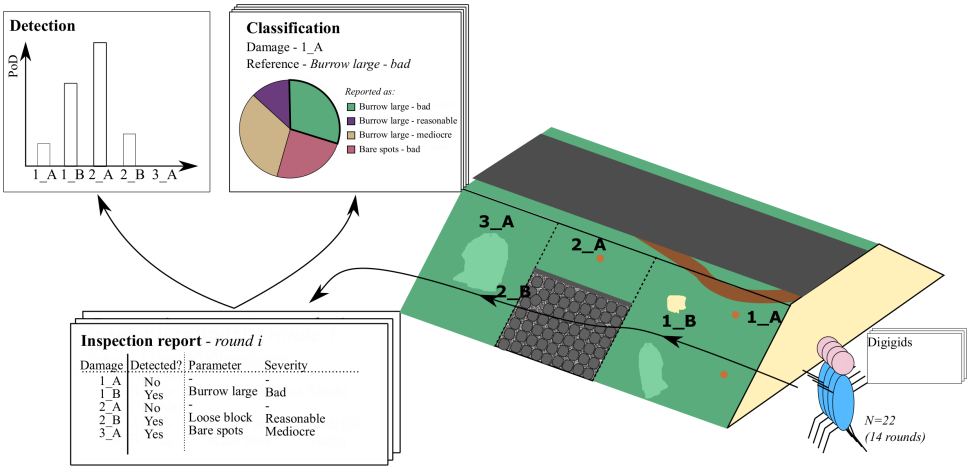


Figure 6.2.: *General approach to the field test. Inspectors inspect several dike sections as by their normal practice. This results in an inspection report for each inspection round where for each damage present it is indicated whether it was detected and how it was classified. This enables analysis of the detection accuracy and classification consistency based on the predefined reference situation.*

6

6.3.2. General set up of field test

The goal of the field test was to answer the three main research questions outlined in the introduction. Therefore it was set up to mimic the actual spring inspection for flood defences using the typical guidelines in the Netherlands as closely as possible. Figure 6.2 shows the general setup of the test. Different dike sections (3 are displayed here) were inspected by different (teams of) inspectors. During the field test inspectors registered defects on their smartphone or tablet. This was done using a cloned version of the ESRI Survey123 application they normally use during inspections, in order to avoid issues in registering defects. Inspectors always had to register the coordinates, parameter, severity, dimensions, other remarks, and add a detail and overview picture of the defect. The database from these registrations enables analysis of the detection accuracy, as well as the consistency of classification using parameter and severity in accordance with the Digigids. To enable the analysis of detection accuracy and classification consistency all considered dike sections were pre-inspected in order to map all defects and determine the reference classifications.

To facilitate the analysis of influential variables, several questionnaires were filled in by the inspectors at different times. Before the test, an extensive questionnaire on a.o. several personal characteristics, training, experience and their common inspection approach was filled in by the participants. Additionally throughout the experiment inspectors were inquired about amongst others their feeling during the day, their experiences during the test and whether they thought the field test was representative for their normal inspections. All questions to individual inspectors

have been listed in Table C.1 in Appendix C. Answers to questions have several categories: numerical values, yes/no answers, multiple choice questions and questions on a 1 to 5 rating (very bad-very good or disagree-agree). In the analysis of influential variables only those of which an effect on inspection performance could be expected have been included. All considered variables including the questions they originate from are listed in Table C.2 in Appendix C.1 and will be further discussed in Section 6.4.3. Before and after inspection rounds, the inspectors involved were inquired (as a pair) about the difficulty of that round, and whether they experienced any time pressure. This was only the case for a few inspection rounds, and these questions have not been analysed further.

During the test a supervisor was present at each dike section. These supervisors posed questions before and after inspection rounds, gave participants general instructions and observed the general behaviour of participants during the test. Supervisors were given smartphone applications to log important events during the inspection, and general remarks on participant behaviour. Examples are people passing by, the walking routes of inspectors, and remarks about the collaboration between participants.

Supervisors ensured that inspection rounds lasted no longer than 25 minutes. This time frame was determined based on the typical time used for spring inspections and was chosen such that no additional time pressure was imposed, as was confirmed from the questions posed to the inspectors. Also, nearly all inspections were finished within the given time frame.

An important working agreement for spring inspections is that defects with severity 'good' are not registered. This means that based on the registered points, it is impossible to determine whether a damage was not detected, or whether it was detected but classified with severity 'good'. The logging of the supervisors was used to gather evidence on points that were detected but not registered in the Survey123 database. Such points where added to the database with classification 'good'. Thus we have a database consisting of two parts: the registrations in Survey123, and additional damage registrations based on the logging. These datasets contain all required information for determining the PoD for damage points and section damages, and analysing the classification consistency. It should be noted that, as the supervisor logging is not entirely complete, the computed PoD-values are still lower bounds.

6.3.3. Damage classes considered

In the analysis we consider two damage classes. The first class are the damage points, which are specific damage spots such as a specific animal burrow. As in some cases the overall damage to a flood defence section is a better measure for the state of the flood defence, we also consider section damages. If for instance a certain dike section contains 3 animal burrows, this yields 3 individual damage points and a section damage 'burrowing' that encompasses all three burrows for that specific section. Whether section damage 'burrowing' has been detected is

computed as:

$$I(\text{burrowing}) = \max(I_{\text{burrow1}}, I_{\text{burrow2}}, I_{\text{burrow3}}), \tag{6.3}$$

where $I(..)$ is an indicator function that indicates whether a damage is detected ($I(..) = 1$) or not. Logically the PoD for section damages will always be equal or larger than for the damage points it encompasses. Determining the PoD for section damages is relevant, as the maintenance of some damage types is carried out at a section level, meaning that all individual damages at a certain section will be repaired. Based on the severity of a damage, and its potential consequences for failure we also distinguish between essential and non-essential section damages and damage points. Essential damages are those that should not be missed in an inspection as these likely have a direct impact on flood defence safety. To summarise: a single damage point (e.g., a specific animal burrow) is part of the subset that forms a section damage 'burrowing' and both the damage point and main damage are categorised as (non-)essential. Note that non-essential damages could develop into essential damage over time. Whether damages are essential has been based upon an expert assessment of their potential consequences for the strength of the flood defence. A description of the damages at the different sections is given in Section 6.3.5.

6.3.4. Approach for analysis of influential variables

The influential variables based on the questionnaires are of two main types: categorical variables typically consisting of 2 or 3 categories, and variables based on questions that were answered using a 5-point scale or using numerical values (such as years of experience).

For the analysis we use Bayesian Parameter Estimation as outlined by Kruschke [176]. With this approach we update our prior assumptions on uncertain parameters using data in order to obtain posterior distributions of for instance different categories. Based on these distributions we can then assess the difference between groups, or the influence of a numerical variable. This estimation of the posterior distribution is done by generating multiple samples using Markov Chain Monte Carlo.

The advantage of Bayesian Parameter Estimation over for instance null hypothesis significance testing is that it provides richer information, for instance with regard to the influence of uncertainty on the differences between groups. As sample sizes are relatively small, this provides more insight in the influence of different variables, rather than a simple acceptation/rejection of the null hypothesis of, for instance, two groups originating from the same distribution. Nevertheless, given the relatively small sample sizes, all results should be interpreted as an exploration of what parameters might influence flood defence inspection performance, rather than statistical evidence of the importance of a certain parameter.

For categorical variables we estimate distribution parameters of a three-parameter Student-t distribution of the number of detected damage points d for both groups (e.g., whether inspectors used a tablet or smartphone for registration). For numerical parameters we use a similar approach, but here we estimate the parameters

Figure 6.3.: *Impression of the 4 sections. Photos for sections 1, 2 and 3 were taken by inspectors during the test. Photo 4 was taken by one of the supervisors. Visible damages are for instance animal burrowing (section 2) and missing joint fill (section 3).*

of a linear regression between the number of damage points detected d and a numerical variable y (e.g. inspector age) such that $d = a + b \cdot y + \epsilon$. More details on the precise formulations and prior assumptions are given in Appendix C.2.

In line with Kruschke [176] we compute the 95% Highest Density Interval (HDI) for all parameters. This is the interval that contains 95% of the posterior density of the distribution with estimated parameters. For categorical variables we primarily look at the effect size, which is defined as: $(\mu_1 - \mu_2)/\sqrt{(\sigma_1^2 + \sigma_2^2)/2}$ for groups 1 and 2. If the HDI is (almost) entirely negative or positive, this means that it is highly likely that there is a difference between the categories. For numerical variables we primarily look at the HDI for slope b. If the HDI for b is (almost) entirely positive or negative, this indicates a relation between the detected damage points d and the considered variable y. Note that we only analyse the influential variables for damage points, not for section damages.

6.3.5. Description of field test location

The field test was conducted at 4 dike sections near the city of Tiel, along the Waal river in the Netherlands. Pictures of the 4 sections are shown in Figure 6.3. The sections were chosen as these are representative for the flood defences in the considered area. For each dike section a set of reference damage points and section

damages was derived based on the pre-assessment and field test. An overview of all reference damages is provided in Table 6.1. Section 1 is approximately 185 meters long. Here the inner slope, which consists of a grass revetment, was inspected. Several small damages to the grass cover are present, both burrows and other types of damage. Additionally the slope is deformed locally. Both the presence of burrowing at a section level, and the slope deformation are essential damages in an inspection of this section. Section 2 is 200 meters long, and here the outer slope was inspected. The general shape of the grass revetment is good except for a few spots with weeds. However, there is a significant number of animal burrows, mostly caused by dogs who enlarged smaller pre-existing mice or rabbit burrows. It has to be noted that damage point 2_6 developed during the field test and could therefore only be observed in 5 of 14 rounds. This is the only damage that developed in the period the field test was executed. Although there was high water during the test, the water level did not influence the inspectability at this dike section. Section 3 is 200 meters long, and the outer slope was inspected. The lower two-third of the slope is covered with a block revetment (consisting of pattern-placed basalt), and the upper third is covered with grass. The block revetment is in a relatively bad shape: there are several loose and missing blocks, the joint fill material has washed out and there are tree trunks that penetrated the revetment and displaced the blocks. Many of the damage points at this section are therefore classified as essential. The grass at the higher part of the slope is also not in good shape. It has to be noted that at section 3 the influence of the high water level conditions influenced the outcome of the test, as it was impossible to walk on the maintenance path at the lower part of the revetment. During some test sessions the inspectability was lowered as the revetment was wet due to rainfall and therefore difficult to access. Section 4 is approximately 80 meters long, but here both inner and outer slope as well as the crest and the inlet structure had to be inspected. In general there is a significant amount of rough vegetation at this section, and the transition between revetment and structure is an important point of attention. Some parts of the outer slope contain a concrete lawn grid which is deformed in 1 location. The high water conditions had no influence on inspectability. Figure 6.4 provides an example of the results from the Survey123 database for dike section 4. Here both registrations from the pre-assessment and the field test are shown for each damage point. All registrations have been manually coupled to the registrations, based on the attached photographs, description and location.

6.3.6. Conditions during the field test

During the field test the river water levels were relatively high. Several inspectors indicated that this impacted their performance and approach, especially at dike section 3. At the 3rd of March, water levels were lower than at the 6th of March. Weather conditions differed slightly between the two dates: the 3rd of March was dry, and mostly sunny, at the 6th of March rainfall in the night before resulted in wet slopes. Especially at section 3 this had influence on the accessibility of the block revetment.

Table 6.1.: *Overview of reference damages for all sections. Descriptions of damages marked as non-essential are displayed in italics. Both damage points and corresponding section damages are given, as well as reference classification. In some cases no reference classification was given, specifically when damage points were not in the pre-assessment or if the classification was ambiguous and multiple classifications could apply.*

	Damage point		**Section damage**		**Reference classification**	
	ID	*Description*	*ID*	*Description*	*Parameter*	*Severity*
Section 1	1_1	*Large burrow 1*	1_B	Burrowing	Burrowing large	Bad
	1_2	*Small burrow 1*	1_B	Burrowing	Burrowing small	Bad
	1_3	*Weeds*	1_A	Grass cover	Weeds	Reasonable
	1_4	*Cover & bare spots*	1_A	Grass cover	Bare spots/Coverage	Mediocre
	1_5	Slope deformation	1_C	Deformation	Slope deformation	Bad
	1_6	*Small burrow 2*[a]	1_B	Burrowing	-	-
	1_7	*Small burrow 3*[a]	1_B	Burrowing	-	-
Section 2	2_1	*Burrow 1*	2_A	Burrowing	Burrowing large	Bad
	2_2	*Burrow 2*	2_A	Burrowing	Burrowing large	Bad
	2_3	*Crack*	2_B	*Grass cover*	Cracks	Mediocre
	2_4	*Burrow 4*	2_A	Burrowing	Burrowing large	Bad
	2_5	*Weeds*	2_B	*Grass cover*	Weeds	Reasonable
	2_6	*Burrow 5*[b]	2_A	Burrowing	Burrowing large	-
	2_7	*Burrow 6*	2_A	Burrowing	Burrowing large	Bad
	2_8	*Other burrows*[a]	2_A	Burrowing	Burrowing large	Bad
Section 3	3_1	*Bare spots*	*3_A*	*Grass cover*	Bare spots	Bad
	3_2	Washed out joint fill	3_B	Loss of clamping force	Joint fill washout	-
	3_3	Tree trunks	3_C	Woody vegetation	Woody vegetation	Bad
	3_4	*Displaced block*[a]	3_B	Loss of clamping force	-	-
	3_5	Loose block 1[a]	3_D	Loose or missing blocks	Loose blocks	Bad
	3_6	Loose block 2[a]	3_D	Loose or missing blocks	Loose blocks	Bad
	3_7	Missing block	3_D	Loose or missing blocks	Holes	Bad
	3_8	Loose block 3[a]	3_D	Loose or missing blocks	Loose blocks	Bad
	3_9	Loose block 4	3_D	Loose or missing blocks	Loose blocks	Bad
	3_10	*Small burrow*[a]	3_A	*Grass cover*	-	-
Section 4	4_1	*Woody vegetation 1*	4_A	Rough & woody vegetation	Woody vegetation	Bad
	4_2	*Woody vegetation 2*	4_A	Rough & woody vegetation	Rough vegetation	Bad
	4_3	Transition 1	4_B	Transition with structure	-	-
	4_4	*Transition 2*	4_B	Transition with structure	-	-
	4_5	*Lawn grid deformed*	4_C	*Grass cover*	-	-
	4_6	*Tree growth*	4_A	Rough & woody vegetation	Woody vegetation	Reasonable
	4_7	Rough slope	4_A	Rough & woody vegetation	Woody vegetation	Bad
	4_8	*Small burrow*	4_C	*Grass cover*	Burrowing small	Bad
	4_9	*Weeds*[a]	4_C	*Grass cover*	-	-

[a] This damage was not reported in the pre-assessment.
[b] This damage developed during field test

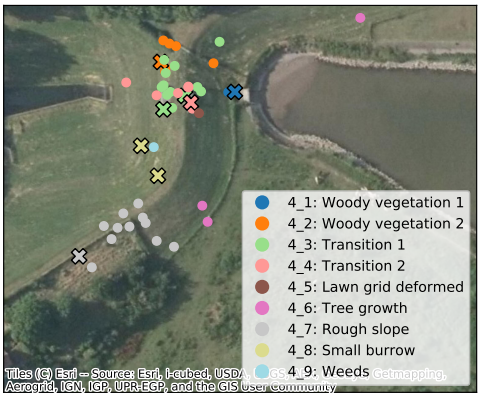


Figure 6.4.: *Results for dike section 4 obtained from the Survey123 database. Dots are registrations during the test, crosses indicate registrations during the pre-assessment.*

All inspectors received a time schedule for their inspections. Time schedules were generated based on a randomised algorithm to ensure each section was inspected once, and inspectors always inspected with a different partner. In a few occasions inspectors arrived slightly late or were rescheduled to a different time slot. Nevertheless, except for 1 inspector who did not inspect section 3, all inspectors inspected all sections, each time with a different partner.

Due to absence of 6 of the total 28 scheduled inspectors not all inspections were done by a pair of inspectors. 25 out of 56 inspection rounds were carried out by 1 inspector (7 at section 3, 6 at the other sections) and each participant inspected 0 to 2 rounds on their own. From the results it is demonstrated that this had no influence on scores at sections 1, 2 and 4, but at section 3 single inspections are found to result in lower detection rates.

An important remark with regards to the supervisors at different sections is that each used a slightly different approach to log events during the test. In some cases supervisors recorded many voice messages, others took plenty of photos. This might have influenced the number of damage registrations that were added based on the logging. Additionally, at the 6th of March inspectors were given specific instructions to pay attention to damage points that were detected but not registered in Survey123. In total 29 damage registrations were added based on the supervisor logging, of which 14 concerned the 5 inspection rounds at the 6th of March, and 15 registrations the 9 rounds on the 3rd of March.

6.4. Results

6.4.1. Accuracy of flood defence inspections

First we look at the overall Probability of Detection (PoD) for damage points and section damages. Figure 6.5 shows the PoD for all damage points per section. Grey bars indicate registrations based on the supervisor logging, and hatched bars

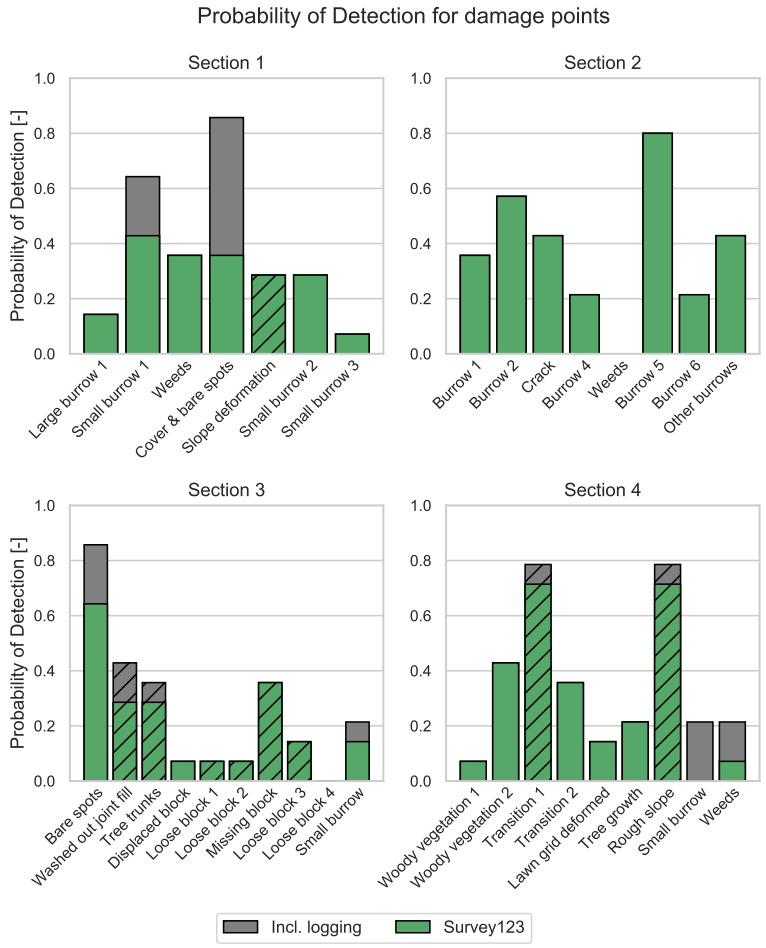


Figure 6.5.: *Probability of Detection for all damage points per section. Hatched bars denote damage points that were categorised as essential damages. Green indicates registrations in Survey123, whereas grey includes registrations added based on the supervisor logging. Note that Burrow 5 at section 2 was only present in the last field test session and was therefore only observable during 5 inspection rounds.*

denote damage points that were marked as essential in the reference. Overall, the PoD varies significantly, ranging from 0 to almost 0.9 per damage point. Most of the registrations added based on the logging concern smaller issues, such as small burrowing and the quality of the grass cover. Also note that there is no clear difference in the PoD for essential and non-essential damages, except for section 4 where the two essential damage points were registered by the majority of the inspectors.

The PoD is lowest for issues with the block revetment at section 3. Here the loose and missing blocks have been sparsely or not detected during the test. There are several explanations for this: first of all, the high water conditions hampered inspectors (which was also indicated in the questionnaires). Secondly, it should be noted that only damage point 3_9 (Loose block 4) was detected during the pre-assessment, when conditions were much more favourable. A second explanation is therefore that inspecting block revetments is generally more difficult as it is harder to see and process all the details. This especially holds for blocks that are loose, but still in their place. This could also explain why the missing block was detected more often, despite its proximity to some of the loose blocks.

At sections 1 and 2, there were multiple burrows for which the PoD varies significantly. The question is whether the main reason is failure to detect, as inspectors indicated that in many cases they do not register all the burrows at a section. This reduces the work load during inspection, and common maintenance works to deal with it will be done at a section level anyway, so all burrows will be repaired together.

From that perspective, it is more relevant to look at burrowing as a section damage. Figure 6.6 shows the PoD for all section damages. By definition these are higher than for individual points (see Equation (6.3)). It can be seen that at section 2 most inspectors (PoD=0.86) registered at least 1 burrow.

Nearly half of the inspection rounds was done by a single inspector instead of a pair. For most damage points and section damages this caused no major differences in the estimated PoD, except for section damages 3_B (loss of clamping force) and 3_D (Loose or missing block), and the corresponding damage points. For these damages individual inspectors scored much lower: for 3_B the PoD for a pair and individual were found to be 0.86 versus 0.14, for 3_D the PoD was 0.71 versus 0.29. Figures for all damage points and section damages for individuals/pairs of inspectors can be found in Appendix C.3.

An important aspect in the context of risk-based inspection is not only the average PoD, but also the variation among different inspectors. Figure 6.7 shows the variation among inspectors for both damage points and section damages, and subsets of the essential damages. For damage points, the PoD ranges between approximately 0.25 and 0.55, with a bit more variation for the essential damages. Note that this is strongly influenced by the low detection percentages of the various loose blocks at section 3. The PoD for section damages ranges between 0.5 and 0.9. The variation is similar to that of damage points, but the average is significantly higher.

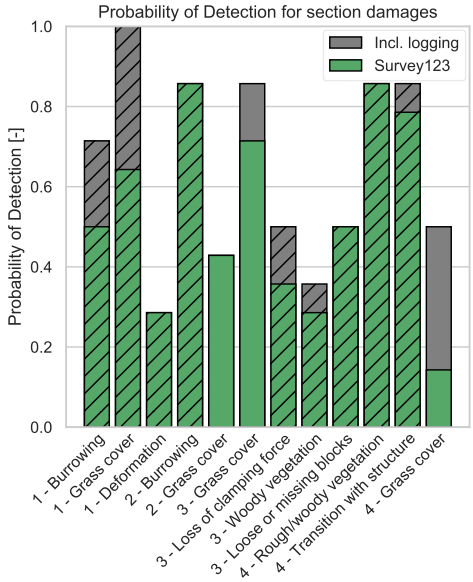


Figure 6.6.: *Probability of Detection for all section damages (numbers at horizontal axis denote the section). Hatched bars denote section damages that were categorised as essential. Green indicates registrations by inspectors in Survey123, whereas grey includes registrations added based on the supervisor logging.*

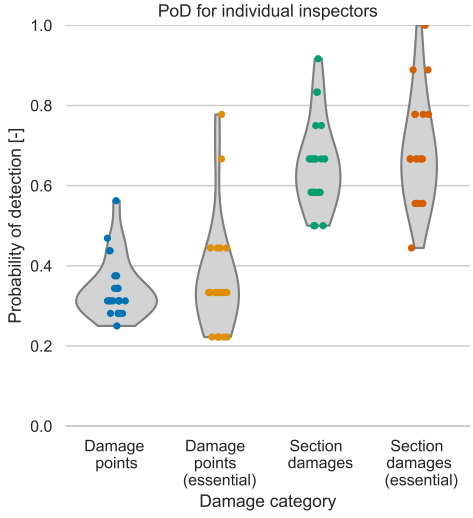


Figure 6.7.: *Probability of Detection (PoD) for individual inspectors. Grey shaded areas provide a density estimate capped at the highest and lowest PoD encountered in the test. Colored dots provide results for individual inspectors for (essential) section damages and damage points.*

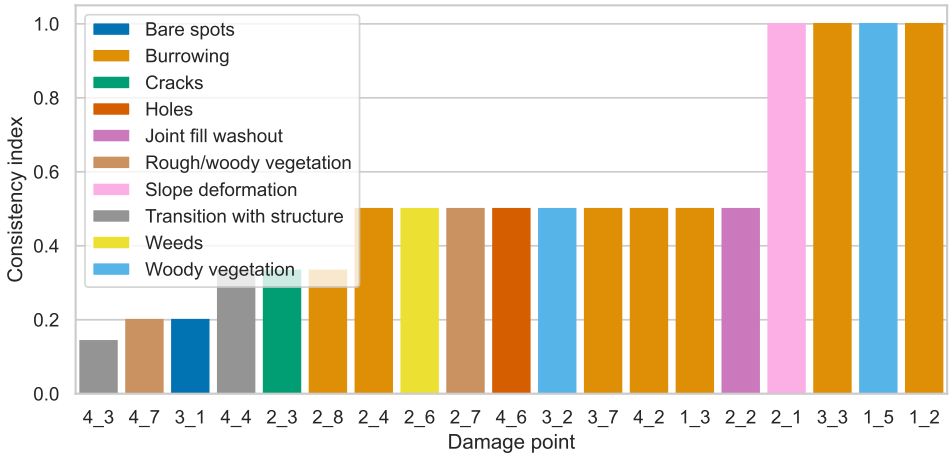


Figure 6.8.: *Consistency index of different damage points with more than 3 registrations in the database. Colors indicate the reference classification parameter (see Table 6.1) except for 'Transition' where no reference parameter could be determined, and burrowing where small and large burrows are combined.*

6

6.4.2. Consistency of classification of flood defence defects

Based on the database of inspection registrations we can determine how consistently damages were classified in the database. By consistency we mean the agreement between inspection reports of different inspectors. This is determined both by whether the damage was detected, and what parameter and severity they assigned. Theoretically, each damage point should have 1 correct parameter, although in some cases multiple parameters of the Digigids might be applicable. For instance, in many places overgrown vegetation consists of both weeds, woody vegetation and generally rough vegetation. Henceforth, when looking at consistency between parameters we do not look at how many inspectors chose the 'correct' parameter, but rather at how many different parameters were used for the same damage point. To that end, we define a consistency index:

$$C = \frac{(N/N_{\mathrm{par}})}{N}, \tag{6.4}$$

where N is the number of records in the database, and N_{par} denotes the number of unique parameters (e.g., weeds, small burrowing, large burrowing) in the damage registrations by inspectors. Higher values for C mean that inspectors were more consistent in their parameter choice. Figure 6.8 shows the consistency index for all damage points with $N > 3$. Damage point 1_4 was not included as this encompasses 2 damage parameters (bare spots and cover). It is found that especially for the damage to the transition (4_3 and 4_4) the registrations are very inconsistent: 4_3 was registered 11 times with 7 different parameters, 4_4 5 times with 3 different parameters. This indicates that there is no clear parameter to define

Figure 6.9.: *Example of the presence of multiple damage parameters at a single damage point (4_3). At least bare spots, weeds and rough vegetation could be applicable here, while the cause of the bare spot (just right of the reference marker) could be burrowing.*

damage to a transition. Note that in this case many of the parameters used by the inspectors are actually visible, as is shown in Figure 6.9. Here it can be clearly seen that there is rough vegetation, weeds, a bare spot, and possibly also burrowing. All these parameters were used by different inspectors and are at least partially representative for the situation.

For burrowing the consistency is 0.5 in many cases, as inspectors used both small and large burrowing as parameters. This indicates that distinguishing these in practice is difficult. Other cases with low consistency (e.g. 2_3, 3_1 and 4_7) are also typically damage points where multiple damage types are present. In practice asset managers always use the photographs to review the actual situation (or do a field visit) before deciding what maintenance is to be done. This procedure likely ensures that, even though the parameters might be inconsistent or incorrect, at least the correct maintenance action is taken.

However, it has to be noted that the reported severity does play an important role in an asset managers decision to review a registered damage, as they mainly focus on reviewing with severity 'bad'. Figure 6.10 shows the fraction of records that was categorised in the different severity categories, including the fraction of inspections where a damage was not registered. The latter consists of damage points where inspectors failed to detect a damage, or where they found it not severe enough to register. For damage points with severity 'bad', more than half of the inspectors that registered such a damage classified it as less severe. For damage points with reference mediocre or reasonable, there is also a large variation in reported severity. It should be noted that this figure is based on the entire database, including points added based on the supervisor logging. If the same figure is made just for the points registered in Survey123, the fraction classified as 'good' is 0 in most cases, which

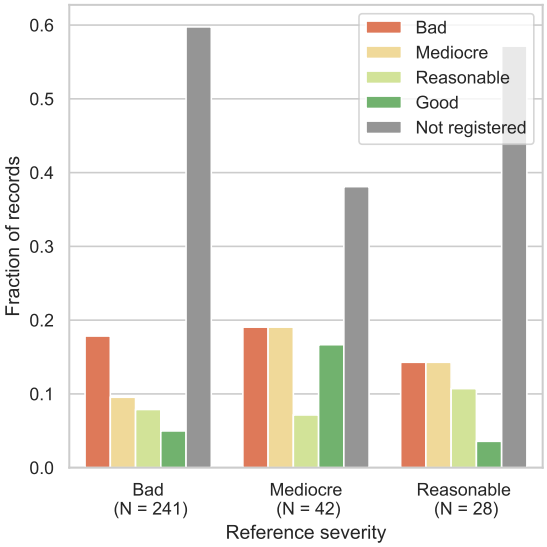


Figure 6.10.: *Fraction of records in the Survey123 database for different severity classifications compared to the reference. N denotes the number of possible registrations for each category.*

emphasises that the grey bars contain many detected damage points classified as good.

6.4.3. Influential variables for flood defence inspection accuracy

As was mentioned in Section 6.3, the overall number of samples in this test limits the extent to which conclusions can be drawn with regard to influential variables. However, a Bayesian Parameter Estimation can give some directions for future research. In general we have two main types of variables: categorical variables that are split into 2 to 4 categories, and numerical variables that are either numerical variables (such as age) or answers given on a 1 to 5 scale. The majority of the analysed variables are based on the questionnaires in Table C.1. All variables are listed in Table C.2, including the origin of the data (question or other data) and whether these have been included in the analysis or not. The influence of variables on inspection performance was estimated using Bayesian Parameter Estimation as described in Section 6.3.4.

Figure 6.11 displays the Highest Density Intervals (HDI) for all numerical variables (green), and all categorical variables (orange) where each group consists of 5 or more participants. The HDI indicates that 95% of the probability density of the posterior distribution falls within this range (grey bar). For numerical variables the HDI for the slope of the regression line is shown. A negative slope means that higher

values for the numerical variable relate to less detected damage points. To compare variables with different ranges all values have been rescaled to a 1 to 5 scale based on the minimal and maximal values in the dataset. For categorical variables the HDI for the effect size is given. If the value is positive it means that the listed category performs better. E.g., inspectors who are also asset manager perform slightly better than those who are not. If the HDI is (almost) entirely negative or positive, this indicates that a variable has effect on inspection performance (number of damage points detected). Additionally the coloured lines (orange for categorical, green for numerical variables) indicate the mean and interval $\mu \pm \sigma$.

The variables are grouped in 3 categories. Experience & training variables are often related: in practice asset managers are involved in maintenance planning and execution, and both daily and emergency inspection. All these variables are found to relate to a somewhat better inspection performance, although in none of the cases the HDI is entirely positive or negative. Logically, more experience in years also relates to better performance. For personal characteristics no clear relations are found. For the test circumstances and approach, it is found that the (negative) effect of inspecting solo at section 3 is relatively large.

6.5. Discussion

This chapter presents a field test where the accuracy and consistency of visual flood defence inspections is investigated. The goal of this field test is to answer three main questions.

The first main question concerns the Probability of Detection (PoD) of flood defence inspections. The field test shows that there is a large variation in the PoD, both between inspectors as well as between different (types of) damage points. The variation between inspectors cannot be explained by the parameters elicited in the various questionnaires, and is likely due to the nature of visual inspection (of flood defences) itself, and the general method used in the field test. This is in line with findings from literature on infrastructure inspections [e.g. 158, 165], where also large variations between inspectors and damage points was found. It should be noted that damage points with severity 'good' were not registered: therefore some of the non-detections might have been detected but classified too leniently. However, most of the cases where this was observed concerned minor damages, and there is a general agreement among inspectors that essential damages (e.g., loose blocks and slope deformations) should be registered.

The results clearly indicate that the block revetment at section 3 was more difficult to inspect. Although circumstances during the field test were difficult due to the high water levels, the fact that multiple damage points were also not detected in the pre-assessment emphasises the difficulty of detecting flaws in block revetments. For other types of damage points, such as animal burrows, the PoD varies significantly per damage point. In some cases this is due to the method of registration: for instance at section 2, there is a large number of burrows and it is time-consuming for inspectors to register all individual burrows. Inspectors indicated that they often register only 1 or 2 burrows, as the commonly applied maintenance method will

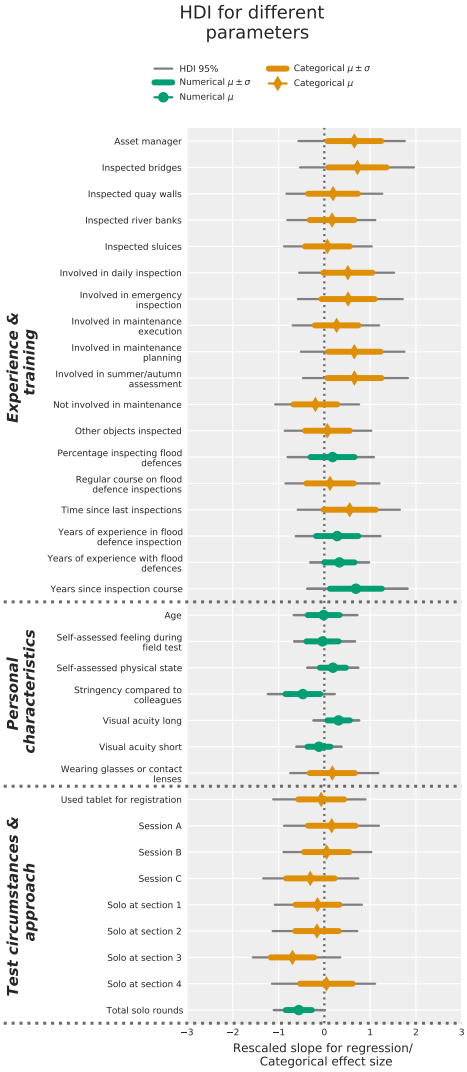


Figure 6.11.: *Results of Bayesian Parameter Estimation for categorical (orange) and numerical (green) variables. For categorical variables the effect size is given, for numerical variables the estimated slope of the linear regression. Diamonds and circles indicate mean values, colored lines indicate the interval $\mu \pm \sigma$. Grey lines indicate the 95% Highest Density Interval. Note that the slope has been rescaled to the interval 1-5 to make numerical variables with different ranges comparable.*

ensure that all burrows over a longer section are repaired. This might explain the variance between the PoD for different burrows at this section. It should be noted that in this particular case a more consistent method of registering damage would be to assess the number of burrows at a section level, rather than at individual points. To improve consistency it is therefore recommended to align the spatial level (section or point) of the damage registration with that of the commonly applied maintenance measure.

The second goal of this field test was to investigate the consistency of damage registrations. Here it is found that the parameter registrations for damage are generally quite consistent except for transitions, where a parameter is lacking, and for animal burrowing where the distinction between small and large burrowing is hard to make in practice. As it is found from overflow tests [177, 178] that transitions between structures are often places where erosion initiates, a specific parameter for transitions between revetment types and/or structures will be a valuable addition for risk-based inspections. In line with the aforementioned variation in PoD for animal burrowing, it is questionable whether distinguishing burrows with two parameters adds any value. In general the number of parameters in the applied inspection guidelines is large, while literature shows that inspections with less parameters are generally more reliable [see e.g. 163–166]. Given the considerable overlap between parameters, a valuable improvement towards improving both accuracy and consistency of inspections would be to reduce the number of parameters in the inspection guidelines. However, it has to be noted that, as asset managers typically check the severe damage points themselves, in many cases a suboptimal parameter choice will not have a major impact on maintenance. Asset managers do however consider the reported severity in prioritizing damage points for maintenance. In that sense, the inconsistency in severity encountered in the field test is more worrisome: if we consider damage points with reference severity 'bad', only 18% of the inspectors registers such a point as 'bad', while approximately 22 % registers such points as mediocre or reasonable, and 60% of the inspectors fails to register the point at all (either due to underestimating the severity, or failure to detect the damage point at all). Practically such inconsistencies lead to inadequate maintenance.

The third goal of this field test is to identify potentially important factors for higher or lower inspection accuracy. Due to the relatively limited sample size (22 inspectors) drawing major conclusions on this is not possible. Using Bayesian Parameter Estimation (BPE) we identified some parameters that might have influence on inspection performance. In general, most of the factors that relate to the work of flood defence asset managers relate to higher inspection performance. A possible explanation is that such inspectors have more practical experience with inspection and maintenance as a whole, which enables them to better assess different types of damage points and their potential consequences. A major finding from the BPE is that there was a significant difference between inspectors who inspected section 3 as a pair, versus those who had to do it alone. Initially all inspection rounds were planned to be done by pairs of inspectors, but as some inspectors were not present during the field test it was decided to maximise the number of inspection rounds for each section, rather than reschedule such that there would only be pair

inspections (but less inspections overall). For future tests it would be better to have all inspections done by the same number of inspectors to ensure consistency of the data. Additionally, from literature it is found that vigilance and tiredness can also be factors that influence inspection accuracy and consistency. Due to the relatively short inspection times, these factors could not be investigated in this experiment, but might be relevant for future tests.

6.6. Conclusions

In the field test described in this chapter 4 dike sections of 200 meters have been inspected by 22 inspectors in order to estimate the accuracy and consistency of visual flood defence inspections. Approximately half of the inspections was done by a single inspector, the other half by a pair of inspectors. The inspected sections are different, but representative for the variety of flood defences encountered in practice. Three of the sections inspected consist of grass revetments, 1 of the sections contains a block revetment. Each section contains approximately 8 damage points that should be detected by inspectors.

The Probability of Detection (PoD) differs significantly for different damage points as well as between inspectors. For different damage points the PoD in the field test ranges between 0 and 0.9. The PoD for damage to the block revetment is found to be lower, both because block revetments are generally more difficult to inspect, and as high water levels during the test reduced accessibility. For section damages, which are subsets of similar damage points at the same dike section (e.g., all burrows at a certain section), the PoD ranges between 0.3 and 1. There is significant variability between inspectors: for damage points the average PoD of different inspectors ranges between 0.25 and 0.6, while for section damages it ranges between 0.45 and 1. Combined with the large variation between different damage points it can be concluded that defining a single PoD for flood defence inspections is difficult.

It should be noted that the estimated PoD is a lower bound, as in some cases inspectors might have detected a damage but decided not to register a damage point, especially for smaller and less important damages. The registrations by inspectors have therefore been supplemented by observations from supervisors during the test. In future tests it is advised to try and distinguish more clearly between non-detections and non-registrations. It should however be noted that the practical implication is the same: a damaged spot remains unknown to the flood defence asset manager.

The consistency of damage registrations was evaluated based on the registered damage parameter and severity. Although there is some inconsistency in damage parameters, the encountered inconsistency in damage severity is of more importance for maintenance. For damage points with the highest severity (bad), only 18 % registered such damage as bad, while 22% registered it as less severe. 60% did not register the point at all, either due to not detecting it, or due to not classifying it as damage. As asset managers often use the reported severity for maintenance prioritisation, this can have significant influence on the effectiveness of risk-based

maintenance. Hence, improving the consistency of severity classification should be an important point of future attention.

A variety of variables that could influence inspection accuracy has been investigated using Bayesian Parameter Estimation. Some indication is found that inspectors who are also asset manager, participate in other types of inspections, and are involved in maintenance tasks, perform slightly better. The fact that none of the investigated variables explains the large variability indicates that variability originates from other sources. Some likely ones are the structure of the currently used inspection guidelines, and general variability among different dike sections. Based on literature [164, 166, e.g.], it is likely that simplifying inspection guidelines and tasks will lead to more consistent and more accurate inspections. Concrete improvements based on this test are to reduce the number of (sometimes overlapping) parameters, introducing a consistent parameter for transitions between structures and different revetments, and carrying out specific inspections for damage types with potentially high consequences, such as damage to block revetments or large animal burrowing. Other types of flood defence inspections (e.g., after high water) can likely be improved in a similar way, although their current PoD might be different.

In general terms, the average PoD found from this field test is in line with the PoD reported in literature on other types of infrastructure inspections such as sewers and bridges. Given the stringent reliability requirements for flood defences in most countries, it is doubtful whether it is sufficiently high to ensure these requirements are met. This is something that should be further considered jointly with knowledge on the influence of damage on flood defence reliability. Irrespective of the precise influence of damage on reliability, improving inspection accuracy and consistency leads to better maintenance planning, and is thus likely an effective means to decrease overall flood risk.

7

Risk-based inspection and maintenance of flood defence segments

Reasonable men "satisfice" much more than they "optimize"

Raiffa & Schlaifer [47]

One of the key topics of this thesis is to investigate how uncertain degradation processes should be accounted for in Inspection & Maintenance (I&M). In the previous chapter, it was found that the damages as a result of these processes are often not detected by visual inspection, and that such damages might occur randomly in time. As a consequence there is a relatively high probability that the flood defence is in a damaged state at the start of an extreme high water situation. This chapter presents a model that can quantify the influence of damage and Inspection & Maintenance (I&M) on the reliability of a riverine flood defence system. We then use this model to evaluate the maintenance concept, which is the policy for I&M, and compare this to the reliability that would be obtained from a safety assessment. Additionally we investigate how the system reliability can be improved through different interventions, both different maintenance concepts and structural upgrades. By analysing their influence on structural robustness and costs we determine what interventions can be done to ensure that requirements are met. Based on this analysis we also quantify what conditions new inspection methods and schedules should satisfy to outperform existing methods.

Inspection & Maintenance (I&M) of flood defences is aimed at detecting and resolving damage caused by a variety of different degradation processes. The policy and rules for I&M are prescribed by the maintenance concept [179, 180]. In the preceding chapter we have seen that inspections of flood defences are imperfect. As such, depending on the specific properties of the system (e.g., geometry and degradation rate), and the applied maintenance concept, there is a probability that a flood defence segment has (weaker) damaged spots. Such damaged spots can negatively influence the system reliability. In this chapter we quantify this influence, in order to achieve our main aim of assessing the effect of mitigating interventions (structural upgrades and changes to the maintenance concept). To evaluate interventions we use a multi-objective approach based on Equivalent Annual Cost and structural robustness. In this context, structural robustness is an indicator for the relative contribution of damaged spots to flood risk. We answer two main questions:

1. How can we assess whether the commonly applied maintenance concept is sufficient to meet the reliability requirement for the considered flood defence segment?

2. What are the most efficient interventions to improve both system reliability and structural robustness of the considered flood defence segment?

As a case study we consider a riverine flood defence segment with a large fetch and the possibility of relatively large waves hitting the slope. We focus on regular I&M and do not specifically look into damages that might occur related to high water situations (e.g., burrowing due to animals seeking dry ground) and do not consider emergency measures. As we consider a riverine flood defence, waves are considered independent from the discharge-dominated water level. Damage can reduce the reliability of the flood defence. We focus on failure due to external erosion at the inner and outer slope. It should be noted that in literature [e.g. 28] also the influence of animal burrowing on the phreatic level, and internal erosion have been considered and found to be of importance for amongst others slope instability. However, given the characteristics of the case study we consider, burrows of the size considered in this literature are unlikely to occur at the case study location and are not considered in this chapter. The approach is developed such that it can be extended with other failure modes as well.

Section 7.1 presents background on degradation, how damage impacts the reliability of flood defences, and some practical challenges in I&M of flood defences. The general approach of this chapter is outlined in Section 7.2, where we discuss how a Dynamic Bayesian Network is utilized to evaluate the impact of degradation and I&M on the condition of the different sections that are part of the larger flood defence segment. By combining this with conditional failure probabilities for external erosion conditional on different types of damage we can obtain the failure probability of the segment in time. Section 7.3 presents the case study and results for a baseline maintenance concept. In Section 7.4 we use the developed model to evaluate a variety of interventions and their impact on robustness, cost and flood risk. Section 7.5 present discussion, conclusions and recommendations.

7.1. Background

7.1.1. Influence of degradation on flood defence condition

Degradation processes can result in a reduction of structural reliability over time. Typically two types of degradation processes are distinguished: shock-based and continuous degradation [13], but also combinations might occur. For flood defences there is a multitude of different degradation processes [3]. Past studies in optimization of flood defence decisions have considered degradation through settlement of the crest as most important time-variant parameter [e.g. 181]. Buijs et al. [62] investigated the temporal reliability of flood defences while accounting for different degradation processes, namely settlement of the crest, cumulative effects in piping erosion (i.e., seepage length reduction) and damage to vegetation due to traffic. It was assumed for all damage types that these were governed by different degradation processes — a gamma process for settlement [similar to 182], but also hierarchical models for traffic damage and seepage length reduction where damage arrives in shocks governed by a Poisson process. Chen & Mehrabani [86] also considered degradation of (coastal) flood defences, where degradation was modelled using the condition grades proposed by Environment Agency [183]. These condition grades correspond to observed conditions to the flood defence state — condition grade 1 represents a hardly damaged flood defence with an intact surface protection, whereas condition grade 5 represents a flood defence with major rutting, bad surface protection and major (fox) holes. It is presumed by Environment Agency [183] that such degradation of the condition is gradual in time and takes several decades — the curves suggest that even with fast degradation condition grade 5 will not be reached within 30 years after construction. As such, Chen & Mehrabani [86] use a sequential Markov chain to model degradation, rather than a progressive Markov chain,[1] thus eliminating the probability of large shocks, e.g. the revetment condition degrading from grade 1 to 5 in a short instance. This can be an accurate approach for some types of degradation, such as settlement of the crest, but it might not be in line with the behaviour associated with for instance animal burrowing. Klerk & Adhi [172] found that the probability that a grass slope with a closed sod degrades to a fragmented sod is larger than that it degrades to an open sod, where fragmented is the worst state considered. As such, degradation of grass revetments is likely better described using a progressive than a sequential Markov chain. It should be noted that experiments by Le et al. [184] indicate that older grass slopes typically become stronger rather than weaker if external factors remain the same. As such, grass revetments might not only degrade, but also become stronger in time without any intervention. We do not consider autonomous recovery here, but this emphasizes the complexity of degradation behaviour of grass revetments.

Obviously, there are factors which can increase the probability of animal burrowing or rutting at a specific location, as was illustrated in Klerk & Adhi [172] for the influence of neighbouring urban areas. Another example is that species such

[1] In a sequential Markov chain the condition can only transition from 1 to 2, 2 to 3 etc. In a progressive Markov chain all transitions are possible, also 1 to 3 and 1 to 4.

as beavers and nutria start burrowing below the waterline [185], such that flood defences that are not directly adjacent to water bodies are much less susceptible to their burrows. Buijs et al. [62] framed such factors on a more general level through the distinction between excitation (i.e., flood defence properties that initiate some kind of degradation), ancillary (i.e., properties that transform the process) and affected features (i.e., properties in the reliability model that are influenced). Unfortunately, this has not yet been translated to quantitative insight into what interplay of features drives the occurrence of different types of damage (e.g., how much does an overgrown toe increase the probability of burrowing).

7.1.2. Influence of condition on flood defence strength

Next to the occurrence of damage, the key question is to what extent it influences the overall reliability of a flood defence. The International Levee Handbook (ILH) [3] distinguishes 3 main failure modes for flood defences: external erosion, internal erosion and instability. As the scope of this chapter is external erosion, we will only briefly discuss other failure modes.

External erosion can be caused by inner slope erosion through overflow or overtopping waves, and erosion due to wave impact on the outer slope. With regards to the influence of the revetment condition, most of the research in experimental settings has focused on quantifying the strength of for instance rubble mound, grass revetments or block revetments in a good state. The resistance of a damaged revetment to withstand external erosion has not been considered but for some exceptions.

For overtopping erosion on grass revetments, van Bergeijk et al. [152] investigated the influence of damage on overtopping erosion failure probabilities and found this to increase the failure probability by (several) orders of magnitude. Similarly, Aguilar-López et al. [177] demonstrated, using a computational fluid dynamics (CFD) model, that transitions between a road and grass slope result in larger erosion rates due to more turbelent flow. Similar behaviour is found for some of the types of damage encountered on grass slopes and thus this indicates that damage can significantly increase the failure probability. For external erosion on grass-covered outer slopes, not much is known about the influence of damaged spots, but some research has been done on the influence of transitions and initial damage using a wave impact simulator [186, 187]. Again it is found that initial damage can lead to a significant reduction of the strength of the revetment, although this can vary significantly between cases. An important point is that it is found that also the underlying clay layers provide additional resistance [188], but experience from tests is that once the sandy core is exposed to wave loads, erosion can proceed very quickly.

For failure due to slope stability and internal erosion Taccari & Van Der Meij [189] and Palladino et al. [28] investigated the influence of animal burrows. They mostly considered very large burrows, and found these to be of relevance for both slope stability and internal erosion. However, modelling the effects is difficult as the influence depends strongly on their shape, location and size. For internal erosion

especially burrows at the lower part of the inner slope, and burrows with entrances at both sides of a flood defence are of relevance. For slope instability the largest influence is found for burrows that increase the hydraulic gradient in the dike body [189].

An important factor which reduces the likelihood of large animal burrowing for riverine flood defences, is the presence of large foreshores that are dry under normal conditions. At such locations, large burrows are unlikely, while there are many records of smaller burrows in field inspection data [172]. For such smaller burrows the main question is whether these extend through the typically present clay cover layer. In such cases, internal erosion might occur due to a.o. concentrated leak erosion or suffusion [190]. From field data it is found that whether this occurs depends on the animal species: rabbits typically penetrate the cover layer, while burrows by dogs are much more superficial. On a more general level, there is a lot of uncertainty regarding the failure processes related to internal erosion and their quantitative risk analysis, such that assessment depends almost exclusively on engineering judgment [191].

7.1.3. Challenges in inspection and maintenance of flood defences

In many countries, several flood defence inspections are carried out throughout the year. Examples are condition inspections, emergency inspections and specific inspections, for instance aimed at detecting drought cracks CIRIA [3]. In the Netherlands the main inspections are relatively superficial biweekly inspections, detailed condition inspections during spring (sometimes in autumn), and inspections before, during and after high water situations. Nearly all inspections are done visually: other methods are used occasionally [e.g. 192–194] but are not common practice.

As was demonstrated in Chapter 6, the accuracy of visual condition inspections is limited and can vary significantly among inspectors and among different damage types. As such, flood risk analysis should also consider the scenario that the condition of a flood defence during an extreme event is influenced by the occurrence of damage. The probability of this scenario is determined by the probability that damage was not detected, and the probability that it developed after the last inspection.

The challenge of detecting and dealing with defects and damages is often encountered in management of infrastructure. For instance, drawing from the many available examples, Straub & Faber [73] consider inspection and maintenance planning in the context of damage to steel structures, Mendoza et al. [195] considered design of offshore steel structures in the context of inspection and maintenance of fatigue, and Barone & Frangopol [99] considered maintenance of bridges subject to corrosion. The main question is then, what distinguishes flood defence structures from other examples, if these are different at all?

There are a couple of differences that distinguish flood defences from most conventional structures. First of all, many flood defences are 'inactive' most of the time in the sense that they are not loaded at all, which means that damages and

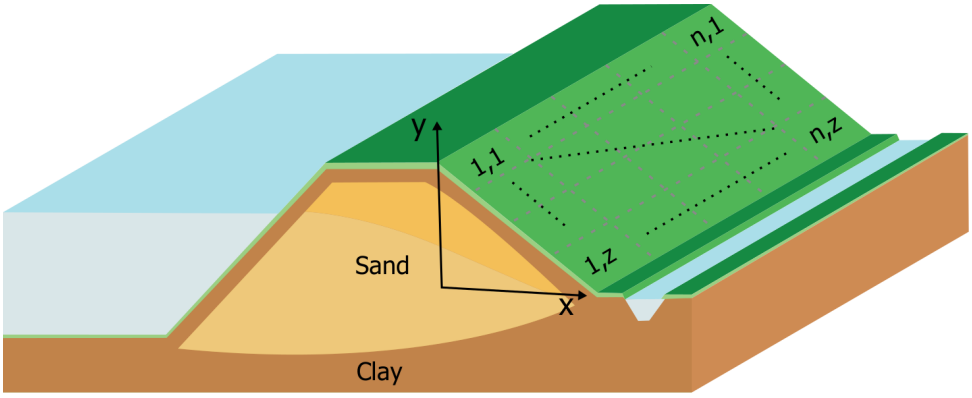


Figure 7.1.: *Schematized flood defence segment consisting of n sections (*1 km *length). The slope of each section is divided into z zones at which occurrence and consequences of damages are considered. The flood defence consists of a clay cover layer with grass revetment and a sandy core.*

anomalies can remain unnoticed for several months or years without causing any failures or other problems. This distinguishes maintenance of flood defences from for instance fatigue-sensitive structures, where cyclic loading can cause behaviour that increasingly deviates from the original behaviour [e.g. 196, 197]. Secondly, flood defences are primarily loaded during specific seasons (i.e., when most rainfall, snow melt or storms occur). As many of these damages can be repaired before an extreme load occurs, even the most detrimental damages do not necessarily have immediate structural consequences in terms of failure or flooding. Thirdly, while many of the examples in literature consider deterioration due to fatigue or corrosion or other continuous progressive processes [e.g. 27, 73, 81, 195, 198], many of the damages that impact the condition of flood defences are caused by random shock-based processes with potentially large shocks. This is specifically the case for grass revetments, where damages such as the occurrence of bare spots due to rutting or drought, or animal burrowing might be more or less likely given local excitation features [62], but their occurrence in time and space is largely random [172]. Also, although hard to quantify based on available data, some damage types might occur primarily during a specific time of the year (e.g., droughts during summer), and grass revetments might recover in time (e.g., bare spots typically cover with grass or other vegetations after some time) [184]. It should be noted that especially the degradation behaviour is very uncertain and can vary a lot per location.

7.2. Modelling inspection and maintenance of flood defence segments

In this section we translate the key aspects identified in the previous sections to a model with which maintenance concepts for flood defence segments can be evalu-

ated. In order to evaluate a maintenance concept, we need to answer two questions:

1. What is the conditional failure probability of a segment given different damages to the revetment at different sections and locations on the slope?
2. What is the probability of a flood defence segment being at a certain state at time t, while considering degradation, inspection and maintenance?

Figure 7.1 shows a flood defence segment that consists of $n \in N$ independent sections — each section n might or might not contain damaged spots, and might also have for instance different geometrical and/or soil properties. Typically such a segment consists of approximately 20 sections of 1 km length. Each section n is divided into $z \in Z$ (perpendicular) zones along the slope, where each zone z denotes the part between the vertical coordinates $[y_i, y_j]$. Zones are considered to be independent, and at each of them damage to the revetment can occur, resulting in a change of the failure probability of that specific zone z at section n. Note that in our case study we apply this approach to both the inner and outer slope. The state of the revetment at section n in zone z is represented by 4 states, which are related to the sod quality of the grass revetment as described in van Hoven & van der Meer [199] and Klerk & Jongejan [200]:

S_1 : the pristine state of the revetment. The sod quality of the grass is 'closed'.

S_2 : some damage to the grass sod due to for instance weeds. The sod quality of the grass is 'open'.

S_3 : major damage to the grass sod due to presence of bare spots or rutting. The sod quality of the grass is 'fragmented' and provides no erosion resistance.

S_4 : major damage to the grass sod and clay layer due to animal burrowing. The clay layer thickness is reduced, the sod quality of the grass is 'fragmented' and provides no erosion resistance. Note that we do not consider burrows that cause large voids in the sandy core of the flood defence.

7.2.1. The failure probability of a damaged flood defence segment

The failure probability of a flood defence section n susceptible to $m \in M$ failure modes at $z \in Z$ slope zones is given by:

$$P_{\mathrm{f}}(n) = P(\min_{m \in M, z \in Z} g_{n,m,z}(\mathbf{X}) < 0), \tag{7.1}$$

where $g(\mathbf{X})$ is the limit state function for failure mode m with uncertain inputs $\mathbf{X}$. The influence of damage is included by modifying specific parameters in $\mathbf{X}$ (in particular the erosion resistance of the grass and clay layer thickness).

If failure modes m are (partially) dependent this can be solved by for instance an integrated Monte Carlo analysis of Equation (7.1), or by combining failure modes

using the Equivalent Planes method [50], such that a failure probability is derived for each combination of z and m at section n. Subsequently, the failure probability can be upscaled to a flood defence segment in a similar way.

In practice, despite influence from the same load variables, failure modes are often considered to be independent for two main reasons [6, 85]: typically one failure mode dominates, or the design point values of shared random variables are different. In both cases assuming independence does not lead to significantly different failure probability estimates, while reducing computational complexity.

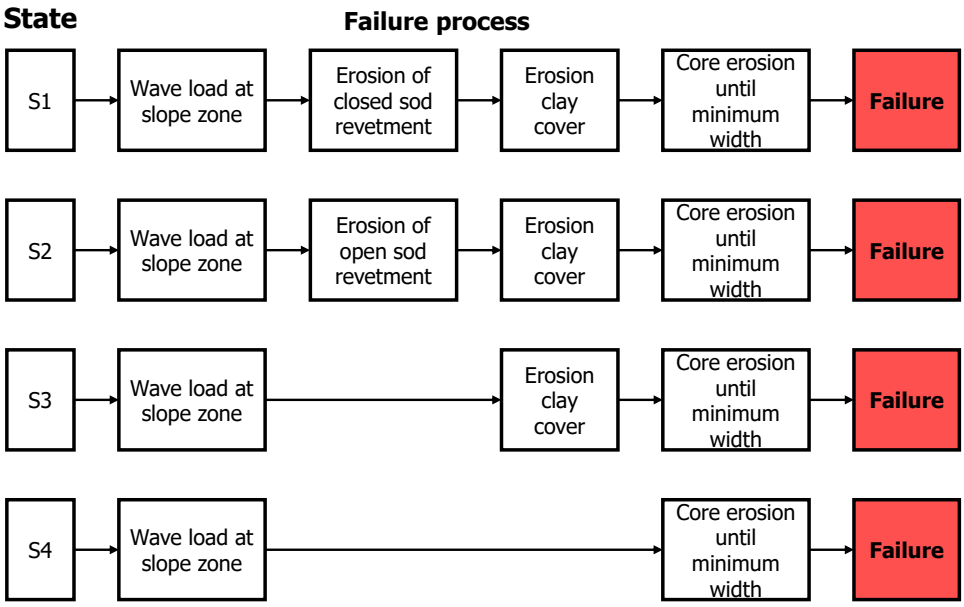
In this chapter we look at external erosion of the inner and outer slope. For riverine flood defences extreme water levels and waves are typically uncorrelated. As such, most failures of the outer slope are caused by events with large waves and moderate water levels, and most failures of the inner slope by events with high water levels and less extreme waves. Therefore, in line with the safety assessment in the Netherlands, we assume the failure of both slopes to be independent.

Figure 7.2a and Figure 7.2b show failure paths (i.e., a form of event trees) for external erosion of the outer and inner slope due to waves, given different initial states S. For outer slope erosion we assume failure occurs when the remaining width at the water level is lower than the original crest width. For inner slope erosion we assume failure if the clay cover layer has been eroded and the sandy core is exposed, as it is found from experiments that after this erosion proceeds extremely quickly. Additionally we assume that for every considered state S the flood defence fails due to overflow if the water level exceeds the crest. Further details on the physical models used for evaluating the failure process for both mechanisms and states are given in Appendix D.

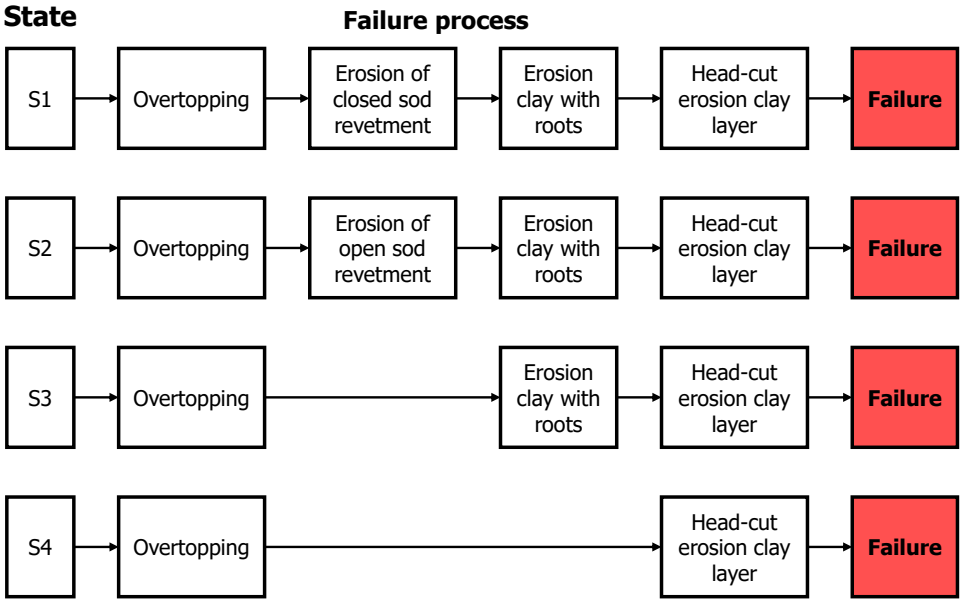
When considering damages at different slope zones z, damage will result in an increase of the failure probability, such that it is typically dominated by the damaged zone. Consider the slope in Figure 7.3: here the failure probability along the outer slope of a section n is shown with (brown) and without damage (green) (we only consider erosion of the outer slope here). The failure probability of the section is defined by Equation (7.1), and primarily determined by the part of the slope with the highest failure probability. If we discretize the slope into 4 zones as shown in the figure, this means that the most unfavourable case is when the second zone z_2 is damaged. In such a case the state of the other zones is irrelevant (provided that the difference in P_f is large). Logically, also the state of the zones for which it holds that $P_f(z_i, s) < P_f(z_2, S_1)$ are irrelevant considering Equation (7.1). We can use this to approximate the failure probability of the section n with $z \in Z$ zones and $s \in S$ states using the following equation:

$$P_f(n) \approx 1 - \prod_{z,s}^{Z,S} (1 - P_f(z,s) \cdot P(z,s) \cdot P(P_f(z,s) > P_f(z^*, s^*))) \tag{7.2}$$

where $P(P_f(z,s) > P_f(z^*, s^*))$ is the probability that zone z in state s is the weakest along the slope (i.e., $P_f(z^*, s^*)$ is the failure probability of the weakest zone at the section). We assume that for a given zone z the failure probability $P_f(z,s)$ is equal to the highest failure probability in that zone, in accordance with Equations (D.6)



(a) *Erosion due to wave impact at the outer slope.*



(b) *Erosion due to wave overtopping at the inner slope.*

Figure 7.2.: *Failure processes for both failure modes for different states of the revetment.*

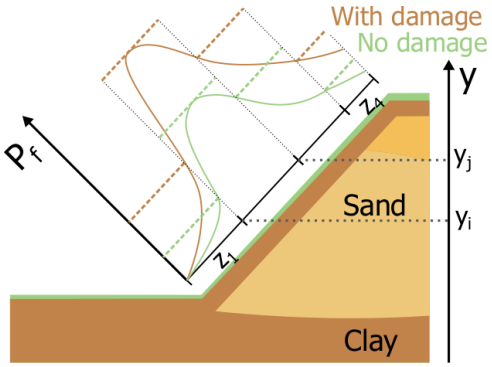


Figure 7.3.: *Failure probability along the slope for 2 states s with (brown) and without (green) damage. Solid lines indicate $P_f(y|s)$, dashed lines indicate the probability of failure for a slope part $P_f(z|s)$.*

and (D.7).

As was displayed in Figure 7.1, we have multiple sections $n \in N$ along the flood defence segment. In order to determine the failure probability of failure mode m at a segment level we look at the probability that zone z at section n is the weakest zone along the entire segment with $n \in N$ sections, such that:

$$P_f(m) = 1 - \prod_{n,z,s}^{N,Z,S} (1 - P_f(n,z,s) \cdot P(n,z,s) \cdot P(P_f(n,z,s) > P_f(n^*,z^*,s^*))) \quad (7.3)$$

where $P(P_f(n,z,s) > P_f(n^*,z^*,s^*))$ is the probability that the zone z at section n is the weakest along the dike segment.

Next, as we have multiple independent failure modes $m \in M$ we can combine them and obtain the failure probability of the flood defence segment:

$$P_{\text{f,segment}} = 1 - \prod_{m \in M} (1 - P_f(m)) \quad (7.4)$$

7.2.2. The influence of inspection and maintenance on reliability in time

Now that we have quantified the probability of failure of a flood defence segment given a combination of states for the different zones z and sections n, we can translate this to a decision model that quantifies the reliability of the segment in time, as well as the cost of different decisions.

Decisions on maintenance and inspection for each element n and each slope part z for a period of $t \in T$ time steps can be structured using the decision tree in Figure 7.4. At time t, there is the option to do an inspection $i_{n,z}(t) \in I_{n,z}(t)$ at section n at slope part z. Next, an observation $o_{n,z}(t) \in O_{n,z}(t)$ is obtained which can be used to update the belief about the state of part z at section n ($s_{n,z}(t) \in$

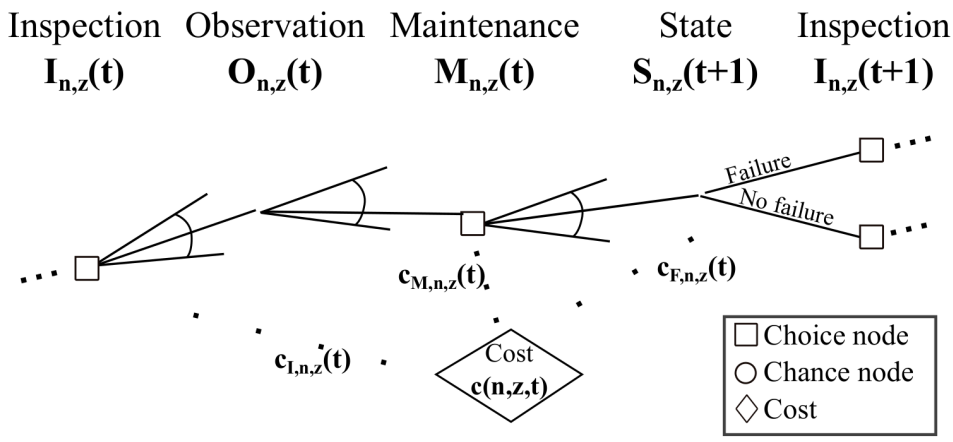


Figure 7.4.: *Generic decision tree for the sequential actions of inspection, observation, maintenance and state development of dike section n and slope zone z.*

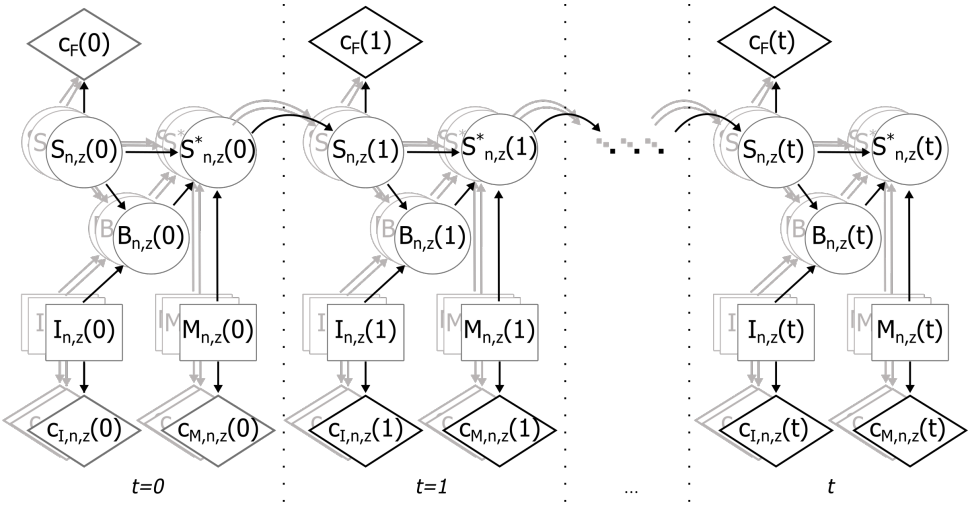


Figure 7.5.: *Influence diagram for inspections I, maintenance M and state development S of dike section n and slope zone z including different cost contributions at every time step. Note that shaded nodes and edges represent different sections n and slope zones z, which together determine the $c_F(t)$ (i.e., the cost of failure of the system).*

$S_{n,z}(t)$). In order to improve $s_{n,z}(t)$, maintenance actions ($m_{n,z}(t) \in M_{n,z}(t)$) can be taken such as repair works, overhaul, or more intensive monitoring of the section and zone. Jointly with the autonomous degradation, these actions influence the state $s_{n,z}(t+1) \in S_{n,z}(t+1)$ after which the sequence of decisions repeats itself. Inspection and maintenance actions contribute to the cost $c_{n,z}(t)$, as is the case for potential failures given the state $s_{n,z}(t)$. Note that this cost contribution should be considered at the level of system failures (i.e., flooding of the hinterland) which will be considered more explicitly in the remainder of this study.

For systems with many elements and sequential choices influence diagrams are more suitable to structure decision problems [26]. Influence diagrams are an extension of Bayesian Networks, and also include decision and utility nodes [79]. Figure 7.5 shows the influence diagram for I&M for a flood defence segment of $n \in N$ sections with $z \in Z$ slope parts for a period of $t \in T$ time steps. Note that the observation ($o_{n,z}(t)$) is not present in the influence diagram but replaced by a belief node $B_{n,z}(t)$, which represents the belief on the state $S_{n,z}$ conditional on the actual state of the slope part and the inspection that has been done. Additionally we add a node $S^*_{n,z}(t)$ that represents the state after maintenance $M_{n,z}(t)$ of the observed damages (represented by $B_{n,z}(t)$). Subsequently, autonomous degradation yields $S_{n,z}(t+1)$.

Provided that the transition matrix does not vary in time, the Dynamic Bayesian Network for each section n and zone z is in fact a Partially Observable Markov Decision Process (POMDP) [201]. However, as we look at a system scale this is no longer the case, as the utility (in particular the cost of failure $c_F(t)$) depends on the state of the flood defence segment rather than that of an individual section and zone. Note that this could be resolved by reformulating the different nodes to include all relevant combinations of states and actions for each section and zone, but this is a non-generic step as this differs per flood defence segment, and solving such a problem is computationally challenging due to the extremely large number of possible combinations.

In the following subsections we outline the different parts of the Dynamic Bayesian Network used to evaluate different maintenance concepts.

State of the flood defence system

The state S of the revetment is modelled as a Markov process where the state of zone z at section n is represented by the 4 previously introduced states (S_1 to S_4). In our implementation we use a transition from $S^*_{n,z}(t)$ to $S_{n,z}(t+1)$ using the following progressive transition matrix:

$$P(S_{n,z}(t+1)|S^*_{n,z}(t)) = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ 0 & p_{22} & p_{23} & p_{24} \\ 0 & 0 & p_{33} & p_{34} \\ 0 & 0 & 0 & p_{44} \end{bmatrix} \tag{7.5}$$

where p_{jk} indicates the probability that the state transitions from state j at t to k at $t+1$. While we assume that degradation is completely random, there are

ample examples in literature where continuous degradation (combined with random degradation) has been considered using transition matrices [27, 61, 81], and this extension could also be made here.

Inspection and maintenance actions

For inspections we primarily consider visual inspections characterized by a Probability of Detection (PoD). We only consider failure to detect damage, although it was demonstrated in Chapter 6 that also incorrect classifications can lead to incorrect beliefs about the state of a flood defence. The PoD can vary, depending on the type of inspection and the state (i.e., not all damages are detected with the same accuracy).

In the influence diagram in Figure 7.5, inspections are implemented by a conditional probability $P(B|S,I)$ which relates the belief $b \in B$ to the actual state $s \in S$ and the inspection action $i \in I$. For an inspection i we use the following belief matrix:

$$P(B|S,i) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1-\text{PoD}(S_2)_i & \text{PoD}(S_2)_i & 0 & 0 \\ 1-\text{PoD}(S_3)_i & 0 & \text{PoD}(S_3)_i & 0 \\ 1-\text{PoD}(S_4)_i & 0 & 0 & \text{PoD}(S_4)_i \end{bmatrix} \tag{7.6}$$

As such, the PoD can differ for each state, and if no damage is observed it is assumed that the state is S_1 (pristine state).

Based upon the belief, maintenance actions can be taken in order to obtain the state after interventions $P(S^*|B,M,S)$. We consider 3 interventions: small-scale repair (M_1), complete overhaul(M_2), and monitoring (M_3) of the revetment. Maintenance actions can be planned both as condition and time-based maintenance, where condition-based maintenance is carried out automatically after an inspection, and time-based at a fixed point in time [169]. If maintenance is executed in the model, all zones with a state equal or worse than the threshold S_c are repaired. We consider perfect repair, such that for a state $s \in S$ that is repaired it holds that:

$$P(s^*) = P(s) - P(B_s) \tag{7.7}$$

$$P(S_1^*) = P(S_1) + P(B_s) \tag{7.8}$$

with $s^* \in S^*$, and $B_s \in B$. In practice, some slightly damaged parts might not be repaired immediately but are monitored. This ensures that any further degradation is detected almost instantaneously, and that these spots can be dealt with through emergency measures in case of high water. Lendering et al. [153] demonstrated that the reliability of such measures mostly depends on errors in detection, followed by placement errors. As monitored spots are already detected and their characteristics are known, we assume that the failure probability of a slope with a monitored spot is equal to that of the intact state $P(F|S_1)$ (i.e. we assume that deployment of emergency measures is successful). Based on this assumption we add an additional state S_m to the matrices of S and S^*. If a zone with state s is to be monitored, it is

assumed that:

$$P(s^*) = P(s) - P(B_s) \tag{7.9}$$

$$P(S_m^*) = P(S_m) + P(B_s) \tag{7.10}$$

Furthermore, we assume that at these monitored slopes other new damage spots might occur with the same probability as for non-monitored slope parts, as the slope parts are large in comparison to the typical scale of damage. As such, if we add the transition probability for S_m as last row and column the transition matrix becomes:

$$P(S_{n,z}(t)|S_{n,z}^*(t-1)) = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & 0 \\ 0 & p_{22} & p_{23} & p_{24} & 0 \\ 0 & 0 & p_{33} & p_{34} & 0 \\ 0 & 0 & 0 & p_{44} & 0 \\ 0 & p_{12} & p_{13} & p_{14} & p_{11} \end{bmatrix} \tag{7.11}$$

The cost of inspection and maintenance have to be accounted for in the evaluation of Total Cost, jointly with the risk costs. The cost at time t can be computed as follows:

$$C(t) = P_{\text{f,segment}} \cdot D + \sum_{n\in N}\sum_{z\in Z} C_{\text{I,n,z}}(t) + \sum_{n\in N}\sum_{z\in Z} C_{\text{M,n,z}}(t) \tag{7.12}$$

where D is the damage due to a flood. By summing all costs for time steps t in a certain year we can obtain the annual cost C_{annual}.

Subsequently, we can also include investments for a longer investment using the Equivalent Annual Cost (EAC). The EAC is the annual cost of an asset. In our case this consists of the C_{annual} of flood risk, inspection and maintenance, and of the annualized costs of investments with a lifespan longer than a year. For investment C_{A} in a structural upgrade with lifespan t_{life} we can compute the equivalent cost of the investment (EAI) [202]:

$$EAI = \frac{C_{\text{A}} \cdot r}{1-(1+r)^{-t_{\text{life}}}}, \tag{7.13}$$

such that

$$EAC = C_{\text{annual}} + EAI, \tag{7.14}$$

where r is the discount rate ($r = 1.6\%$ [141]). The implementation of the cost computation is further discussed in Section 7.3.1 and Appendix D.2.

7.3. Influence of damage on flood defence segment reliability

7.3.1. Description of case study

For our case study we consider a flood defence segment consisting of 20 identical sections of 1 kilometre length. If the segment fails, the flood damage is 3.5 billion €.

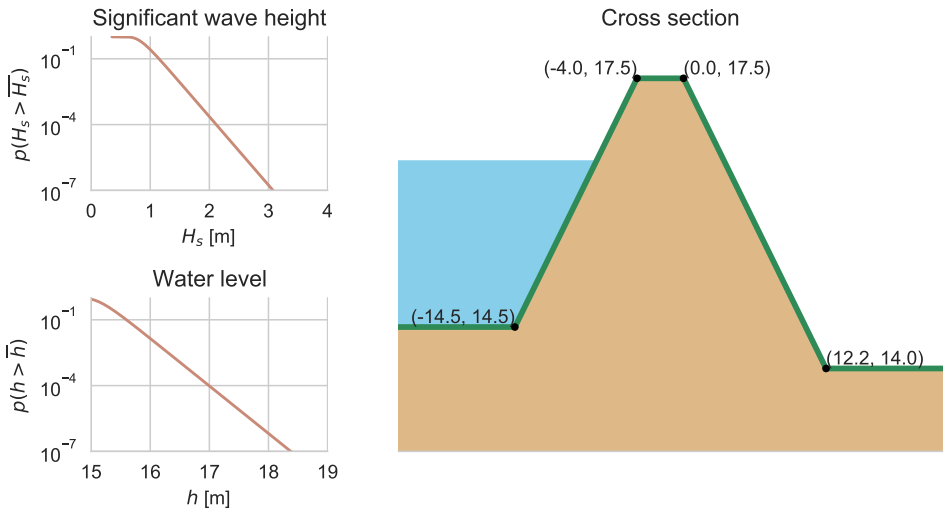


Figure 7.6.: *Left: distributions of significant wave height and water level for the case study. Right: geometry of the flood defence section.*

We assume the failure probability requirement ($P_{\mathrm{f,req}}$) for the segment is $1/10000$ per year, and $1/35000$ per year for external erosion in line with Jongejan et al. [6]. The right pane in Figure 7.6 shows the geometry of the considered cross section. We assume the sandy core is completely covered by a grass revetment and a clay layer of 50 centimetres thick. The left panes show the marginal distributions for the significant wave height H_s and water level h. Both h and H_s are represented by Gumbel distributions fitted to values for a location along the Rhine, as obtained from the hydraulic models used for the Dutch statutory safety assessment [203]. For wave loads we use a simplified criterion for depth-induced breaking on the foreshore: $H_s < 0.5(h - h_{\mathrm{foreshore}})$. For erosion at the inner slope we consider the entire slope from inner crest to inner toe. For erosion at the outer slope we consider the slope from the outer toe until 0.5 meters below the outer crest line — if the top part of the slope is loaded by waves of any relevance the overtopping volume will be so large that this will lead to failure of the inner slope. Input values and distributions for the different failure modes are described in Appendix D.

Degradation

To determine the probability of damage at a section, we analyse 6 years of inspection data of 470 km of primary flood defences along the Dutch Rhine. During the spring inspections in these years, inspectors registered all observed damages using the Digigids system (see Chapter 6) which is a classification system used for flood defence inspections in the Netherlands. Inspectors register the damage parameter (e.g., burrowing or rutting) and severity (good, reasonable, mediocre or bad) and take pictures of damaged spots. Additionally they can indicate the urgency of repair

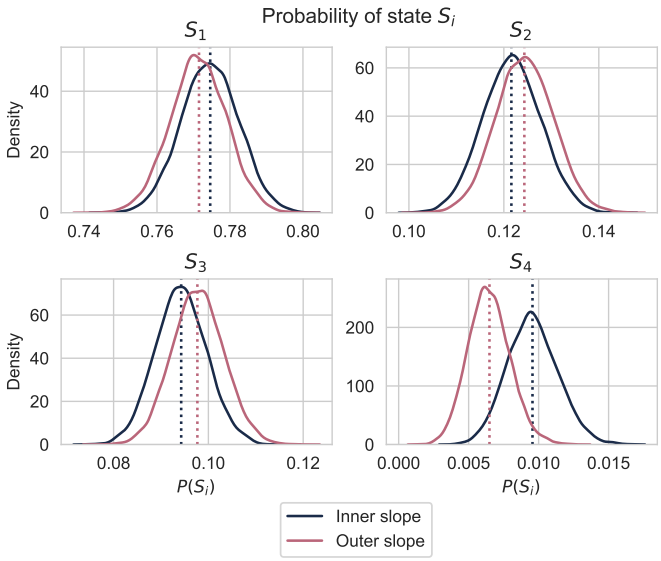


Figure 7.7.: $P(S_i)$ *for a flood defence section of 1000 meters for both inner and outer slope. Kernel Density Estimates obtained from 10000 bootstrap samples of the original dataset (dotted line).*

(e.g., emergency repair or medium urgency). We use these parameters to couple damage spots to estimates of the state of flood defence sections of 1 km length. We split the 470 km into sections of 1000 m, and couple damage registrations for each year to the nearest section. Based on the worst reported damage, we determine the sod quality of the revetment in a given year, for both slopes. S_4 is assigned if there is a burrowing damage with severity bad, and the urgency indicates that it is to be repaired before the next winter season. S_3 consists of all other damages with severity bad, S_2 of all damages with severity mediocre, and S_1 is assigned to all other sections.

There are a few remarks towards the data used. First of all, as these are field observations, the data not only represents the influence of degradation but also maintenance interventions influence the transition of the state between subsequent years. Secondly, as the inspections are approximately 1 year apart, there might be other recovery and degradation processes on a shorter time scale, e.g. due to seasonal influences, that do not emerge clearly from the data. Thirdly, it was demonstrated in Chapter 6 that the registrations in spring inspections are inconsistent in two ways: not all damages are detected, and the severity of damaged spots is often misclassified. As such, deriving transition probabilities between states based on yearly observations of a visual inspection is not a realistic approach.

However, we do have an estimate of how many sections are in a certain state S_i at the time of the inspection, and that this is typically repaired quite soon afterwards.

Table 7.1.: *Overview of all inspection and maintenance actions considered in this study, including the baseline maintenance concept. Note that the baseline parameters indicate the first time an action is taken (T_0) and the interval (ΔT), both in weeks. Costs are per kilometre per slope. Note that monitoring costs are assumed to be included in the general inspection by car.*

Action	Description	Kind	Threshold	PoD			Cost	Baseline	
			S_c	S_2	S_3	S_4	€/km/slope	T_0	ΔT
I_1	General inspection by car	Periodic		0	0	0.05	30	0	2
I_2	Condition inspection by foot	Periodic		0.6	0.6	0.6	120	13	52
I_3	Specific burrowing inspection	Periodic		0	0	0.6	120	-	-
M_1	Repair of damaged spot	Condition-based	S_4				680		
M_2	Overhaul of all damaged spots	Time-based	S_2				1920	15	52
M_3	Monitoring of a damaged spot	Condition-based	S_2				0		

As such we can derive the probability that, after a year, a section is in state S_i.[2] Figure 7.7 shows the $P(S_i)$ for the inner and outer slope at a random section, including Kernel Density Estimations obtained from 10000 bootstrap samples of the dataset. It can be observed that $P(S_i)$ differs slightly per slope, in particular for S_4, but for S_4 the (relative) variation from the bootstrap samples is also larger than for the other states. For the analysis we use the same transition probabilities for each slope. Based on the mean $P(S_i)$ of both slopes we use the following transition matrix for degradation of both slopes:

$$P(S_{n,z}(t)|S_{n,z}(t-1)) = \begin{bmatrix} 0.773 & 0.123 & 0.096 & 0.008 \\ 0 & 0.912 & 0.096 & 0.008 \\ 0 & 0 & 0.992 & 0.008 \\ 0 & 0 & 0 & 1.0 \end{bmatrix} \tag{7.15}$$

Note that this is the transition matrix for a year and for the entire slope, the transition matrix for an individual time step and zone is obtained by rescaling under the assumption that the probability of degradation is constant over a year and equal along the different slope zones z. Maintenance and monitoring is accounted for as described in Section 7.2.2.

Baseline inspection and maintenance

We assume that inspection and maintenance actions are always carried out for the entire flood defence segment of 20 kilometres. For our case we consider a baseline maintenance concept that consists of 2 inspection actions (I_1 and I_2) with different

[2] It should be noted that there are still false negatives due to inaccurate inspections.

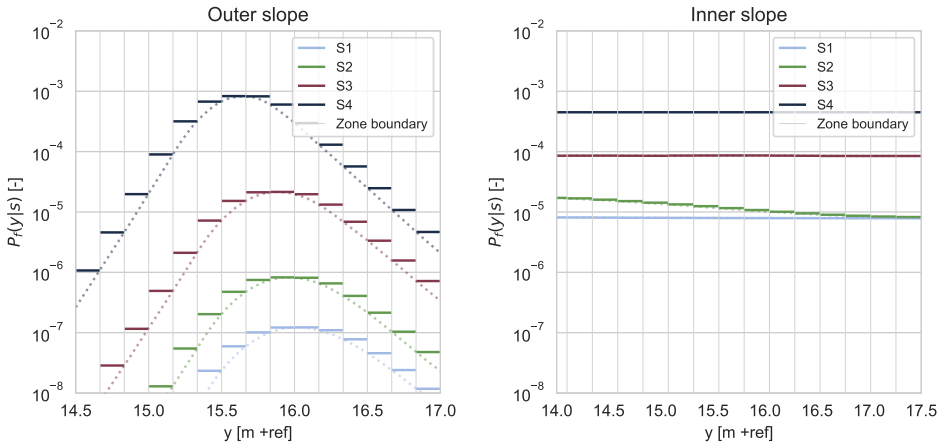


Figure 7.8.: *Failure probability $P_f(y|s)$ for external erosion at the outer (left) and inner (right) slope (with $s \in S$). Dotted coloured lines indicate $P_f(y|s)$, solid lines indicate the discrete values per zone z along the slope. Values are for a single flood defence section.*

accuracy and frequency. I_1 is an inspection by car that is typically done biweekly and not specifically aimed at condition inspection. Due to the general character of this inspection only major damages (S_4) will be discovered with a relatively low PoD. Note that as inspections are considered independent and carried out biweekly, the PoD of a given damaged spot is still relatively high (e.g., for a damaged spot with biweekly inspections the probability of the spot being detected after 26 weeks is $\approx 50\%$). I_2 is a condition inspection by foot as considered in Klerk et al. [151], and is aimed at detecting all meaningful damages. Additionally we consider 3 types of maintenance actions: M_1 is aimed at repairing a single damage spot, M_2 at overhauling all slightly damaged spots at a slope. Repair actions are carried out if it is detected that the state of a section $S_i \leq S_c$, where S_c is the critical threshold. M_3 is a specific action that ensures damages are continuously monitored, such that their further development or degradation is known. Such spots can then be repaired later during time-based maintenance actions such as M_2. Monitoring has no additional cost as in practice this will be done as part of the general inspection (I_1). Table 7.1 displays costs and other parameters for the different actions, further details are given in Appendix D.2.

7.3.2. Results for baseline inspection and maintenance

Figure 7.8 shows the failure probability along the slope $P_f(y|s)$ for external erosion at the inner and outer slope for the different states $s \in S$ for a section n. For the outer slope it can be observed that there is a zone around approximately 15.7 m +ref where P_f is highest, such that even if there is damage at other parts of the slope

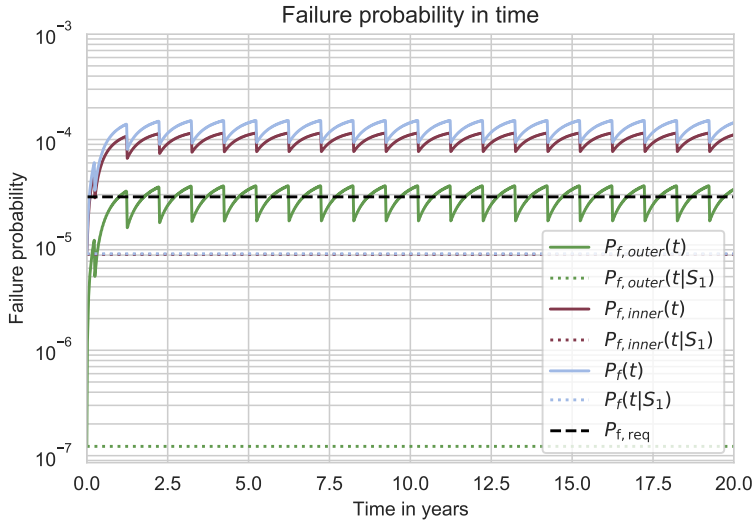


Figure 7.9.: *Annual failure probability in time $P_f(t)$, for inner and outer slope and overall. Dotted coloured lines indicate the failure probability without damage $P_f(t|S_1)$.*

(e.g., at 14.5 m +ref) the overall failure probability will typically not be impacted. This is primarily caused by the fact that the slope is loaded most severely at this level. For the inner slope there is a strong influence of S_3 and S_4 on the reliability. A notable difference is also visible for S_1 and S_2: $P_f(y|S_2)$ varies along the slope as the speed of the overtopping wave front increases further down the slope. For $P_f(y|S_1)$ this is not visible as failure is dominated by overflow failure. It should be noted that $P_f(y|S_3)$ is also dominated by wave overtopping, but the U_c of clay is very low, such that the failure probability does not significantly depend on y.

Without accounting for damage, the requirement of $1/35000$ for the overall failure probability is met both for S_1 and S_2. However, other damages at both slopes might result in a failure probability that is unacceptable. The question is whether the currently implemented maintenance concept is sufficient to ensure that the failure probability requirement is met. To assess that we run the Dynamic Bayesian Network using this maintenance concept.

Next we use Equations (7.3) and (7.4) to obtain the failure probability for the entire flood defence segment including damage and I&M. This is shown in Figure 7.9, and compared with the failure probability without any damage $P(F|S_1)$. Without damage, this segment meets the requirement $(1/35000)$ for external erosion, provided that the slopes are in state S_1 or S_2. It is clearly observed that damages to both the inner and outer slope cause a significant increase in failure probability, and when accounting for damage and the baseline maintenance concept the requirement is no longer met: the maximum failure probability is approximately 20 times higher when accounting for damage, inspection and maintenance. Note that we evaluate a period of 20 years, but after approximately 3 years an equilibrium

situation is reached. As such, we can evaluate the costs for both cases by looking at the average costs per year C_{annual} as obtained through Equation (7.12).

Table 7.2 shows the average costs per year for the period of 10 to 20 years in the simulation. Here we observe that the initial risk costs without damage are similar to the expenditure on inspection and maintenance. However, accounting for damage to the flood defence gives a large increase in flood risk costs. As such, additional measures are required, and likely cost-efficient.

Table 7.2.: *Annual cost of inspection, maintenance and flood risk. Flood risk costs are split into the initial risk (without damage) and the risk increase due to damage.*

Cost component	€/yr
Maintenance	13385
Inspection	18000
Initial risk	28848
Risk increase	4.74×10^5
C_{annual}	5.34×10^5

7.4. Improving robustness and reliability of flood defence segments

In the preceding section we have identified that the contribution of revetment damage to flood risk can be significant, and is much higher than the expenditure on Inspection & Maintenance (I&M) for the considered case. This indicates that the current maintenance concept is insufficient for this flood defence segment. In this section we evaluate several interventions that could improve the overall performance, which we evaluate through total cost and structural robustness to damage. We consider changes to the maintenance concept as well as investments in structural upgrades of the flood defence.

7.4.1. Approach

Structural robustness is typically defined as that 'the consequences of structural failure should not be disproportional to the effect causing the failure' [40], and is mostly used in the context of designing structures to prevent progressive collapse [204]. For structures with large potential consequences of failure, the Eurocode [40] specifically advises to include a risk-based analysis of the capability to withstand accidental loads caused by for instance explosions or human error. For flood defence systems robustness can be particularly useful as a performance measure for estimating the relative contribution of accidental damage and human error by inspectors, and its consequences for system reliability. Baker et al. [39] proposed a quantitative indicator for structural robustness that relates direct and indirect consequences of structural damage. We apply this indicator here, and presume that direct consequences are flood risk costs for the flood defence segment in a pristine

Table 7.3.: *Different inspection schedules evaluated as part of the analysis of different maintenance concepts. PoD and cost of inspection actions are listed in Table 7.1. All units are in weeks.*

	I1		I2		I3		I4		M3	
Policy name	T_0	ΔT	T_0	ΔT	T_0	ΔT	T_0	ΔT	T_0	ΔT
Base	0	2	13	52					15	52
Autumn inspection	0	2	39	52					41	52
Base & autumn inspection	0	2	13	26					15	52
Base & burrowing inspection	0	2	13	52	45	52			15	52

state, while indirect consequences are the additional consequences due to the potential presence of damage during a high water, such that the robustness indicates the sensitivity of system risk estimates to structural damage through degradation. This yields the following equation for robustness indicator I_R:

$$I_R = \frac{R_{\text{init}}}{R_{\text{init}} + \Delta R} \tag{7.16}$$

Where R_{init} is the initial flood risk if damage is not taken into account (i.e., the state is always S_1), and ΔR the risk increase due to damage to the flood defence (both in €/yr). Similar to the previous section we use the average R_{init} and ΔR from year 10 onward. As Baker et al. [39] state, robustness might decrease if the (relative) direct consequences of a failure increase. Therefore we both look at the robustness indicator I_R and the costs in a multi-objective setting. It should be stressed that the relevance of solely looking at I_R is limited, as it is a relative indicator. As an indicator for cost we use Equivalent Annual Cost (EAC), using Equation (7.14).

In our analysis we consider (combinations of) two types of measures: structural upgrades of the flood defence, and changes to the maintenance concept. Table 7.3 shows the different maintenance concepts considered. In a separate analysis it was evaluated whether changing the threshold S_c for repair (M_1) to a lower value would yield benefits — as this was not the case this is not included in the analysis and the only difference between the different concepts considered is the inspection schedule. We also evaluate several (combinations of) structural upgrades for the different slopes. We consider 4 types of such interventions:

1. Burrow protection: construction of a protective layer (e.g. grid or geotextile) such that burrowing animals can not penetrate the top layer. It is assumed that for this upgrade $P(F|S_4) = P(F|S_1)$. This is applicable at both inner and outer slope.

2. Increase clay layer thickness: reinforcement by increasing the clay layer thickness at the inner slope to 1 meter in order to increase the erosion resistance of the slope.

3. Reduction of inner slope angle: reinforcement by decreasing the inner slope from 1:3.5 to 1:4.5, such that overtopping waves have a lower flow velocity

Table 7.4.: *Considered interventions and their equivalent cost of investment (EAI).*

	Intervention	EAI €/km/yr
Maintenance concepts	Base	900
	Autumn inspection	900
	Base & autumn inspection	1020
	Base & burrowing inspection	1020
Structural interventions	Increase of inner slope clay cover to 1 meter	12067
	Decrease inner slope to 1:4.5	10056
	Burrow-preventive measures at outer slope	4223
	Burrow-preventive measures at inner slope	4223
	Crest heightening 0.5 m	28157
	Crest heightening 1.0 m	36202

and cause less erosion. Additionally this slightly increases the allowed erosion volume for wave impact at the outer slope as the width of the flood defence is increased.

4. Crest heightening: increase of the crest height by 0.5 or 1 meter. Reduces the probability of overtopping and increases the allowed erosion volume for wave impact at the outer slope.

Table 7.4 shows the equivalent cost of investment for each considered intervention. Details on the computation of these values are given in Appendix D.2. Figure 7.10 shows the failure probability for the different structural adaptation alternatives. Note that burrow-preventive measures are not included in the figure, but can be combined with the displayed measures such that $P(F|S_4) = P(F|S_1)$. Worthwhile to note is that for instance a clay cover of 1 meter has a relatively large impact on $P(F|S_3)$ and $P(F|S_4)$ for inner slope erosion, while increasing the crest height impacts the failure probability for all states.

7.4.2. Comparison of interventions

Now we can compare the effectiveness of (combinations of) different maintenance concepts and structural interventions. Figure 7.11 of I_{R} and EAC displays the result of all considered combinations, where different maintenance concepts are distinguished by coloured dots. Axis represent the robustness indicator I_{R} and the EAC consisting of flood risk cost, and costs of structural upgrade and/or the maintenance concept as determined using Equation (7.14). Stars indicate different maintenance concepts without any structural upgrade. The Pareto front is indicated by the black line and dots, for which the underlying structural upgrades are shown in the table. With regards to the maintenance concept it can be seen that especially an inspection in autumn can both increase robustness and reduce EAC: each Pareto optimal solution has a maintenance concept with condition inspections both in spring and autumn.

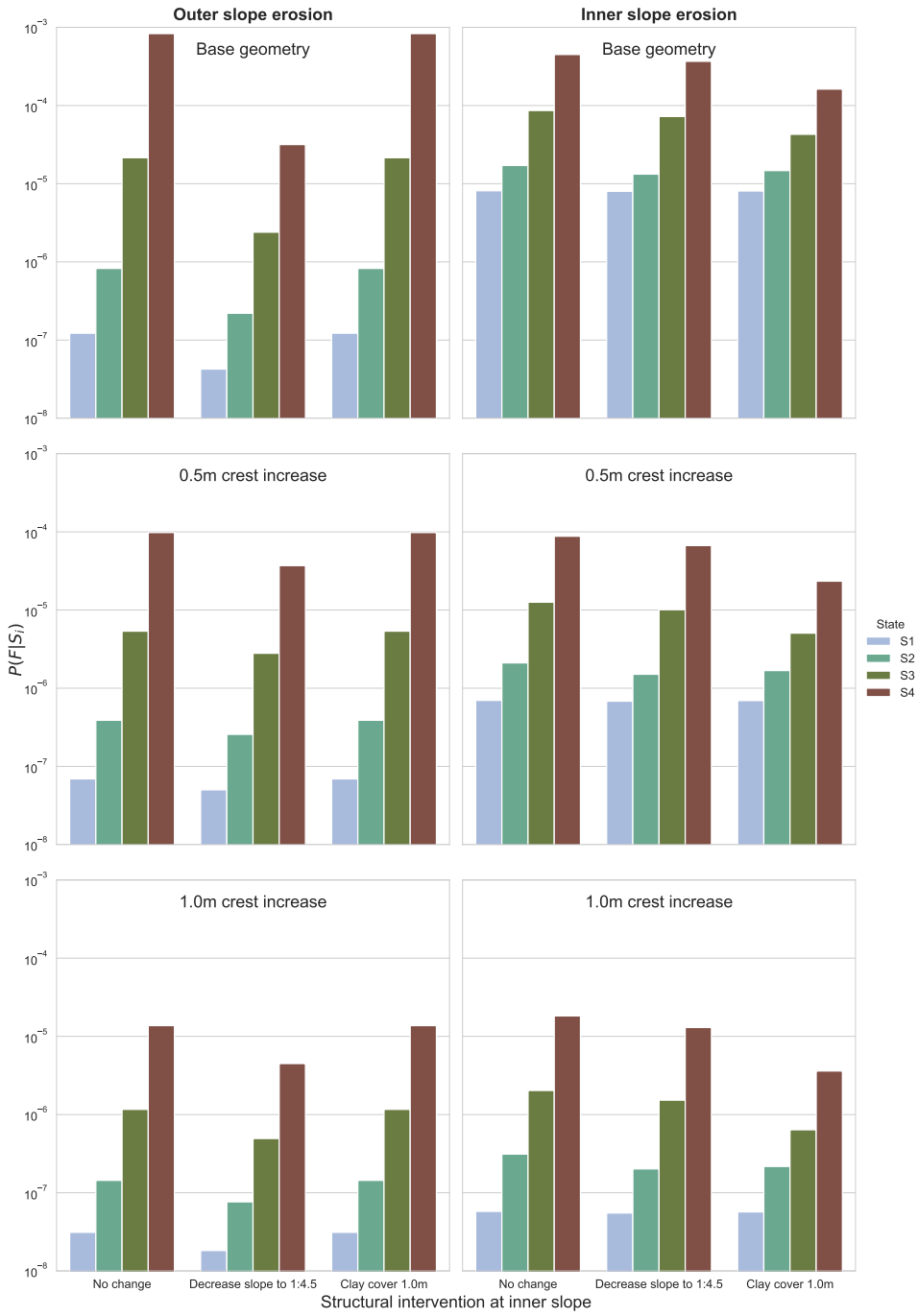


Figure 7.10.: $P(F|S_i)$ *for different combinations of crest increase and changes to the inner slope for the different states.*

A structural intervention that greatly improves robustness is to increase the thickness of the clay cover at the inner slope. Installing additional protection against burrowing at either or both slopes can further increase robustness, albeit at the expense of a slightly higher EAC.

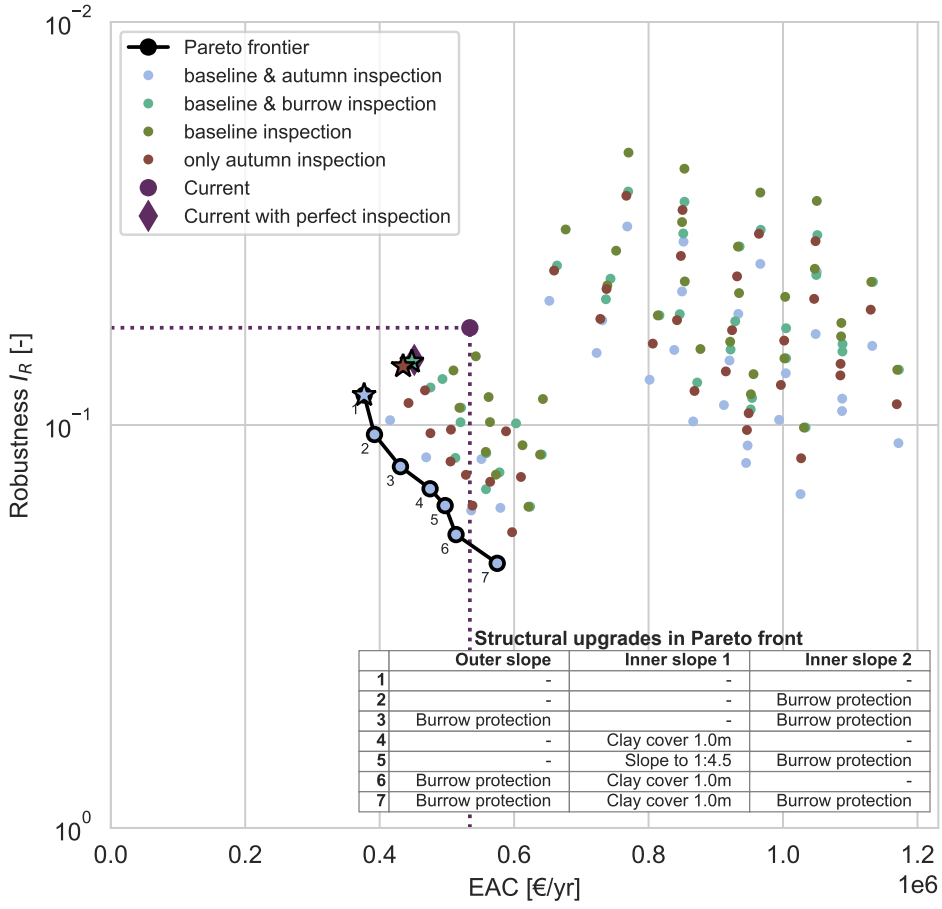
Increasing the crest level was also considered in the analysis: while this decreases the overall failure probability, the high costs of this intervention result in an increase in EAC and a decrease in I_R. In other cases this might be different for costs, but it is expected that a decrease in robustness will also be encountered in other cases due to the lower R_{init}. With the definition of structural robustness that we use here, a safer flood defence is thus not necessarily more robust.

7.4.3. Requirements for new inspection methods and policies

In recent years many pilots and tests have been conducted with a variety of different methods for inspection of infrastructure. Examples are satellite monitoring of deformations [192], use of remote sensing data [205] and inspection by helicopter or using drones [194]. While the potential of such methods is high, the point at which they can outperform or supplement existing visual inspections in a cost-effective manner is unknown. In this section we explore at what combination of inspection frequency and PoD a new inspection method would outperform the optimal maintenance concept found in the preceding section (baseline + autumn inspection). To that end, we consider a maintenance concept with a recurring drone inspection, and no other inspections. Rules for maintenance are the same as in the preceding section, and we do not consider structural upgrades. We assume the cost of a drone inspection is 250 €/km, twice as high as for a visual inspection. In practice actual costs can vary greatly per method and for instance depend strongly on the scale at which they are applied.

Figure 7.12 shows the EAC for different inspection intervals and PoDs for a maintenance concept with drone inspections. The orange line denotes the optimal inspection frequency conditional on the PoD, as each PoD potentially leads to a different optimal inspection interval. The dotted and dashed lines show how this policy compares to the situation with base & autumn inspection (see Table 7.3 for details): for all combinations below the dashed line the structural robustness is higher than for the reference policy. For all combinations below the dotted line also the EAC is lower than for the reference situation. For instance, if a drone inspection would have a PoD of 0.5, an inspection interval of 12 weeks would lead to lower costs and a more robust flood defence segment. Obviously such values vary depending on system characteristics and costs of inspections.

One particular point of attention is the definition of the PoD. In the model in this chapter observations from inspections are assumed to be independent. For visual inspection (see Chapter 6) but especially also for automated inspections there are several aspects that likely cause (partial) correlation between different observations. Examples are the method of inspection, the approach for interpreting/processing the results and the characteristics of the flood defence (e.g., location/length of vegetation or presence of hard to reach places). As such, it is likely that there



	Outer slope	Inner slope 1	Inner slope 2
1	-	-	-
2	-	-	Burrow protection
3	Burrow protection	-	Burrow protection
4	-	Clay cover 1.0m	-
5	-	Slope to 1:4.5	Burrow protection
6	Burrow protection	Clay cover 1.0m	-
7	Burrow protection	Clay cover 1.0m	Burrow protection

Figure 7.11.: *Pareto front for robustness index I_R and Equivalent Annual Cost (EAC) (annual cost of flood risk, inspection, maintenance and structural interventions). Pink diamond and dot indicate current situation with/without perfect inspection. Colours indicate different maintenance concepts, with stars indicating combinations without any structural upgrades. Structural interventions of the Pareto optimal solutions (black line) are given in the table at the bottom left.*

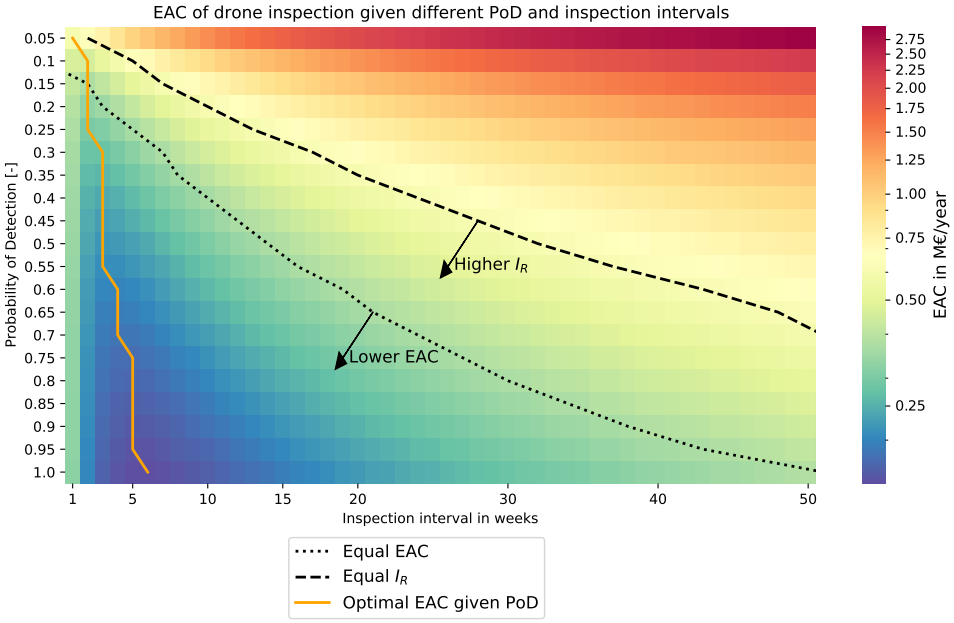


Figure 7.12.: *Heatmap of EAC of a drone inspection for different intervals and PoD. Orange line denotes the optimal inspection interval for different PoD, dashed line denotes the values for which the I_R is equal to that of the optimal maintenance concept from the preceding section, dotted line denotes equal EAC. Arrows indicate the area where the optimal maintenance concept is outperformed by the drone inspection.*

is some correlation between observations. If that is the case, one would have to account for this correlation by considering an effective PoD which accounts for this effect.

To assess the potential impact of this effect we model the correlation between inspection observations using a Gaussian autocorrelation function, for which the correlation in time $\rho(t)$ is given by:

$$\rho(t) = (1 - \rho_0) \cdot \exp^{-\frac{t}{t_0}} + \rho_0, \tag{7.17}$$

where ρ_0 is the non-ergodic part of the correlation (i.e., the part that does not vary in time), and t_0 is the correlation time (i.e., the time period for which observations are correlated). As such, we can consider a time-variant (through t_0) and time-invariant (through ρ_0) component of correlation. It should be noted that for inspections other autocorrelation functions might be equally or more valid, but this has not been investigated, and our goal here is to illustrate the potential influence on the effectiveness of inspection policies. By sampling correlated observations in time, we can assess what the probability is that damage is detected during time period T for a combination of ρ_0 and t_0, and a given PoD.

We consider an inspection scheduled at an interval of 4 weeks for which the PoD of an individual inspection is 0.5. Figure 7.13 shows the effective PoD, which is the PoD of each inspection while accounting for correlation with preceding inspections. Red and green lines indicate the theoretical values for fully (un)correlated inspections. It can be seen that both the time-variant (t_0) and time-invariant (ρ_0) correlation have a significant influence on the effective PoD. Concretely, if for a given inspection method with a PoD of 0.5 it would hold that $\rho_0 = 0.5$ and $t_0 = 13$ weeks, the effective PoD of a single inspection would be ≈ 0.23 if applied at an interval of 4 weeks. It should be noted that due to the high frequency of inspections, the PoD over a period of a year would be ≈ 0.97. This means that in the development of new, and improvement of existing inspection methods one also has to carefully consider the dependence between observations, as this greatly influences their practical performance. Practically, especially for inspections with a high frequency, such as highly frequent visual inspections, and/or high time-invariant correlation, for instance if drone images are processed by the same algorithm, this is of importance in properly assessing the Value of Information of such inspections.

7.5. Discussion & Conclusions

7.5.1. Discussion

This chapter has considered the influence that damage can have on the failure probability of a flood defence segment subject to external erosion. The model that was developed can be used to estimate the effect of different interventions on the failure probability at a segment level: both structural upgrades and changes to the maintenance concept.

Based on the available inspection data for similar riverine flood defences, degradation was modelled as a random process using a progressive Markov process. This

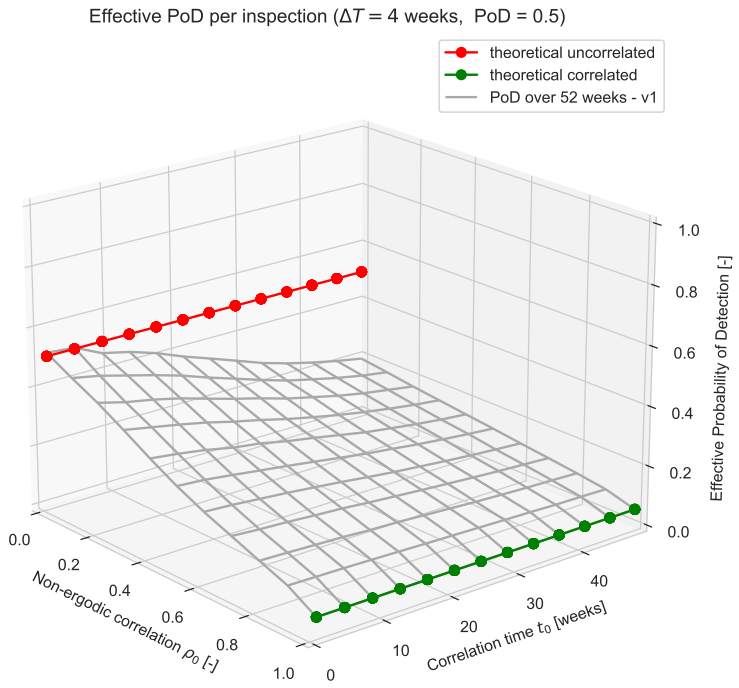


Figure 7.13.: *Effective PoD for inspections with* $\Delta T = 4$ *weeks and PoD*$= 0.5$*. Red and green lines indicate bounds for fully correlated and uncorrelated inspections.*

is different from previous studies [e.g. 62, 86], but aligns well with both data and field experience. It should be noted that degradation rates for specific locations are difficult to obtain due to a lack of complete inspection datasets that span a considerable time range. Main reasons are the limitations of existing inspection techniques (i.e., not all damages are detected), and the fact that inspection records have not been recorded consistently over longer periods of time. From practical experience it is qualitatively known that certain types of damage occur in specific seasons (e.g., weeds in spring and summer). However, there is no data available to substantiate such seasonal patterns. As it was found to have limited influence on overall outcomes, degradation rates were assumed to be constant over the year. Provided that estimates are available, this can be accounted for using the presented method and might lead to further optimization of timing of inspections throughout the year, or focussing inspections in certain months at specific types of damage.

In our analysis we considered a flood defence segment consisting of 20 identical sections. However, the general approach is also applicable for segments with sections of different lengths and properties. This is of relevance, as some types of damage might occur more frequently in specific locations: it was for instance found in Klerk & Adhi [172] that animal burrowing occurs more frequently in rural areas. In such cases it might be preferable to consider specific inspection policies for sections with specific excitation and ancillary features that can cause different degradation patterns and rates [62].

While we only considered external erosion in this analysis, it is found from literature that also other failure modes can be affected by the types of damage considered here, most notably inner slope instability and internal erosion [e.g. 28, 189]. However, due to the large variation in effects of different (especially larger) burrows and the uncertainties in modelling for instance internal erosion [e.g. 191] it was decided to not consider these in this study, despite their practical relevance.

Based on the findings in Chapter 6 we set the PoD of visual condition inspection to 0.6, but it should be noted that the PoD can vary greatly in practice. The PoD of inspection by car and specific inspections for animal burrowing were based on expert estimates. It should however be noted that the PoD varies greatly for different situations, per inspection type, and also per inspector. To better estimate the effectiveness of different types of inspections further investigations are required.

In the analysis of different interventions we combined both structural interventions and changes to the inspection policy. The influence on both the structural robustness and costs was assessed. To that end the robustness index proposed by Baker et al. [39] was reformulated for flood defences. From this it was found that both inspections and several targeted structural interventions can lead to lower costs and higher robustness of flood defence segments compared to the baseline situation.

In this chapter we formulated a structural robustness indicator with regards to the influence of damage to the revetment, which differs from robustness indicators used for flood defences in the past. An example is the Dutch robustness allowance proposed in ENW [206], where the flood defence height was increased to make flood defences more robust to potential uncertainty in water level estimates. As all

uncertainties in hydraulic loads are now explicitly taken into account this allowance is no longer used. Yet, robustness is often associated with increasing dimensions of a flood defence. The findings from this chapter with regards to increasing flood defence height show that this is not always correct, and that a broader interpretation of flood defence robustness may be required. It should be noted that the robustness indicator used in this chapter should always be considered jointly with an absolute indicator related to flood defence reliability (i.e., EAC or segment reliability).

Most studies into the effectiveness of inspection policies do not account for the correlation between different inspection observations. However, there are several factors that might cause observations between inspections to be (partially) correlated, for which the effects were conceptually estimated. For such methods, especially if inspections are executed at a high frequency, the effective Probability of Detection per inspection can be significantly lower. This might be an important factor in the evaluation of inspection policies, and it emphasizes the practical relevance of ensuring that observations from inspections are as independent as possible, for instance by combining different inspection techniques and rotating individual inspectors.

A limitation of this study is that it only considers regular inspections, but emergency inspections during high water situations are also an important part of flood defence asset management in general. In our analysis it is found that inspections in autumn are more effective than inspections in spring, as autumn inspections are closer to the high water season. This is even more so for emergency inspections as these are done right before or during a high water situation. As such, improvements of emergency inspections can likely significantly reduce costs and increase robustness of flood defence segments [153], and could be an important extension of the model.

7.5.2. Conclusions

In this chapter we presented a model that can account for the influence of degradation and Inspection & Maintenance (I&M) on the reliability of flood defence segments subject to external erosion. Based on a large dataset from inspections it was found that degradation of flood defences is mostly random in time. Using a Dynamic Bayesian Network, a riverine flood defence segment subject to relatively large waves was evaluated using a baseline inspection policy commonly applied in the Netherlands. From this it was found that without accounting for damage the segment satisfies the requirements. However, when accounting for damage the implemented I&M policy is insufficient to mitigate the consequences of damage to the revetment. For the case study, damage increases the failure probability by approximately a factor 20. On a more general level this means that flood defence design optimizations that do not account for the occurrence of damage likely lead to insufficiently safe flood defences.

Next, several changes to the maintenance concept, as well as structural upgrades were considered. Combinations of measures were evaluated based on resulting structural robustness and Equivalent Annual Cost (EAC). It was found that especially

increasing the clay cover layer thickness, and doing additional inspections in autumn are effective measures that both reduce EAC and increase structural robustness.

Finally the potential requirements for new inspection methods (e.g., a drone inspection) were considered in order to assess what PoD such an inspection should achieve for different inspection intervals in order to outperform existing policies. Here it is clearly demonstrated that improving inspection methods can not only lead to lower EAC but also to a higher structural robustness. It was also conceptually shown what the impact of correlation between inspection observations can be on the effective PoD: if observations are strongly correlated the added value of additional inspections is limited, which should be accounted for when evaluating new inspection policies and/or new inspection techniques.

While the degradation rates of different flood defence segments can vary significantly, the developed approach can both aid in assessing the impact of damage, and identifying efficient strategies for dealing with damage in flood defence segments. Potential extensions of the analysis are the inclusion of different sections with differentiated inspection policies and including emergency inspections in the model. A more fundamental improvement required for better assessment of the effectiveness of I&M is a better understanding of degradation patterns of flood defences (e.g., relation to structural properties and seasonal variation). Additionally a better understanding of the impact of damage on other failure modes such as inner slope instability and internal erosion is required to fully assess the impact of damage on flood defence system reliability.

8

Conclusion

A decision was wise, even though it led to disastrous consequences, if the evidence at hand indicated it was the best one to make; and a decision was foolish, even though it led to the happiest possible consequences, if it was unreasonable to expect those consequences.

Herodotus

This thesis has considered advances on three key topics that hamper efficient asset management of flood defence systems: the practical application of optimization of risk-based design at a system level, the reduction of time-invariant uncertainty, and dealing with uncertainty in time-variant processes. While each of the chapters in this thesis contains a discussion of specific results and conclusions, Section 8.1 summarizes the main findings for the three key topics, and Section 8.2 summarizes the main recommendations for both further research, as well as practical implementation. Section 8.3 provides some closing remarks.

8.1. Main findings

8.1.1. Optimization of system design

Past research on optimization of flood defence systems has mainly focused on deriving optimal reliability requirements. In this thesis we have advanced this towards the optimization of flood defence system reinforcements (Chapter 3). By adopting a system optimization approach rather than a design approach per individual section, Life Cycle Cost (LCC) of a flood defence system reinforcement was reduced by approximately 40 %. The main benefit is that through the adopted system perspective, reinforcement interventions can be optimally chosen in space and time. This is different from a reliability-based design approach per section, and significantly enlarges the solution space leading to more efficient investments. The developed method has been applied to an on-going reinforcement project, which emphasizes its practical applicability.

The size and complexity of flood defence systems, together with the connection to design, require optimization methods that can deal with large interdependent systems. In Chapter 3 we developed a greedy search algorithm that is both quick and accurate and can be applied for this type of optimization problems. Although the algorithm does not guarantee that a global optimum is found, differences in Total Cost (TC) of optimal design solutions are on average 0.04%. As such, from a practical perspective, the inaccuracies in the optimization method are eclipsed by other uncertainties (e.g., time-invariant strength uncertainty) in reliability estimates, and the method can be of great practical value in finding optimal design solutions for flood defence systems.

8.1.2. Reduction of time-invariant uncertainty

Chapters 4 and 5 focused on the cost reduction achieved by reducing time-invariant uncertainty based on the Value of Information (VoI). We primarily looked at reduction of uncertainty through pore pressure monitoring and proof load tests. Through a Bayesian pre-posterior analysis the VoI of both measures was determined, and found to be positive in many cases. Alongside determining whether reducing uncertainty is cost-effective, such an analysis also gives insight in the optimal sequence for different interventions for uncertainty reduction, and at which sections in a larger system uncertainty reduction efforts are most beneficial.

There are two key conclusions that emerge from the analysis in Chapter 4. First of all, in the context of a reinforcement, the benefits of uncertainty reduction strongly depend on the type of available reinforcement interventions: if only mitigation measures are available (such as stability berms, see Section 2.6 for definitions), uncertainty reduction is likely of value for the optimal decision. However, if cost-effective substitution measures (e.g., Vertical Sandtight Geotextile) are available, the benefits of uncertainty reduction might be negligible, as was for instance encountered in the case study in Section 4.3. This demonstrates that efforts to reduce uncertainty can only be properly valued within the context of the decisions they have an impact on. Secondly, it is found that the relative uncertainty contribution of different parameters has a large influence on the VoI — in Chapter 4 it

was for instance encountered that the VoI of pore pressure monitoring after a proof load test was higher than without a proof load test. This emphasizes the relevance of jointly considering multiple types of interventions for uncertainty reduction in a single decision analysis.

In Chapter 5 it was found that uncertainty in current estimates of flood defence performance not only influences the cost of upcoming investments, but might extend to flood defence reinforcements further in the future as well. As such, reducing (time-invariant) uncertainty in performance of flood defences should be a key part of plans to prepare flood protection systems for adaptation to an uncertain distant future.

In this thesis a distinction was made between time-dependent and time-independent interventions for uncertainty reduction. An example of time-dependent uncertainty reduction is pore pressure monitoring, where outcomes, and hence the VoI, depend on the duration of monitoring as it depends on randomly occurring high water levels (contrary to for instance Cone Penetration Tests which is time-independent). As short pore pressure monitoring campaigns might not yield any information, it is often better to execute them separately from flood defence reinforcement projects: projects might not allow for lengthy monitoring campaigns, which can significantly reduce their potential value for decisions.

8.1.3. Managing uncertainty in time-variant processes

The last key topic considered in this thesis is how uncertainty in time-variant processes, primarily degradation, is accounted for. Typically, reliability analysis of flood defence systems does not take into account the influence of most degradation processes and the implemented maintenance concept. In this thesis it has been demonstrated that the combined effect of imperfect inspections and randomly occurring damage to revetments can lead to a much lower reliability than is typically assumed in current assessments based on relatively 'perfect' revetments.

Degradation processes can have many different drivers and characteristics, but experiences from the field observations in Chapter 6 and the data analysis in Chapter 7 demonstrate that a major part of the degradation of flood defence reliability is caused by shock-based degradation processes with random shocks in time. Damage to the revetment caused by amongst others animal burrowing, rutting and drought is one of the causes for a reduction in reliability. Due to their random character and potentially significant effect on flood defence strength, the effective influence on flood defence reliability is potentially much higher than for instance that of gradual settlement of the crest. This is also demonstrated in Chapter 7 where the probability of failure for external erosion of a flood defence with and without damage was considered. For the considered flood defence, animal burrowing increases the failure probability of the inner slope by up to a factor 100, and for the outer slope by up to a factor 1000 compared to the undamaged situation. Hence, methods to timely detect such shocks are required to ensure flood defences meet their requirements.

In Chapter 6 the accuracy of visual condition inspections of flood defences was studied during a field experiment. Multiple inspections of several flood defence

sections revealed large variability in the Probability of Detection (PoD) for different individual damage spots (0 to 0.9), as well as different inspectors (on average 0.25 to 0.6). If damages are considered at a section level, the estimates of the PoD are somewhat higher. Although these estimates are lower bounds, it reveals a large uncertainty in the PoD of a single visual inspection, even though on average the PoD is somewhat similar to that of inspection of other types of engineering assets (e.g., bridges and aircraft). Practically the main consequence of non-detection of damage is that it leads to inadequate maintenance, which is further reinforced by the low consistency of damage registrations. For damage spots with reference severity 'bad', only half of the inspectors who detected registered it as 'bad', others were more lenient. A key factor that (partially) explains the inconsistency and non-detection of damage is the complexity of the used inspection guideline, in particular the large number of sometimes overlapping damage parameters, and the ambiguity in severity categories.

Based on the findings from Chapter 6, in Chapter 7 the effectiveness of I&M of a flood defence segment was evaluated using a Dynamic Bayesian Network (DBN). The focus was on three key aspects of I&M. First of all, it was determined what the effectiveness of the current maintenance concept is for a specific case study. From this it was found that if degradation and I&M are not accounted for, the failure probability for a flood defence segment of 20 sections (1 km length) was underestimated by approximately a factor 20. Obviously this factor can be different for other cases, but it demonstrates that failure probability estimates that do not account for damage are likely too optimistic. Secondly, it was evaluated what combination of structural upgrades and changes to the maintenance concept yield the largest cost reduction and increase in structural robustness to damage. This demonstrated that additional condition inspections in autumn are a very cost-effective measure for reducing the consequences of degradation on the system failure probability. Additionally, targeted structural upgrades to mitigate the impact of damage on the failure probability (e.g., thicker clay layer, measures to prevent animal burrowing) can significantly increase structural robustness as they limit the potential consequences of damage. Given uncertainties in the influence and occurrence probability of damage, structural robustness is found to be a valuable additional performance indicator next to Equivalent Annual Cost (or Total Cost). Finally, it was explored what requirements in terms of accuracy would need to be met for new inspection techniques to outperform the I&M schedules considered in the previous steps. This can help in evaluating new techniques such as inspections using drones. It is demonstrated that for both new and existing techniques, the correlation between observations from different inspections can significantly reduce the efficiency of an I&M schedule. This emphasizes the importance of efforts to ensure independence between inspections. Overall, it is demonstrated that not considering degradation, inspection and maintenance leads to optimistic reliability estimates and inefficient flood defence asset management decisions.

8.2. Recommendations

Following the conclusions provided in the preceding section there are several recommendations based on the research in this thesis. For each key topic we provide two types of recommendations: both for future research and more directly for engineering practice, such as the Dutch Flood Protection Programme (HWBP). It should be noted that in some cases these might relate or overlap.

8.2.1. Optimization of system design

Recommendations for future research

In Chapter 3 it has been demonstrated that it is possible to obtain solutions for flood defence system reinforcements with (near-)optimal total cost. While this ensures optimal reinforcement measures, investments in reducing uncertainty were not included in the analysis. In the example in Section 4.3 it was demonstrated that inclusion of uncertainty reduction can change optimal investment decisions for reinforcement. It is recommended to explore how uncertainty reduction can be included in the optimization approach from Chapter 3.

Another potential improvement is to explore how other objectives can be included. From the perspective of flood defence reinforcements this could for instance be requirements for sustainability, or valuing adaptive capacity of flood defence design. This might require extension of the considered time horizon, and inclusion of additional future investments.

Recommendations for engineering practice

The design approach in Chapter 3, as well as the investigations into the benefits of uncertainty reduction in Chapters 4 and 5 have clear benefits for flood defence design. It is recommended to adopt the system optimization approach from Chapter 3 as a tool in an iterative design approach for flood defence reinforcements. In such an approach, different iterations in the design process (e.g. new design solutions) can be evaluated based on their effect on a system level. Adopting such an approach, of which the practical feasibility has been demonstrated, can provide at least three major benefits: projects become much more cost-effective, have less impact on the environment, and the approach significantly increases the transparency of underlying decisions and assumptions and their effect on the preferred design alternative.

An additional application for which the approach is particularly suitable is the prioritization of flood defence reinforcement projects on a national or catchment scale, or more broadly: any prioritization of investments in flood risk reduction. Typically there are more projects than can be executed within a year. By considering many potential investments in individual flood defence sections or parts of segments and prioritizing them using the approach from Chapter 3, one can find a priority order than ensures that budgets are spent such that projects maximally contribute to reduction of flood risk on a catchment or national scale. This can improve the

societal benefits of flood protection investment programmes such as the Dutch Flood Protection Programme.

8.2.2. Reduction of time-invariant uncertainty

Recommendations for future research

In Chapters 4 and 5 we considered different methods for reducing time-invariant uncertainty. A limitation of these studies is that these only consider a cross section, while in practice these uncertain parameters might have large spatial variability as well. It is advised to further investigate the impact of spatial variability on monitoring outcomes, and how these translate to reliability of flood defence sections. A point of attention is for instance how to optimize sensor spacing such that the Value of Information (VoI) is optimized.

An important finding in Chapter 5 was that uncertainty reduction has a large influence on long term investments. This is typically not considered in studies into long term adaptation of flood defence systems. Considering that most flood defence structures (in the Netherlands) are disapproved for failure modes often dominated by parameters with large time-invariant uncertainty (mainly slope instability and piping erosion), it is likely that systematically including interventions to reduce uncertainty has a large influence on future investments. This would not only impact the level of investments to keep flood risk at an acceptable level, but also the range of scenarios for which this is economically feasible.

Recommendations for engineering practice

Both in Chapters 4 and 5 it was demonstrated that in many cases the VoI of uncertainty reduction (both pore pressure monitoring and proof load testing) is positive in most cases. Additionally, for instance Schweckendiek [8] found similar results, also for site investigation. Yet, in engineering practice this consideration is not always made. One reason why the uptake of uncertainty reduction as an alternative to structural reinforcement or adaptation is slow, might be funding arrangements: while (the majority of) reinforcement costs is funded, flood defence asset managers often have to fund the costs of uncertainty reduction themselves, if these are not a clear part of a reinforcement project. As such, uncertainty reduction is preferably applied in a brief period before a reinforcement is executed. As especially for pore pressure monitoring the VoI is dependent on the duration of monitoring, this significantly limits the VoI. Restructuring funding arrangements might go a long way in increasing the positive impact of uncertainty reduction interventions on flood defence asset management.

While in this thesis the VoI is computed as a combination of cost and flood risk reduction, practical considerations on investment in uncertainty reduction might not attach the same value to risk reduction or prevented investment costs as the direct costs of uncertainty reduction itself. In other words: funding arrangements and financial incentives might stimulate preferences that do not align with the societal benefits of uncertainty reduction as considered in this thesis (for extensive discussion on this topic see Kahneman & Tversky [69] and van Erp [70]). This can

mean that either the value of societal benefits should be evaluated differently (e.g., society is more risk-seeking or risk-averse), or that funding arrangements are not aligned with societal preferences. For efficient flood defence asset management it is imperative that funding arrangements, such as those for funding of flood defence reinforcements (e.g., in the Flood Protection Programme), stimulate behaviour that aligns well with increasing societal value. A practical improvement that can help is to not only quantify the benefits of uncertainty reduction as the expected reduction in Total Cost (i.e. VoI), but also the distribution of cost reduction (as done in Chapter 5), potentially split among different cost components as has been done in Chapter 4.

8.2.3. Managing uncertainty in time-variant processes

Recommendations for future research

This thesis provided the first ever experimental estimate for the accuracy of visual condition inspection of flood defences. This revealed that both the Probability of Detection (PoD) and consistency vary a lot per inspector and per type of damage, and that accuracy is lower than is commonly assumed by practitioners. It is recommended to further investigate the accuracy of (visual) inspections. This could for instance focus on a further comparison of different variables that are found to be important, such as the number of inspectors per inspection (e.g., single or pair), the scope of the inspection, and the inspection guidelines that are used. As has been demonstrated by research done in other applications (see Chapter 6 for examples), this can greatly improve the overall performance of flood defence inspectors. Additionally, such investigations can be used to demonstrate and optimize the added value of using technological support such as drones for inspections.

Inspections are only one piece of the puzzle of managing uncertainty in time-variant flood defence reliability. For most of the damages considered in condition inspections, there is very limited insight into the influence of such damage on flood defence reliability. Such insights are pivotal in risk-based Inspection & Maintenance (I&M) of flood defence systems. It is recommended to do more fundamental research into the impact of damage on reliability, in particular for revetments (also block/asphalt revetments), slope instability and internal erosion.

A final recommendation for future research relates to the assumptions on degradation behaviour in Chapter 7. With the available inspection data it is difficult to quantify some of the expert knowledge on degradation, such as that degradation rates vary in time, grass revetments might recover, and the influence of location-specific factors that increase or decrease the rates of certain types of degradation. Improved insight into degradation rates and patterns in time and space can be a huge improvement towards finding optimal I&M policies for flood defence systems.

Recommendations for engineering practice

One of the key issues addressed in the chapters on I&M (Chapters 6 and 7) is the distance between the practice of I&M and how this is addressed in reliability analysis of flood defence systems. On a general level it has been clearly illustrated

that accounting for the possibility of degradation in reliability analysis is essential to obtain realistic reliability estimates. Given the existing uncertainties in accounting for time-variant processes in reliability, practical implementation might seem a bridge to far. However, the following targeted steps and improvements can pave the way to gradually improve how time-variant uncertainty is taken into account in risk-based flood defence asset management:

- *Better understand degradation behaviour through structured and targeted collection of (inspection) data.*
  Currently, inspection data is incomplete and inconsistent due to the limitations of visual inspections. Additionally, data is only gathered at specific times (typically only in spring). As such, deriving seasonal patterns from this data is not possible. Insight in degradation behaviour is pivotal to understand why, when and how flood defences degrade, and what can be done about it.

- *Improve visual inspection of flood defences.*
  In this thesis several recommendations have been given to improve visual inspection of flood defences, both in terms of the inspection itself as in terms of determining maintenance concept. For inspections it is recommended to include an autumn inspection aimed at detecting damage to the flood defence. Currently autumn inspections are typically aimed at ensuring maintenance works have been done properly, but this should be broadened to a full condition inspection (if not done already). Additionally, it should be ensured that observations are as independent as practically feasible, for instance by rotating inspectors. Recommended actions for improving visual inspections are to make them more targeted (less parameters) and to invest in making guidelines more consistent (no overlap/conflict between parameters) and less complicated (less parameters). Additionally, it should be investigated how (formal) feedback mechanisms on inspector performance can be implemented, as this has shown to give tremendous improvement in other applications [e.g. 154]. For damage types that are difficult to detect by visual inspection it is advised to explore the potential added value of other means of inspection such as drones, potentially with automated defect detection. On that note, the findings from Chapter 7 demonstrate that such techniques do not need to have perfect accuracy in order to improve flood defence asset management as a whole.

- *Use specific performance metrics to evaluate the impact of uncertainty in time-variant processes.*
  In Chapter 7 structural robustness was used in order to evaluate the relative impact of damage to flood defence system reliability. Vulnerability [38] might be another indicator that can be used to assess this impact, and how the impact is reduced through changes in the maintenance concept as well as structural upgrades. This is particularly important as there is considerable uncertainty on estimates of degradation behaviour. An additional and specific performance metric can then support decision makers in decisions on risk-based I&M. It should be noted that structural robustness as used in this thesis

should only be used together with a performance indicator that represents the reliability of a flood defence system (e.g, failure probability or total flood risk cost).

8.3. Closing remarks

The main theme of this thesis is how to make (optimal) risk-based decisions in flood defence asset management. This revolves around 3 key topics, and in each of these topics, the approaches used have been demonstrated to both increase efficiency and effectiveness of decisions, and increase transparency in decision making.

Since 2017 new reliability requirements have been put on the central stage in flood defence asset management in the Netherlands. The novelty in these requirements is primarily that the definition of the requirements has changed. Consequently, the major challenge is to translate this definition into decisions and approaches that align with the safety philosophy behind these requirements: balancing flood risk and investment cost in flood defence systems, based on our best performance estimates, given all relevant uncertainties. This approach as a whole provides great potential for decision optimization.

In the examples in this thesis it has been demonstrated that the application of these principles is currently not at the level it should be in a country for which reliable flood defence systems are existential: system optimization of flood defence reinforcements is nowhere near common practice, and issues encountered in Inspection & Maintenance (I&M) are typically not accounted for in reliability estimates. With the insights presented in this thesis these topics should now become a part of our reliability estimates, as well as our asset management decisions: they are within the realm of our knowledge.

Obviously, there are unresolved issues and uncertainties in the approaches used, and how to apply these in practice. This might lead to the reasoning that knowledge is too uncertain to be applied, and that we should refrain from using it until it has matured further. However, from the perspective of risk-based decisions under uncertainty this is an invalid argument that often stems from the prospect of 'fear of taking the wrong decision'. However, can a decision where all uncertainties were carefully weighted and considered really be wrong? Not with the mindset conveyed in the quote by Herodotus at the beginning of this chapter. It is this mindset that could really make us utilize the full potential of a risk-based approach to ensure efficient and effective asset management of flood defence systems, now and in the future.

Bibliography

1. Winsemius, H. C., Aerts, J. C. J. H., van Beek, L. P. H., Bierkens, M. F. P., Bouwman, A., Jongman, B., Kwadijk, J. C. J., Ligtvoet, W., Lucas, P. L., van Vuuren, D. P. & Ward, P. J. Global drivers of future river flood risk. *Nature Climate Change* **6,** 381–385. doi:10.1038/nclimate2893 (2016).
2. Jongman, B., Ward, P. J. & Aerts, J. C. Global exposure to river and coastal flooding: Long term trends and changes. *Global Environmental Change* **22,** 823–835. doi:10.1016/j.gloenvcha.2012.07.004 (2012).
3. CIRIA. *International Levee Handbook* isbn: 9780860177340 (2013).
4. ISO. *NEN-ISO 55000: Asset management - Overview, principles and terminology* tech. rep. (Nederlands Normalisatie-instituut, Delft, 2014).
5. Kind, J. Economically efficient flood protection standards for the Netherlands. *Journal of Flood Risk Management* **7,** 103–117. doi:10.1111/jfr3.12026 (2014).
6. Jongejan, R., Diermanse, F., Kanning, W. & Bottema, M. Reliability-based partial factors for flood defenses. *Reliability Engineering & System Safety* **193,** 106589. doi:10.1016/j.ress.2019.106589 (2020).
7. Vrijling, J. Probabilistic design of water defense systems in The Netherlands. *Reliability Engineering & System Safety* **74,** 337–344. doi:10.1016/S0951-8320(01)00082-5 (2001).
8. Schweckendiek, T. *On Reducing Piping Uncertainties - A Bayesian Decision Approach* PhD thesis (Delft University of Technology, 2014). isbn: 9789053358801. doi:doi.org/10.4233/uuid:f9be2f7e-7009-4c73-afe5-8b4bb16e956f.
9. Vonk, B., Klerk, W. J., Fröhle, P., Gersonius, B., den Heijer, F., Jordan, P., Radu Ciocan, U., Rijke, J., Sayers, P. & Ashley, R. Adaptive Asset Management for Flood Protection: The FAIR Framework in Action. *Infrastructures* **5,** 109. doi:10.3390/infrastructures5120109 (2020).
10. Sayers, P., Gersonius, B., den Heijer, F., Klerk, W. J., Fröhle, P., Jordan, P., Ciocan, U. R., Rijke, J., Vonk, B. & Ashley, R. Towards adaptive asset management in flood risk management: A policy framework. *Water Security* **12,** 100085. doi:10.1016/j.wasec.2021.100085 (2021).
11. Frangopol, D. M. Life-cycle performance, management, and optimisation of structural systems under uncertainty: accomplishments and challenges 1. en. *Structure and Infrastructure Engineering* **7,** 389–413. doi:10.1080/15732471003594427 (2011).

12. Biondini, F. & Frangopol, D. M. Life-Cycle Performance of Deteriorating Structural Systems under Uncertainty: Review. *Journal of Structural Engineering* **142,** F4016001. doi:10.1061/(ASCE)ST.1943-541X.0001544 (2016).
13. Sanchez-Silva, M., Klutke, G. A. & Rosowsky, D. V. Life-cycle performance of structures subject to multiple deterioration mechanisms. *Structural Safety* **33,** 206–217. doi:10.1016/j.strusafe.2011.03.003 (2011).
14. Kammouh, O., Nogal, M., Binnekamp, R. & Wolfert, A. R. Multi-system intervention optimization for interdependent infrastructure. *Automation in Construction* **127.** doi:10.1016/j.autcon.2021.103698 (2021).
15. Dupuits, E., Schweckendiek, T. & Kok, M. Economic optimization of coastal flood defense systems. *Reliability Engineering & System Safety* **159,** 143–152. doi:10.1016/J.RESS.2016.10.027 (2017).
16. Eijgenraam, C., Brekelmans, R., Den Hertog, D. & Roos, K. Optimal strategies for flood prevention. *Management Science* **63,** 1644–1656. doi:10.1287/mnsc.2015.2395 (2017).
17. Voortman, H. *Risk-based design of large-scale flood defence systems* PhD Thesis (Delft University of Technology, 2003). http://resolver.tudelft.nl/uuid:31d3672a-0062-465d-b30a-bacc2ed4b79d.
18. Zwaneveld, P., Verweij, G. & van Hoesel, S. Safe dike heights at minimal costs: An integer programming approach. *European Journal of Operational Research* **270,** 294–301. doi:10.1016/j.ejor.2018.03.012 (2018).
19. Bischiniotis, K., Kanning, W., Jonkman, S. & Kok, M. Cost-optimal design of river dikes using probabilistic methods. *Journal of Flood Risk Management* **11,** S1002–S1014. doi:10.1111/jfr3.12277 (2018).
20. Schweckendiek, T., van der Krogt, M., Teixeira, A., Kanning, W., Brinkman, R. & Rippi, K. *Reliability Updating with Survival Information for Dike Slope Stability Using Fragility Curves* in *Geo-Risk 2017* (American Society of Civil Engineers, Reston, VA, 2017), 494–503. doi:10.1061/9780784480700.047.
21. Schweckendiek, T. & Vrouwenvelder, A. Reliability updating and decision analysis for head monitoring of levees. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* **7,** 110–121. doi:10.1080/17499518.2013.791034 (2013).
22. Thöns, S. On the Value of Monitoring Information for the Structural Integrity and Risk Management. *Computer-Aided Civil and Infrastructure Engineering* **33,** 79–94. doi:10.1111/mice.12332 (2018).
23. Zonta, D., Glisic, B. & Adriaenssens, S. Value of information: impact of monitoring on decision-making. *Structural Control and Health Monitoring* **21,** 1043–1056. doi:10.1002/stc.1631 (2014).

24. Frangopol, D. M. & Liu, M. Maintenance and management of civil infrastructure based on condition, safety, optimization, and life-cycle cost. *Structure and Infrastructure Engineering* **3,** 29–41. doi:10.1080/15732470500253164 (2007).

25. Straub, D. *Generic approaches to risk based inspection planning for steel structures* PhD Thesis (ETH Zurich, 2004). https://doi.org/10.3929/ethz-a-004783633.

26. Luque, J. & Straub, D. Risk-based optimal inspection strategies for structural systems using dynamic Bayesian networks. *Structural Safety* **76,** 68–80. doi:10.1016/j.strusafe.2018.08.002 (2019).

27. Morato, P., Papakonstantinou, K., Andriotis, C., Nielsen, J. & Rigo, P. Optimal inspection and maintenance planning for deteriorating structural components through dynamic Bayesian networks and Markov decision processes. *Structural Safety* **94,** 102140. doi:10.1016/j.strusafe.2021.102140 (2022).

28. Palladino, M. R., Barbetta, S., Camici, S., Claps, P. & Moramarco, T. Impact of animal burrows on earthen levee body vulnerability to seepage. *Journal of Flood Risk Management* **13,** 1–21. doi:10.1111/jfr3.12559 (2020).

29. Klerk, W. J., Schweckendiek, T., den Heijer, F. & Kok, M. Value of Information of Structural Health Monitoring in Asset Management of Flood Defences. *Infrastructures* **4,** 56. doi:10.3390/infrastructures4030056 (2019).

30. Amadi-Echendu, J. E., Brown, K., Willett, R. & Mathew, J. *Definitions, Concepts and Scope of Engineering Asset Management* 416. isbn: 1849961786 (Springer Science & Business Media, 2010).

31. Jonkman, S. N., Voortman, H. G., Klerk, W. J. & van Vuren, S. Developments in the management of flood defences and hydraulic infrastructure in the Netherlands. *Structure and Infrastructure Engineering,* 1–16. doi:10.1080/15732479.2018.1441317 (2018).

32. Blanchard, B. S. & Fabrycky, W. J. *Systems Engineering and Analysis* Fifth Edit, 148–162. isbn: 978-0132217354 (Pearson/Prentice Hall, Upper Saddle River, NJ, USA, 2011).

33. Stapelberg, F. R. *Handbook of Reliability, Availability, Maintainability and Safety in Engineering Design* 148–162. isbn: 978-1-84800-174-9. doi:10.1007/978-1-84800-175-6 (Springer London, London, 2009).

34. Van Dantzig, D. Economic Decision Problems for Flood Prevention. *Econometrica* **24,** 276. doi:10.2307/1911632 (1956).

35. Kok, M., Jongejan, R., Nieuwjaar, M. & Tanczos, I. *Fundamentals of Flood Protection* isbn: 978-90-8902-160-1 (2017).

36. Jonkman, S., van Gelder, P. & Vrijling, J. An overview of quantitative risk measures for loss of life and economic damage. *Journal of Hazardous Materials* **99,** 1–30. doi:10.1016/S0304-3894(02)00283-2 (2003).

37. Van Alphen, J. The Delta Programme and updated flood risk management policies in the Netherlands. *Journal of Flood Risk Management* **9,** 310–319. doi:10.1111/jfr3.12183 (2016).

38. Lind, N. C. A measure of vulnerability and damage tolerance. *Reliability Engineering & System Safety* **48,** 1–6. doi:10.1016/0951-8320(95)00007-O (1995).

39. Baker, J. W., Schubert, M. & Faber, M. H. On the assessment of robustness. *Structural Safety* **30,** 253–267. doi:10.1016/j.strusafe.2006.11.004 (2008).

40. European Committee for Standardization. *EN 1991-1-7 Eurocode 1 - Actions on structures - Part 1-7: General actions - Accidental actions* tech. rep. (Brussels, 2006).

41. Maier, H. R., Guillaume, J. H., van Delden, H., Riddell, G. A., Haasnoot, M. & Kwakkel, J. H. An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environmental Modelling and Software* **81,** 154–164. doi:10.1016/j.envsoft.2016.03.014 (2016).

42. Mens, M. J., Klijn, F., de Bruijn, K. M. & van Beek, E. The meaning of system robustness for flood risk management. *Environmental Science and Policy* **14,** 1121–1131. doi:10.1016/j.envsci.2011.08.003 (2011).

43. Van Gelder, P. *Statistical methods for the risk-based design of civil structures* en. PhD thesis (2000). isbn: 9090134522. http://resolver.tudelft.nl/uuid:6a62d6fa-cbcc-4c38-af8a-027c3d191a9d.

44. Kiureghian, A. D. & Ditlevsen, O. Aleatory or epistemic? Does it matter? *Structural Safety* **31,** 105–112. doi:10.1016/j.strusafe.2008.06.020 (2009).

45. Phoon, K. K., Prakoso, W. A., Wang, Y. & Ching, J. in *Reliability of Geotechnical Structures in ISO2394* (eds Phoon, K. K. & Retief, J.) 1st ed., 49–87 (CRC Press/Balkema, Leiden, 2016).

46. Walker, W., Harremoës, P., Rotmans, J., van der Sluijs, J., van Asselt, M., Janssen, P. & Krayer von Krauss, M. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment* **4,** 5–17. doi:10.1076/iaij.4.1.5.16466 (2003).

47. Raiffa, H. & Schlaifer, R. *Applied statistical decision theory* isbn: 978-0-471-38349-9 (Cambridge University Press, Cambridge, Mass, 1961).

48. Van der Krogt, M. G., Schweckendiek, T. & Kok, M. Uncertainty in spatial average undrained shear strength with a site-specific transformation model. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* **13,** 226–236. doi:10.1080/17499518.2018.1554820 (2019).

49. Kanning, W. *The Weakest Link: Spatial Variability in the Piping Failure Mechanism of Dikes* en. PhD thesis (Delft University of Technology, 2012). isbn: 9789461860880. doi:https://doi.org/10.4233/uuid:5fb7b121-dc00-48aa-bda2-b163f10513bf.

50. Roscoe, K., Diermanse, F. & Vrouwenvelder, T. System reliability with correlated components: Accuracy of the Equivalent Planes method. *Structural Safety* **57,** 53–64. doi:10.1016/j.strusafe.2015.07.006 (2015).

51. Vanmarcke, E. H. Matrix formulation of reliability analysis and reliability-based design. *Computers and Structures* **3,** 757–770. doi:10.1016/0045-7949(73)90056-4 (1973).

52. Ministerie van Infrastructuur en Milieu. *Regeling veiligheid primaire waterkeringen 2017 Bijlage III Sterkte en veiligheid (in Dutch)* tech. rep. (2017). https://www.helpdeskwater.nl/publish/pages/157017/bijlage%7B%5C_%7Diii%7B%5C_%7Dsterkte%7B%5C_%7Den%7B%5C_%7Dveiligheid%7B%5C_%7Dregeling%7B%5C_%7Dveiligheid%7B%5C_%7Dprimaire%7B%5C_%7Dwaterkeringen%7B%5C_%7D2017.pdf.

53. Straub, D., Schneider, R., Bismut, E., Kim, H.-j., Risk, E., Group, A. & München, T. U. Reliability analysis of deteriorating structural systems. *Structural Safety* **82,** 101877. doi:10.1016/j.strusafe.2019.101877 (2020).

54. Moubray, J. *Reliability-Centered Maintenance* 2nd, 440. isbn: 9780750633581 (Butterworth-Heinemann, 1999).

55. Kottegoda, N. & Rosso, R. *Applied statistics for civil and environmental engineers* 2nd. isbn: 9781405179171 (Blackwell Publishing Ltd, 2008).

56. European Committee for Standardization. *Eurocode 0: Basis of structural design* tech. rep. December 2008 (Brussels, 2002), 89.

57. Holicky, M., Diamantidis, D. & Sykora, M. *Determination of Target Safety for Structures* in *12th International Conference on Applications of Statistics and Probability in Civil Engineering* (2015), 1–8. doi:10.14288/1.0076258.

58. Vrouwenvelder, A. Developments towards full probabilistic design codes. *Structural Safety* **24,** 417–432. doi:10.1016/S0167-4730(02)00035-8 (2002).

59. Roubos, A., Steenbergen, R., Schweckendiek, T. & Jonkman, S. Risk-based target reliability indices for quay walls. *Structural Safety* **75,** 89–109. doi:10.1016/j.strusafe.2018.06.005 (2018).

60. Finkelstein, M. *Failure Rate Modelling for Reliability and Risk* Springer S, 302. isbn: 1848009860. doi:10.1007978-1-84800-986-8 (Springer London, 2008).

61. Klerk, W. J., Kanning, W., Vos, R. J. & Kok, M. *Risk-based maintenance of asphalt revetments on flood defences* in *Life-Cycle Civil Engineering: Innovation, Theory and Practice: Proceedings of the 7th International Symposium on Life-Cycle Civil Engineering (IALCCE 2020)* (eds Chen, A., Ruan, X. & Frangopol, D. M.) (CRC Press / Balkema - Taylor & Francis Group, Shanghai, China, 2020).

62. Buijs, F., Hall, J., Sayers, P. & Van Gelder, P. Time-dependent reliability analysis of flood defences. *Reliability Engineering & System Safety* **94,** 1942–1953. doi:10.1016/j.ress.2009.06.012 (2009).

63. Pol, J. C., Kanning, W. & Jonkman, S. N. Temporal Development of Backward Erosion Piping in a Large-Scale Experiment. *Journal of Geotechnical and Geoenvironmental Engineering* **147,** 04020168. doi:10.1061/(ASCE)GT.1943-5606.0002415 (2021).

64. Pol, J., Kanning, W. & Jonkman, B. *Strength Recovery in a Large Scale Backward Erosion Piping Experiment* in *European Working Group on Internal Erosion* (submitted, 2022).

65. Koelewijn, A., De Vries, G., van Lottum, H., Förster, U., van Beek, V. & Bezuijen, A. *Full-scale testing of piping prevention measures : Three tests at the IJkdijk* in *Physical Modelling in Geotechnics* (2014), 891–897. isbn: 9781138001527. doi:10.1201/b16200-124.

66. Van der Krogt, M. G., Schweckendiek, T. & Kok, M. Improving dike reliability estimates by incorporating construction survival. *Engineering Geology* **280,** 105937. doi:10.1016/j.enggeo.2020.105937 (2021).

67. Mansur, C. I., Postol, G. & Salley, J. R. Performance of relief well systems along Mississippi River levees. *Journal of Geotechnical and Geoenvironmental Engineering* **126,** 727–738. doi:10.1061/(ASCE)1090-0241(2000)126:8(727) (2000).

68. Von Neumann, J. & Morgenstern, O. *Theory of games and economic behavior* isbn: 0691119937. doi:10.2307/3610940 (Princeton University Press, Princeton, NJ, 1947).

69. Kahneman, D. & Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* **47,** 263. doi:10.2307/1914185 (1979).

70. Van Erp, N. *A Bayesian framework for risk perception* PhD thesis (2017). isbn: 9789055841745. doi:https://doi.org/10.4233/uuid:1ff6ae46-c2bd-4375-aeb1-a4a9313ec560.

71. Dupuits, E. J. C., Klerk, W. J., Schweckendiek, T. & de Bruijn, K. M. Impact of including interdependencies between multiple riverine flood defences on the economically optimal flood safety levels. *Reliability Engineering & System Safety* **191.** doi:10.1016/j.ress.2019.04.028 (2019).

72. Lindley, D. V. *Bayesian Statistics: A Review* 6th. doi:10.2307/2490159 (Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1972).

73. Straub, D. & Faber, M. H. Risk based inspection planning for structural systems. *Structural Safety* **27,** 335–355. doi:10.1016/j.strusafe.2005.04.001 (2005).

74. Lempert, R. J. *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis* 208. isbn: 0833034855. doi:10.1016/j.techfore.2003.09.006 (2003).

75. Kwakkel, J. H., Haasnoot, M. & Walker, W. E. Developing dynamic adaptive policy pathways: a computer-assisted approach for developing adaptive strategies for a deeply uncertain world. *Climatic Change* **132,** 373–386. doi:10.1007/s10584-014-1210-4 (2015).

76. Giuliani, M. & Castelletti, A. Is robustness really robust? How different definitions of robustness impact decision-making under climate change. *Climatic Change* **135,** 409–424. doi:10.1007/s10584-015-1586-9 (2016).

77. Tversky, A. & Kahneman, D. Judgment under Uncertainty: Heuristics and Biases. *Science* **185,** 1124–1131. doi:10.1126/science.185.4157.1124 (1974).

78. Schöbi, R. & Chatzi, E. N. Maintenance planning using continuous-state partially observable Markov decision processes and non-linear action models. *Structure and Infrastructure Engineering* **12,** 977–994. doi:10.1080/15732479.2015.1076485 (2016).

79. Jensen, F. V. & Nielsen, T. D. *Bayesian Networks and Decision Graphs* (eds Jordan, M., Kleinberg, J. & Scholkopf, B.) isbn: 0-387-68281-3 (Springer, New York, 2007).

80. Luque, J. & Straub, D. *Algorithms for optimal risk-based planning of inspections using influence diagrams* in *Proceedings of the 11th International Probabilistic Workshop* (Brno, Czech Republic, 2013), 247–261.

81. Bismut, E. & Straub, D. Optimal adaptive inspection and maintenance planning for deteriorating structural systems. *Reliability Engineering & System Safety* **215,** 107891. doi:10.1016/j.ress.2021.107891 (2021).

82. Saltelli, A. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications* **145,** 280–297. doi:10.1016/S0010-4655(02)00280-1 (2002).

83. Woodward, M., Kapelan, Z. & Gouldby, B. Adaptive flood risk management under climate change uncertainty using real options and optimization. *Risk Analysis* **34,** 75–92. doi:10.1111/risa.12088 (2014).

84. Van den Boomen, M., Spaan, M. T., Schoenmaker, R. & Wolfert, A. R. Untangling decision tree and real options analyses: a public infrastructure case study dealing with political decisions, structural integrity and price uncertainty. *Construction Management and Economics* **37,** 24–43. doi:10.1080/01446193.2018.1486510 (2019).

85. Klerk, W. J., Kanning, W., Kok, M. & Wolfert, R. Optimal planning of flood defence system reinforcements using a greedy search algorithm. *Reliability Engineering & System Safety* **207,** 107344. doi:10.1016/j.ress.2020.107344 (2021).

86. Chen, H.-p. & Mehrabani, M. B. Reliability analysis and optimum maintenance of coastal flood defences using probabilistic deterioration modelling. *Reliability Engineering & System Safety* **185,** 163–174. doi:10.1016/j.ress.2018.12.021 (2019).

87. Jongejan, R. & Maaskant, B. *The Use of Quantitative Risk Analysis for Prioritizing Flood Risk* in *CDA 2013 Annual Conference* (Montreal, Quebec, 2013).

88. Bohnenblust, H. & Slovic, P. Integrating technical analysis and public values in risk-based decision making. *Reliability Engineering & System Safety* **59,** 151–159. doi:10.1016/S0951-8320(97)00136-1 (1998).

89. Lengkeek, H. J., Post, M., Breedeveld, J. & Naves, T. Eemdijk full-scale field test programme : ground dike and sheet pile dike failure test. *Proceedings of the XVII ECSMGE-2019.* doi:10.32075/17ECSMGE-2019-0454 (2019).

90. Dupuits, E. J. C. *Optimal System Safety Targets - Incorporating Hydrodynamic Interactions in an Economic Cost-Benefit Analysis for Flood Defence Systems* PhD thesis (2019), 178. isbn: 9789055841745. doi:10.4233/uuid.

91. Voortman, H. & Vrijling, J. Optimal design of flood defence systems in a changing climate. *Heron* **49** (2004).

92. Brekelmans, R., Den Hertog, D., Roos, K. & Eijgenraam, C. Safe dike heights at minimal costs: The nonhomogeneous case. *Operations Research* **60,** 1342–1355. doi:10.1287/opre.1110.1028 (2012).

93. Haasnoot, M., Kwadijk, J., Van Alphen, J., Le Bars, D., Van Den Hurk, B., Diermanse, F., Van Der Spek, A., Oude Essink, G., Delsman, J. & Mens, M. Adaptation to uncertain sea-level rise; how uncertainty in Antarctic mass-loss impacts the coastal adaptation strategy of the Netherlands. *Environmental Research Letters* **15.** doi:10.1088/1748-9326/ab666c (2020).

94. Woodward, M., Gouldby, B., Kapelan, Z. & Hames, D. Multiobjective optimization for improved management of flood risk. *Journal of Water Resources Planning and Management* **140,** 201–215. doi:10.1061/(ASCE)WR.1943-5452.0000295 (2014).

95. Ciullo, A., de Bruijn, K. M., Kwakkel, J. H. & Klijn, F. Accounting for the uncertain effects of hydraulic interactions in optimising embankments heights: Proof of principle for the IJssel River. *Journal of Flood Risk Management* **12,** 1–15. doi:10.1111/jfr3.12532 (2019).

96. Ciullo, A., Kwakkel, J. H., De Bruijn, K. M., Doorn, N. & Klijn, F. Efficient or Fair? Operationalizing Ethical Principles in Flood Risk Management: A Case Study on the Dutch-German Rhine. *Risk Analysis.* doi:10.1111/risa.13527 (2020).

97. Kollat, J. B. & Reed, P. M. The value of online adaptive search: A performance comparison of NSGAII, ϵ-NSGAII and ϵMOEA. *Lecture Notes in Computer Science* **3410,** 386–398. doi:10.1007/978-3-540-31880-4_27 (2005).

98. Bocchini, P. & Frangopol, D. M. A probabilistic computational framework for bridge network optimal maintenance scheduling. *Reliability Engineering & System Safety* **96,** 332–349. doi:10.1016/J.RESS.2010.09.001 (2011).

99. Barone, G. & Frangopol, D. M. Life-cycle maintenance of deteriorating structures by multi-objective optimization involving reliability, risk, availability, hazard and cost. *Structural Safety* **48,** 40–50. doi:10.1016/j.strusafe.2014.02.002 (2014).

100. Cavdaroglu, B., Hammel, E., Mitchell, J. E., Sharkey, T. C. & Wallace, W. A. Integrating restoration and scheduling decisions for disrupted interdependent infrastructure systems. *Annals of Operations Research* **203,** 279–294. doi:10.1007/s10479-011-0959-3 (2013).

101. Bagloee, S. A., Sarvi, M., Patriksson, M. & Asadi, M. Optimization for Roads' Construction: Selection, Prioritization, and Scheduling. *Computer-Aided Civil and Infrastructure Engineering* **33,** 833–848. doi:10.1111/mice.12370 (2018).

102. Leontaris, G., Morales-Nápoles, O., Dewan, A. & Wolfert, A. R. M. R. Decision support for offshore asset construction using expert judgments for supply disruptions risk. *Automation in Construction* **107,** 102903. doi:10.1016/j.autcon.2019.102903 (2019).

103. Poshdar, M., González, V. A., Raftery, G. M., Orozco, F. & Cabrera-guerrero, G. G. A multi-objective probabilistic-based method to determine optimum allocation of time buffer in construction schedules. *Automation in Construction* **92,** 46–58. doi:10.1016/j.autcon.2018.03.025 (2018).

104. IBM. *IBM ILOG CPLEX Optimization Studio CPLEX User's Manual* tech. rep. (2019).

105. Mitchell, J. E. Branch-and-Cut Algorithms for Combinatorial Optimization Problems. *Handbook of Applied Optimization,* 65–77 (2002).

106. Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C. *Introduction to Algorithms, Third Edition* 3rd. isbn: 0262033844, 9780262033848 (The MIT Press, 2009).

107. Špačková, O. & Straub, D. Cost-Benefit Analysis for Optimization of Risk Protection Under Budget Constraints. *Risk Analysis* **35,** 941–959. doi:10.1111/risa.12310 (2015).

108. Ministerie van Infrastructuur en Milieu. *Achtergronden bij de normering van de primaire waterkeringen in Nederland (in Dutch)* tech. rep. (2016). https://www.helpdeskwater.nl/publish/pages/132791/normering%7B%5C_%7Dprimaire%7B%5C_%7Dwaterkeringen%7B%5C_%7Dhoofdrapport%7B%5C_%7D28juni2016.pdf.

109. Wong, T. E. & Keller, K. Deep Uncertainty Surrounding Coastal Flood Risk Projections: A Case Study for New Orleans. *Earth's Future* **5.** doi:10.1002/eft2.255 (2017).

110. Jevrejeva, S., Frederikse, T., Kopp, R. E., Le Cozannet, G., Jackson, L. P. & van de Wal, R. S. *Probabilistic Sea Level Projections at the Coast by 2100* 2019. doi:10.1007/s10712-019-09550-y.

111. Deltares. *KOSten voor versterking WATerkeringen. KOSWAT, systeemdocumentatie.* tech. rep. (2014). https://www.helpdeskwater.nl/publish/pages/130414/koswat%7B%5C_%7Dv2%7B%5C_%7D3%7B%5C_%7D-%7B%5C_%7Dsysteemdocumentatie.pdf.

112. Van der Krogt, M. G., Klerk, W. J., Kanning, W., Schweckendiek, T. & Kok, M. Value of information of combinations of proof loading and pore pressure monitoring for flood defences. *Structure and Infrastructure Engineering,* 1–16. doi:10.1080/15732479.2020.1857794 (2020).

113. Klerk, W. J., Kanning, W., van Veen, N.-J. & Kok, M. *Influence of monitoring on investment planning of flood defence systems* in *Proceedings of the 7th International Symposium on Geotechnical Safety and Risk (ISGSR 2019): State-of-the-Practice in Geotechnical Safety and Risk* (eds Ching, J., Li, D.-Q. & Zhang, J.) (Research Publishing, Singapore, Taipei, Taiwan, 2019), 792–797. doi:10.3850/978-981-11-2725-0_IS4-10-cd.

114. Phoon, K. K. & Retief, J. *Reliability of Geotechnical Structures in ISO2394* 1st ed. (eds Phoon, K. K. & Retief, J.) isbn: 9781138029118 (CRC Press/Balkema, Leiden, 2016).

115. Zhang, L. Reliability Verification Using Proof Pile Load Tests. *Journal of Geotechnical and Geoenvironmental Engineering* **130,** 1203–1213. doi:10.1061/(ASCE)1090-0241(2004)130:11(1203) (2004).

116. Frangopol, D. M., Strauss, A. & Kim, S. Use of monitoring extreme data for the performance prediction of structures: General approach. *Engineering Structures* **30,** 3644–3653. doi:10.1016/j.engstruct.2008.06.010 (2008).

117. Zhang, J., Zhang, L. M. & Tang, W. H. Slope Reliability Analysis Considering Site-Specific Performance Information. *Journal of Geotechnical and Geoenvironmental Engineering* **137,** 227–238. doi:10.1061/(ASCE)GT.1943-5606.0000422 (2011).

118. Li, D.-Q., Zhang, F.-P., Cao, Z.-J., Tang, X.-S. & Au, S.-K. Reliability sensitivity analysis of geotechnical monitoring variables using Bayesian updating. *Engineering Geology* **245,** 130–140. doi:10.1016/j.enggeo.2018.07.026 (2018).

119. Schweckendiek, T., Vrouwenvelder, A. & Calle, E. Updating piping reliability with field performance observations. *Structural Safety* **47,** 13–23. doi:10.1016/j.strusafe.2013.10.002 (2014).

120. Diamantidis, D., Sykora, M. & Sousa, H. *Quantifying the Value of Structural Health Information for Decision Support: Guide for Practising Engineers* tech. rep. (2019), 1–61. https://www.cost-tu1402.eu/-/media/sites/cost-tu1402/documents/deliverables/guidelines/cost-action-tu1402-guidelines%7B%5C_%7Dpract%7B%5C_%7Dengineers.pdf.

121. Sousa, H., Wenzel, H. & Thöns, S. *Quantifying the Value of Structural Health Information for Decision Support: Guide for Operators* tech. rep. (2019). https://www.cost-tu1402.eu/-/media/sites/cost-tu1402/documents/deliverables/guidelines/tu1402-guide-for-operators-v2.pdf.

122. Klerk, W. J., Kanning, W., van der Meer, M. T. & Nieuwenhuis, J. W. *Structural health monitoring in life-cycle management of dikes: a case study in the north of the Netherlands* in *Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of the Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016)* (eds Bakker, J., Frangopol, D. M. & van Breugel, K.) (CRC Press / Balkema - Taylor & Francis Group, Delft, the Netherlands, 2016). isbn: 9781138028470.

123. Spross, J. & Johansson, F. When is the observational method in geotechnical engineering favourable? *Structural Safety* **66,** 17–26. doi:10.1016/j.strusafe.2017.01.006 (2017).

124. Thöns, S. *Quantifying the Value of Structural health Information for Decision Support: Guide for Scientists* tech. rep. (COST Action TU1402, 2019). https://www.cost-tu1402.eu/-/media/sites/cost-tu1402/documents/deliverables/guidelines/tu1402-guide-for-scientists-v3.pdf.

125. Tavenas, F., Mieussens, C. & Bourges, F. Lateral displacements in clay foundations under embankments. *Canadian Geotechnical Journal* **16,** 532–550. doi:10.1139/t79-059 (1979).

126. Schofield, A. & Wroth, P. *Critical State Soil Mechanics* (McGraw Hill, Maidenhead, 1968).

127. Ladd, C. & Foott, R. New Design Procedure for Stability of Soft Clay. *Journal of the Geotechnical Engineering Division* **100,** 763–786 (1974).

128. Rijkswaterstaat. *Schematiseringshandleiding Macrostabiliteit WBI-2017 version 3.0, in Dutch* tech. rep. (Ministerie van Infrastructuur en Milieu, 2019). https://www.helpdeskwater.nl/publish/pages/157033/sh%7B%5C_%7Dmacrostabiliteit%7B%5C_%7D3%7B%5C_%7D0%7B%5C_%7D1.pdf.

129. Van Hoven, A. & Noordam, A. *POVM Infiltratieproef II -Analyse Infiltratieproef IJsseldijk (in Dutch)* tech. rep. (POV Macrostabiliteit, 2018). https://www.povmacrostabiliteit.nl/wp-content/uploads/2019/12/POVM-Analyse-Infiltratieproef-IJsseldijkII%7B%5C_%7Dv03-definitief.pdf.

130. Kanning, W. & van der Krogt, M. *Memo: Pore water pressure uncertainties for slope stability* tech. rep. (Deltares, 2016). http://publications.deltares.nl/1230090%7B%5C_%7D034.pdf.

131. Bruins Slot, H. *Enhancing reliability of dikes - an approach for assessing benefits of pore pressure monitoring and pressure relief wells in spatially variable soils* MSc. thesis (Delft University of Technology, 2021).

132. Le Bars, D., Drijfhout, S. & De Vries, H. A high-end sea level rise probabilistic projection including rapid Antarctic ice sheet mass loss. *Environmental Research Letters* **12.** doi:10.1088/1748-9326/aa6512 (2017).

133. Ko, J. M. & Ni, Y. Q. Technology developments in structural health monitoring of large-scale bridges. *Engineering Structures* **27,** 1715–1725. doi:10.1016/j.engstruct.2005.02.021 (2005).

134. Thöns, S. *Monitoring Based Condition Assessment of Offshore Wind Turbine Support Structures* PhD. thesis (ETH Zurich, 2012). isbn: 9783905826111. doi:ethz-b-000265761.

135. Pyayt, A. L., Kozionov, A. P., Mokhov, I. I., Lang, B., Meijer, R. J., Krzhizhanovskaya, V. V. & Sloot, P. M. Time-frequency methods for structural health monitoring. *Sensors* **14,** 5147–5173. doi:10.3390/s140305147 (2014).

136. Jongejan, R. B. & Maaskant, B. Quantifying flood risks in the Netherlands. *Risk analysis : an official publication of the Society for Risk Analysis* **35,** 252–64. doi:10.1111/risa.12285 (2015).

137. Klerk, W. J., den Heijer, F. & Schweckendiek, T. *Value of information in lifecycle management of flood defences* in *Safety and Reliability of Complex Engineered Systems. Proceedings of the ESREL 2015 conference* (eds Sudret, B., Podofillini, L., Stojadinovic, B., Zio, E. & Kroger, W.) (CRC Press, Zurich, Switzerland, 2015), 931–938.

138. Bachmann, D., Huber, N. P., Johann, G. & Schüttrumpf, H. Fragility curves in operational dike reliability assessment. *Georisk* **7,** 49–60. doi:10.1080/17499518.2013.767664 (2013).

139. Edwards, T. L., Brandon, M. A., Durand, G., Edwards, N. R., Golledge, N. R., Holden, P. B., Nias, I. J., Payne, A. J., Ritz, C. & Wernecke, A. Revisiting Antarctic ice loss due to marine ice-cliff instability. *Nature* **566,** 58–64. doi:10.1038/s41586-019-0901-4 (2019).

140. Bassis, J. N., Berg, B., Crawford, A. J. & Benn, D. I. Transition to marine ice cliff instability controlled by ice thickness gradients and velocity. *Science* **372,** 1342–1344. doi:10.1126/science.abf6271 (2021).

141. Werkgroep Discontovoet. *Rapport werkgroep discontovoet 2020* tech. rep. (The Hague, The Netherlands, 2020), 95. https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/kamerstukken/2020/11/10/rapport-werkgroep-discontovoet-2020/rapport-werkgroep-discontovoet-2020.pdf.

142. Hanemann, W. M. Information and the concept of option value. *Journal of Environmental Economics and Management* **16,** 23–37. doi:10.1016/0095-0696(89)90042-9 (1989).

143. Hohenbichler, M. & Rackwitz, R. First-order concepts in system reliability. *Structural Safety* **1,** 177–188. doi:10.1016/0167-4730(82)90024-8 (1983).

144. DeConto, R. M. & Pollard, D. Contribution of Antarctica to past and future sea-level rise. *Nature* **531,** 591–597. doi:10.1038/nature17145 (2016).

145. Nerem, R. S., Beckley, B. D., Fasullo, J. T., Hamlington, B. D., Masters, D. & Mitchum, G. T. Climate-change-driven accelerated sea-level rise detected in the altimeter era. *Proceedings of the National Academy of Sciences of the United States of America* **115,** 2022–2025. doi:10.1073/pnas.1717312115 (2018).

146. Vermeersen, B. L., Slangen, A. B., Gerkema, T., Baart, F., Cohen, K. M., Dangendorf, S., Duran-Matute, M., Frederikse, T., Grinsted, A., Hijma, M. P., Jevrejeva, S., Kiden, P., Kleinherenbrink, M., Meijles, E. W., Palmer, M. D., Rietbroek, R., Riva, R. E., Schulz, E., Slobbe, D. C., Simpson, M. J., Sterlini, P., Stocchi, P., Van De Wal, R. S. & Van Der Wegen, M. Sea-level change in the Dutch Wadden Sea. *Geologie en Mijnbouw/Netherlands Journal of Geosciences* **97,** 79–127. doi:10.1017/njg.2018.7 (2018).

147. Baart, F., Rongen, G., Hijma, M., Kooi, H., de Winter, R. & Nicolai, R. *Zeespiegelmonitor 2018* tech. rep. (Deltares report 11202193-000-ZKS-0004, Delft, 2019).

148. Van Pelt, S. C., Beersma, J. J., Buishand, T. A., van den Hurk, B. J. & Schellekens, J. Uncertainty in the future change of extreme precipitation over the Rhine basin: the role of internal climate variability. *Climate Dynamics* **44,** 1789–1800. doi:10.1007/s00382-014-2312-4 (2015).

149. Haasnoot, M., Bouwer, L. & van Alphen, J. *Planning for rapidly accelerating sea-level rise for the Dutch coast* in *EGU General Assembly Conference Abstracts* **20** (2018), 7840.

150. Kind, J., Baayen, J. & Botzen, W. Benefits and Limitations of Real Options Analysis for the Practice of River Flood Risk Management Water Resources Research. *Water Resources Research,* 1–19. doi:10.1002/2017WR022402 (2018).

151. Klerk, W. J., Kanning, W., Kok, M., Bronsveld, J. & Wolfert, A. R. M. Accuracy of visual inspection of flood defences. *Structure and Infrastructure Engineering,* 1–15. doi:10.1080/15732479.2021.2001543 (2021).

152. Van Bergeijk, V. M., Verdonk, V. A., Warmink, J. J. & Hulscher, S. J. M. H. The Cross-Dike Failure Probability by Wave Overtopping over Grass-Covered and Damaged Dikes. *Water* **13,** 690. doi:10.3390/w13050690 (2021).

153. Lendering, K., Jonkman, S. & Kok, M. Effectiveness of emergency measures for flood prevention. *Journal of Flood Risk Management* **9,** 320–334. doi:10.1111/jfr3.12185 (2016).

154. See, J. E., Drury, C. G., Speed, A., Williams, A. & Khalandi, N. The Role of Visual Inspection in the 21 st Century. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **61,** 262–266. doi:10.1177/1541931213601548 (2017).

155. Quirk, L., Matos, J., Murphy, J. & Pakrashi, V. Visual inspection and bridge management. *Structure and Infrastructure Engineering* **2479,** 1–13. doi:10.1080/15732479.2017.1352000 (2017).

156. Ter Berg, C. J., Leontaris, G., van den Boomen, M., Spaan, M. T. & Wolfert, A. R. Expert judgement based maintenance decision support method for structures with a long service-life. *Structure and Infrastructure Engineering* **15,** 492–503. doi:10.1080/15732479.2018.1558270 (2019).

157. Drury, C. G. & Fox, J. G. in *Human reliability in quality control* (eds Drury, C. G. & Fox, J. G.) 1st editio, 11–16 (Taylor and Francis, London, UK, 1975).

158. Graybeal, B. A., Phares, B. M., Rolander, D. D., Moore, M. & Washer, G. Visual inspection of highway bridges. *Journal of Nondestructive Evaluation* **21,** 67–83. doi:10.1023/A:1022508121821 (2002).

159. Moore, M., Phares, B., Graybeal, B., Rolander, D. & Washer, G. *Reliability of Visual Inspection for Highway Bridges* tech. rep. (US Department of Transportation - Federal Highway Administration. Technical Report No. FHWA-RD-01-020, McLean, , VA, USA, 2001).

160. Spencer, F. W. *Visual Inspection Research Project Report on Benchmark Inspections* tech. rep. (FAA Technical Center, DOT/FAA/AR-96/65, 1996).

161. Drury, C. G. & Spencer, F. W. *Measuring Human Detection Performance in Aircraft Visual Inspection* in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (1997), 304–308. doi:10.1177/107118139704100168.

162. Harris, D. Effect of equipment complexity on inspection performance. *Journal of Applied Psychology* **50,** 236–237. doi:10.1037/h0023419 (1966).

163. Gallwey, T. J. & Drury, C. G. Task complexity in visual inspection. *Human Factors* **28,** 595–606. doi:10.1177/001872088602800509 (1986).

164. Dalton, J. & Drury, C. G. Inspectors' performance and understanding in sheet steel inspection. *Occupational Ergonomics* **4,** 51–65. doi:10.3233/OER-2004-4105 (2004).

165. Dirksen, J., Clemens, F. H., Korving, H., Cherqui, F., Le Gauffre, P., Ertl, T., Plihal, H., Müller, K. & Snaterse, C. T. The consistency of visual sewer inspection data. *Structure and Infrastructure Engineering* **9,** 214–228. doi:10.1080/15732479.2010.541265 (2013).

166. Van der Steen, A. J., Dirksen, J. & Clemens, F. H. Visual sewer inspection: detail of coding system versus data quality? *Structure and Infrastructure Engineering* **10,** 1385–1393. doi:10.1080/15732479.2013.816974 (2014).

167. See, J. E. *Visual Inspection: A Review of the Literature* tech. rep. October (Sandia Report SAND2012-8590, 2012), 77.

168. Wiener, E. in *Sustained attention in Human Performance* (ed Warm, J.) 207–246 (Wiley, Chicester, 1984).

169. Crespo Márquez, A. *The Maintenance Management Framework: Models and Methods for Complex Systems Maintenance* 1st ed., 333. isbn: 9781846288203 (Springer, London, 2007).

170. Bakkenist, S., van Dam, O., van der Nat, A., Thijs, F. & de Vries, W. *Principles of Professional Inspection - organizational part* tech. rep. (STOWA, Amersfoort, the Netherlands, 2012). https://www.stowa.nl/sites/default/files/assets/PUBLICATIES/Publicaties%202012/STOWA%202012-13%20PIW%20organizational%20part%20UK.pdf.

171. Flikweert, J. & Simm, J. Improving performance targets for flood defence assets. *Journal of Flood Risk Management* **1,** 201–212. doi:10.1111/j.1753-318x.2008.00021.x (2008).

172. Klerk, W. J. & Adhi, R. A. *Degradation of grass revetments: a comparison of field observations and structured expert judgement* in *Science and practice for an uncertain future* (Budapest University of Technology and Economics, Online, 2021). doi:10.3311/FloodRisk2020.8.2.

173. Long, G., Mawdesley, M. & Simm, J. *Improved approaches to condition assessment - Volume 2: Detailed technical report* tech. rep. June (2006), 150. http://www.fcerm.net/sites/default/files/resources/ur10-vol2.pdf.

174. Het Waterschapshuis. *Digigids* 2016. http://digigids.hetwaterschapshuis.nl/ (2021).

175. Keprate, A. & Chandima Ratnayake, R. Probability of Detection as a Metric for Quantifying NDE Reliability: The State of The Art. *Journal of Pipeline Engineering* (2015).

176. Kruschke, J. K. Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General* **142,** 573–603. doi:10.1037/a0029146 (2013).

177. Aguilar-López, J. P., Warmink, J. J., Bomers, A., Schielen, R. M. & Hulscher, S. J. Failure of grass covered flood defences with roads on top due to wave overtopping: A probabilistic assessment method. *Journal of Marine Science and Engineering* **6,** 1–28. doi:10.3390/jmse6030074 (2018).

178. Steendam, G. J., Van Hoven, A., Van der Meer, J. & Hoffmans, G. *Wave overtopping simulator tests on transitions and obstacles at grass covered slopes of dikes* in *Coastal Engineering Proceedings* **1** (2014), 79. doi:10.9753/icce.v34.structures.79.

179. Gits, C. Design of maintenance concepts. *International Journal of Production Economics* **24,** 217–226. doi:10.1016/0925-5273(92)90133-R (1992).

180. Dekker, R. Applications of maintenance optimization models: a review and analysis. *Reliability Engineering & System Safety* **51,** 229–240. doi:10.1016/0951-8320(95)00076-3 (1996).

181. Speijker, L. J. P., van Noortwijk, J. M., Kok, M. & Cooke, R. M. OPTIMAL MAINTENANCE DECISIONS FOR DIKES. *Probability in the Engineering and Informational Sciences* **14,** 101–121. doi:10.1017/S0269964800141087 (2000).

182. Van Noortwijk, J. M. & van Gelder, P. H. Optimal maintenance decisions for berm breakwaters. *Structural Safety* **18,** 293–309. doi:10.1016/S0167-4730(96)00023-9 (1996).

183. Environment Agency. *Technical Report – FCRM Assets: Deterioration Modelling and WLC Analysis* tech. rep. (2013).

184. Le, H. T., Verhagen, H. J. & Vrijling, J. K. Damage to grass dikes due to wave overtopping. *Natural Hazards* **86,** 849–875. doi:10.1007/s11069-016-2721-2 (2017).

185. Bayoumi, A. & Meguid, M. A. Wildlife and safety of earthen structures: A review. *Journal of Failure Analysis and Prevention* **11,** 295–319. doi:10.1007/s11668-011-9439-y (2011).

186. Van Steeg, P. *Residual strength of grass on river dikes under wave attack* tech. rep. (Deltares, 2013).

187. Van Steeg, P., Klein Breteler, M. & Labrujere, A. *Use of Wave Impact Generator and Wave Flume To Determine Strength of Outer Slopes of Grass Dikes Under Wave Loads* in *Coastal Engineering Proceedings* **1** (2014), 60. isbn: 9780989661126. doi:10.9753/icce.v34.structures.60.

188. Klein Breteler, M. *Residual strength of grass on clay in the wave impact zone* tech. rep. (Deltares, Delft, 2015).

189. Taccari, M. L. & Van Der Meij, R. Study of the effect of burrows of European Badgers (Meles meles) on the initiation of breaching in dikes. *E3S Web of Conferences* **7.** doi:10.1051/e3sconf/20160703001 (2016).

190. Fell, R., Wan, C. F., Cyganiewicz, J. & Foster, M. Time for Development of Internal Erosion and Piping in Embankment Dams. *Journal of Geotechnical and Geoenvironmental Engineering* **129,** 307–314. doi:10.1061/(asce)1090-0241(2003)129:4(307) (2003).

191. Robbins, B. A. & Griffiths, D. V. *Internal Erosion of Embankments: A Review and Appraisal* in *Rocky Mountain Geo-Conference 2018* (American Society of Civil Engineers, Reston, VA, 2018), 61–75. doi:10.1061/9780784481936.005.

192. Özer, I. E., Rikkert, S. J. H., van Leijen, F. J., Jonkman, S. N. & Hanssen, R. F. Sub-seasonal Levee Deformation Observed Using Satellite Radar Interferometry to Enhance Flood Protection. *Scientific Reports* **9,** 2646. doi:10.1038/s41598-019-39474-x (2019).

193. Chlaib, H. K., Mahdi, H., Al-Shukri, H., Su, M. M., Catakli, A. & Abd, N. Using ground penetrating radar in levee assessment to detect small scale animal burrows. *Journal of Applied Geophysics* **103,** 121–131. doi:10.1016/j.jappgeo.2014.01.011 (2014).

194. Miquel, T., Sorin, J.-L., Maurin, J., Tourment, R., Pons, F., Bohard, J. & Biscay, J.-F. DIDRO Project – New means for surveying dikes and similar flood defense structures. *E3S Web of Conferences* **7** (eds Lang, M., Klijn, F. & Samuels, P.) 14002. doi:10.1051/e3sconf/20160714002 (2016).

195. Mendoza, J., Bismut, E., Straub, D. & Köhler, J. Risk-Based Fatigue Design Considering Inspections and Maintenance. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* **7,** 04020055. doi:10.1061/AJRUA6.0001104 (2021).

196. Kala, Z. Global sensitivity analysis of reliability of structural bridge system. *Engineering Structures* **194,** 36–45. doi:10.1016/j.engstruct.2019.05.045 (2019).

197. Zheng, R. & Ellingwood, B. R. Role of non-destructive evaluation in time-dependent reliability analysis. *Structural Safety* **20,** 325–339. doi:10.1016/S0167-4730(98)00021-6 (1998).

198. Hamida, Z. & Goulet, J.-a. Quantifying the effects of interventions based on visual inspections from a network of bridges. *Structure and Infrastructure Engineering* **0,** 1–12. doi:10.1080/15732479.2021.1919149 (2021).

199. Van Hoven, A. & van der Meer, J. *Onderbouwing kansverdeling kritisch overslagdebiet ten behoeve van het OI2014v4* tech. rep. (Deltares, Delft, the Netherlands, 2017).

200. Klerk, W. J. & Jongejan, R. B. *Semi-probabilistic assessment of wave impact and runup on grass revetments* tech. rep. (Deltares, 2016).

201. Papakonstantinou, K. & Shinozuka, M. Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part II: POMDP implementation. *Reliability Engineering & System Safety* **130,** 214–224. doi:10.1016/j.ress.2014.04.006 (2014).

202. Schoemaker, M. A., Verlaan, J. G., Vos, R. & Kok, M. *The use of equivalent annual cost for cost-benefit analyses in flood risk reduction strategies* in *FLOODRisk 2016* **20005** (2016). doi:10.1051/e3sconf/20160720005.

203. Slomp, R., Knoeff, H., Bizzarri, A., Bottema, M. & de Vries, W. Probabilistic Flood Defence Assessment Tools. *E3S Web of Conferences* **7** (eds Lang, M., Klijn, F. & Samuels, P.) 03015. doi:10.1051/e3sconf/20160703015 (2016).

204. Ellingwood, B. R. & Dusenberry, D. O. Building design for abnormal loads and progressive collapse. *Computer-Aided Civil and Infrastructure Engineering* **20,** 194–205. doi:10.1111/j.1467-8667.2005.00387.x (2005).

205. Tarrant, O. Identifying the signs of weakness , deterioration , and damage to fl ood defence infrastructure from remotely sensed data and mapped information. **11,** 317–330. doi:10.1111/jfr3.12326 (2018).

206. ENW. *Leidraad Rivieren* tech. rep. (2007). http://resolver.tudelft.nl/uuid:c14b8a5d-322d-4e2e-8780-cd03a9973cdb.

207. Jonkman, S. N., Hillen, M. M., Nicholls, R. J., Kanning, W. & van Ledden, M. Costs of Adapting Coastal Defences to Sea-Level Rise— New Estimates and Their Implications. *Journal of Coastal Research* **290,** 1212–1226. doi:10.2112/jcoastres-d-12-00230.1 (2013).

208. Sellmeijer, H., de la Cruz, J. L., van Beek, V. M. & Knoeff, H. Fine-tuning of the backward erosion piping model through small-scale, medium-scale and IJkdijk experiments. *European Journal of Environmental and Civil Engineering* **15,** 1139–1154. doi:10.1080/19648189.2011.9714845 (2011).

209. Jongejan, R. *WBI2017 Code Calibration - Reliability-based code calibration and semi-probabilistic assessment rules for the WBI2017* tech. rep. (Rijkswaterstaat, 2017). https://publications.deltares.nl/WeL1942.pdf.

210. Salvatier, J., Wiecki, T. V. & Fonnesbeck, C. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* **2,** e55. doi:10.7717/peerj-cs.55 (2016).

211. Borges, J. & Castanheta, M. Statistical Definition and Combination of Loads, in Probabilistic Design of Reinforced Concrete Buildings. *ACI* **31,** 43–61 (1972).

212. Klein Breteler, M., Capel, A., Kruse, G., Mourik, G. & Kaste, D. *Erosie van een dijk na bezwijken van de steenzetting door golven* tech. rep. (Deltares, 2012).

213. Van Hoven, A. *Residual dike strength after macro-instability* tech. rep. (Deltares, Delft, the Netherlands, 2014).

214. Peters, D. *Design of pattern-placed Revetments* PhD thesis (2017). doi:10.4233/UUID:0B67A0DD-A951-46F3-BBAA-86270E546C4E.

215. Hoffmans, G. *The influence of turbulence on soil erosion* (Eburon Uitgeverij BV, 2012).

216. Van der Meer, J., Hardeman, B., Steendam, G., Schüttrumpf, H. & Verheij, H. *Flow depths and velocities at crest and landward slope of a dike, in theory and with the wave overtopping simulator* in *Coastal Engineering Proceedings* (2010).

217. Natural Resources Conservation Service. in *National Engineering Handbook Part 628 Dams* 210-vi-neh (1997).

218. Hughes, S. A. & Shaw, J. M. Continuity of Instantaneous Wave Overtopping Discharge with Application to Stream Power Concepts. *Journal of Waterway, Port, Coastal, and Ocean Engineering* **137,** 12–25. doi:10.1061/(ASCE)WW.1943-5460.0000057 (2011).

219. TAW. *Technisch rapport klei voor dijken* tech. rep. (Delft, 1996), 81.

220. CBS. *Grond-, weg- en waterbouw (GWW); inputprijsindex 2015=100* 2021. https://www.cbs.nl/nl-nl/cijfers/detail/84538NED (2021).

A

Validation of the greedy search algorithm

The aim of this section is to validate the greedy search algorithm introduced in Section 3.5.1. We compare the obtained results with those of a method guaranteed to find the global optimum (in this case the Mixed Integer Programming implementation of the problem described in Section 3.4). Appendix A.1 presents a general discussion on input data and the approach taken for verifying the algorithm performance. Appendix A.2 presents the results of the validation. More detailed information on input data can be found in Appendix B.

A.1. Input data and approach

As test data we use data from 73 dike sections at the river Lek in the Netherlands (including the sections from the case study in Section 3.6). By randomly selecting subsets of dike sections we can generate many different realistic segment configurations with different numbers of sections. For each section we have information on current reliability, the reliability after taking different measures, the cost of measures and the damage in case of flooding. Note that for the validation of the search routine we do not include Equation (3.9) as a constraint, as this is merely an optional additional stopping criterion.

We consider the reliability for overtopping, piping erosion and inner slope instability failures. Reliability estimates were obtained by back calculating implicated reliability indices using the semi-probabilistic assessment rules in the applicable statutory safety assessment tools [6]. It has to be noted that the approach can as easily be used with any failure model as long as it provides a probability of failure for a dike section. Note that it holds that a reliability index $\beta \approx -\Phi^{-1}(P_{\mathrm{f}})$. To properly assess reliability over time we include relevant temporal changes that impact different failure modes. Higher outside water levels reduce reliability for all failure modes; increases in wave run-up due to higher wind speeds, as well as settlement of the crest reduce overtopping reliability; and settlement of the hinterland results

Table A.1.: *Different combinations of sets of measures considered. x indicates the measure is included. Set 6 considers all measures at t=0, other sets also consider measures at t=20. Extent and type refer to the classification in Figure 3.2.*

Measure	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Extent	Type	Mechanisms impacted
Diaphragm Wall	x	x	x			2025	Full	Renewal	All
Soil based reinforcement	x	x	x	x	x	2025	Full/Partial	Renovation	All
Stability screen	x		x	x		2025	Partial	Renewal	Inner slope instability
Vertical Sandtight Geotextile	x	x		x		2025	Partial	Renewal	Piping erosion

in increased hydraulic heads which reduces piping erosion reliability. Reliability in time (see right part of Figure 3.1) was derived for each of the 73 sections based on local data (see Appendix B for formulations).

In the validation we want to consider the influence that including different types of measures has on the performance of the algorithm. Therefore we consider different sets of available measures as shown in Table A.1. Set 1 is a set of all available options for investment years 2025 and 2045. In sets 2 through 5 different measures are excluded. Set 6 only considers investments in 2025. The costs are obtained from standard cost functions [111], except for soil based reinforcement. For soil based reinforcement we consider starting costs, variable costs based on the volume of added soil, and costs dependent on the number of adjacent properties to be removed for each section, which has a large impact on reinforcement costs. For all computations the economic consequences of flooding are assumed to be 5 billion €, the annual discount rate is assumed to be 3%. More detailed formulations for cost and reliability computations can be found in Appendix B.

The combination of different system configurations and sets of measures gives us a large variety of realistic cases for which we can assess the performance of the greedy search algorithm by comparing with a MIP implementation in CPLEX 12.9.

A.2. Validation of the greedy algorithm

A typical run of the greedy search algorithm yields a stepwise prioritization of dike reinforcement measures that eventually ends at or very close to the global optimal solution, which consists of the minimum sum of $TLCC$ and TR. Figure A.1 shows results for a system with 5 dike sections. We see that the greedy search (red) reaches the global optimum, and follows the Pareto front for $TLCC$ and TR (black) computed using the MIP implementation with variable budget limits (i.e., where $TLCC$ is constrained). This shows that, especially closer to the optimal solution (blue diamond), the investment path of the greedy search not only finds the optimal solution, but the intermediate steps are also (near-)optimal for that budget.

Next we consider a large set of system configurations from our dataset and combine these with different sets of measures as defined in Table A.1. We randomly sample configurations of N sections. The largest system size considered is 11 sections, which is smaller than typically encountered in practice but the largest practically feasible with the available 16 GB RAM. We consider regular cases (i.e. directly sampled from our dataset) that are typically dominated by failures due to inner

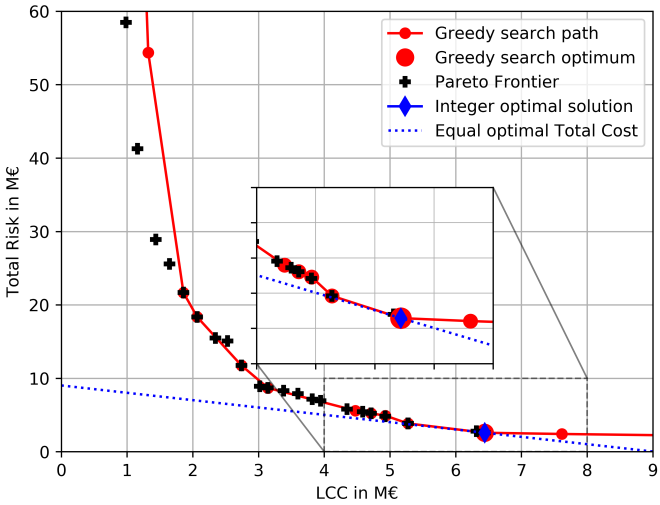


Figure A.1.: *Example result of a system with 5 dike sections. Red line shows the path of the greedy search, where the large dot denotes the optimal solution. Blue diamond indicates the optimum found using branch-and-cut in CPLEX 12.9. The black pluses denote the Pareto Frontier derived from several branch-and-cut computations with a budget limit. Note that here the TR and TLCC are displayed as the conflicting objectives, whereas in the optimization routine these are summed and considered as a single objective.*

Table A.2.: *Results for 2800 different system configurations with different measure sets and system sizes. Systems derived from the case study in Section 3.6 ('Regular cases') are distinguished from systems that are dominated by failures due to overflow. Δ denotes the relative difference between the two methods. Δ* denotes a threshold value exceeded by only 5% of all cases. For both Total Cost and LCC three indicators are used: the % of cases with a Δ larger than 1 %, the Δ* only exceeded by 5 % of the cases and the average difference Δ.*

	Case properties				**Total Cost**			**LCC**		
	Measure set	No. of sections	Total runs	Runs where global optimum was found	$\Delta > 1\%$	Δ^* for $P(\Delta > \Delta^*) < 0.05$	Average Δ	$\Delta > 1\%$	Δ^* for $P(\Delta > \Delta^*) < 0.05$	Average Δ
Regular cases	Set1	mixed	400	91.0%	2.25%	0.37%	0.07%	6.0%	2.2%	0.37%
	Set2	mixed	400	93.3%	1.50%	0.17%	0.05%	6.0%	2.2%	0.35%
	Set3	mixed	400	92.0%	0.50%	0.15%	0.03%	3.8%	0.4%	0.22%
	Set4	mixed	400	93.0%	0.75%	0.16%	0.03%	5.5%	2.2%	0.60%
	Set5	mixed	400	96.5%	0%	0%	0.01%	2.5%	0%	0.06%
	Set6	mixed	400	93.3%	1.00%	0.37%	0.04%	3.0%	0.6%	0.23%
	Size5	5	600	97.8%	0.17%	0%	0.02%	1.7%	0%	0.21%
	Size7	7	600	95.8%	1.00%	0%	0.03%	3.3%	0%	0.22%
	Size9	9	600	91.0%	1.33%	0.26%	0.05%	6.3%	2.2%	0.46%
	Size11	11	600	88.0%	1.50%	0.37%	0.05%	6.5%	2.1%	0.33%
	All	mixed	2400	93.2%	1.00%	0.18%	0.04%	4.5%	0.7%	0.30%
Overflow dominant cases	Set2	5	100	99.0%	0%	0%	0.01%	1.0%	0%	0.06%
	Set2	7	100	98.0%	0%	0%	0.01%	1.0%	0%	0.05%
	Set2	9	100	94.0%	2.00%	0.13%	0.07%	4.0%	0.40%	0.34%
	Set2	11	100	86.0%	3.00%	0.39%	0.11%	11.0%	4.66%	0.49%
	All	mixed	400	94.3%	1.25%	0.19%	0.05%	4.3%	0.55%	0.24%
Total	All	mixed	2800	93.3%	1.04%	0.18%	0.04%	4.4%	0.71%	0.29%

slope instability and piping erosion. We also consider system configurations that are dominated by overflow failures. Here crest levels of sections were modified such that the initial overflow reliability index ranges between 2.8 and 3.5, making it the dominant failure mechanism.

Results are shown in Table A.2, where f_c of the greedy search is set to 1.5. Overall we see that in about 93% of all cases the greedy search finds the global optimum. If the outcome differs from the global optimal solution, only in 1% of the cases the difference in TC is higher than 1%, and only in 5% of the cases the difference is larger than 0.18%. On average the difference is 0.04% which is negligible compared to the often large uncertainties in dike reinforcement projects. The performance for regular cases and cases dominated by overflow failures is very similar.

It is found that differences in Life Cycle Cost are larger. This is explained by the fact that many cases with large differences in LCC, are cases that are often very close to the global optimum. Overall the differences are small: only 5 % of the cases has a difference larger than 0.71%. On average differences are only 0.29%. In most cases the differences arise from cases with different investment cost but close to optimal values for the objective of minimizing TC. Practically this means that even though the solution is not exactly optimal, there are many different combinations of investments that are close to optimal, even though the investment costs are different. This has the practical advantage that it allows policy makers to choose from various near optimal solutions. Also it has to be noted that small

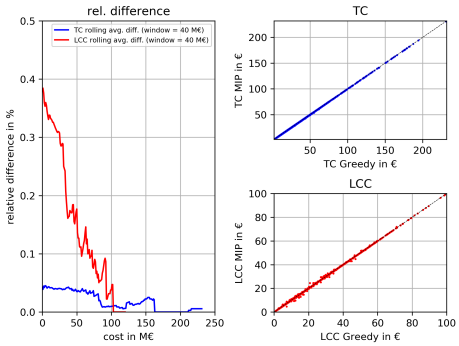


Figure A.2.: *left: Relative error for Total Cost and Life Cycle Cost compared to MIP optimum. Red and blue lines denote a moving average relative difference for all evaluated cases based on a window for absolute TC or LCC of 40 M€. right: comparison of TC and LCC for MIP and Greedy computations. It can be observed that for the LCC values the scatter is slightly larger, especially for small absolute values.*

choices in for instance the schematization of different failure modes, or uncertainty in for instance sea level rise rates have a potentially much larger effect on TC [207].

Another observation from Table A.2 is that the percentage of runs where the global optimum is found decreases slightly for larger systems, which is sensible from the perspective that there are many more different investment paths that can be taken. However, when looking at the investment costs, the deviation does not increase significantly, which is confirmed by the results in Figure A.2. The left pane shows that for both TC and LCC there is a clear decreasing trend of the average relative error for larger values of TC and LCC. This is also represented in a slightly different way in the scatter plots in the right panes. Thus we can conclude that for larger dike segments consisting of many sections that are relatively expensive to improve, the algorithm can be expected to provide accurate results, even though the exact global optimal solution might not be found.

As was explained in Section 3.5.1 the factor f_c can be used to vary the step size taken by the algorithm. The same cases have been evaluated using f_c equal to 1.0, 1.5 and 3 respectively. Table A.3 shows the results for TC for all cases with different settings. Aside from the individual evaluations we also use a combination of the three settings where each time the greedy solution with the lowest TC is used. It turns out that the different settings all perform quite well, but $f_c = 1.5$ results in the highest accuracy, even to the extent that the performance metrics for the combined case are the same as for $f_c = 1.5$. It turns out that in some cases $f = 1.0$ or $f = 3.0$ yield a better result than $f = 1.5$, but these have no bearing on the overall performance (typically because the inaccuracy in these cases is already very small).

One of the advantages of the greedy search algorithm is that it doesn't suffer as much from state space explosion as the MIP approach. Typically for the cases considered we observe that the runtime doubles for each two sections added. At

Table A.3.: *Comparison of results for three different settings of f_c as well as the best combination of the three. Indicators are the same as in Table A.2.*

		Case type		
	Setting	**Normal**	**Overflow dominant**	**All cases**
Runs where global optimum was found	*f=1.0*	90.63%	92.50%	90.89%
	f=1.5	93.17%	94.25%	93.32%
	f=3.0	90.04%	92.75%	90.93%
	combined	93.17%	94.25%	93.32%
Cases with $\Delta > 1\%$	*f=1.0*	2.75%	3.00%	2.79%
	f=1.5	1.00%	1.25%	1.04%
	f=3.0	3.13%	2.75%	2.79%
	combined	1.00%	1.25%	1.04%
Δ^* **for** $P(\Delta > \Delta^*) < 0.05$	*f=1.0*	11.76%	7.40%	11.76%
	f=1.5	5.53%	3.62%	5.53%
	f=3.0	11.76%	7.40%	11.76%
	combined	5.53%	3.62%	5.53%
Avg. difference	*f=1.0*	0.10%	0.12%	0.10%
	f=1.5	0.04%	0.05%	0.04%
	f=3.0	0.11%	0.11%	0.10%
	combined	0.04%	0.05%	0.04%
Total runs		2400	400	2800

the hardware that was used (16 GB of RAM), the largest system that could be solved with MIP contained 15 dike sections (with the most extensive measure set). Furthermore it should be noted that aside from the time required to solve the MIP problem, the initialization of the problem (i.e., defining al the constraints, especially Equation (3.8)) also costs significantly more time using MIP than with the greedy algorithm. Overall, initializing and solving a system of 15 dike sections using MIP took approximately 600 seconds, whereas the greedy search routine can find a solution in approximately 6 seconds. With more extensive hardware the computable system size could be extended, but not easily to the size of practical problems that often consist of over 30 sections.

B

Formulations for computations of cost and reliability over time

This supplement describes the formulations used for computation of cost and reliability over time for the reinforcement measures considered in Chapter 3 and Section 4.3. First we will discuss the reliability computation for the three failure modes, then we discuss the reliability for different reinforcement measures, and how costs have been computed.

B.1. Reliability computation

For failure due to overflow and overtopping we use the statutory safety assessment tools available for the Netherlands. With these tools we can compute the probability of failure given a certain level of the dike crest and an allowable overtopping discharge over the dike crest. An example is shown in Figure B.1. These curves have been derived for each dike section for the water level and the additional load due to waves h_{load}. Combined with temporal changes in crest height $\Delta h_{\text{c}}(t)$ and load $\Delta h_{\text{load}}(t)$ this is then used as input for the following limit state function:

$$Z(t) = (h_{\text{crest}} - \Delta h_{\text{c}}(t)) - (h_{\text{load}} - \Delta h_{\text{load}}(t)), \tag{B.1}$$

where h_{crest} is the initial height of the dike crest. It has to be noted that $\Delta h_{\text{load}}(t)$ is different from the increase in water level. For the area considered it holds that there is a factor f_{load} such that it holds that:

$$\Delta h_{\text{load}}(t) = \Delta h(t) \cdot f_{\text{load}}, \tag{B.2}$$

where $\Delta h(t)$ is the increase in water level over time. The factor f_{load} is approximately 1.65 for the case study area. Of course more detailed local information could be used if needed.

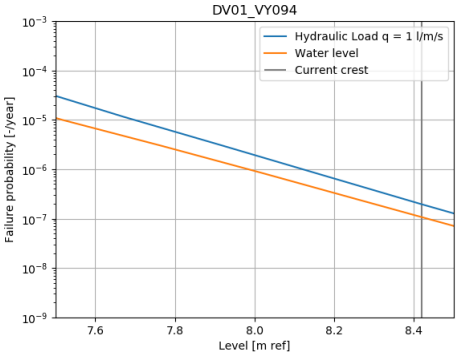


Figure B.1.: *Example of relations between exceedence probability and hydraulic load and water level.*

For failure due to piping erosion we derive frequency lines for water levels similar to the curves used for the hydraulic load (which include waves). This is then used as input for a semi-probabilistic assessment based on the model by Sellmeijer et al. [208], based on assessment of uplift, heave and piping. The semi-probabilistic safety format is described in Jongejan [209]. The computed safety factors are then translated to an implicated reliability index for a cross-section. It has to be noted that this is a relatively conservative estimate. To incorporate temporal changes we consider the change in water level over time Δh and the subsidence of the polder at the inner side of the dike. This increases the hydraulic head over time.

For failure due to inner slope instability we use a similar approach as for piping erosion, where inner slope stability factors for the year 2025 and 2075 have been computed using the semi-probabilistic safety format described in Jongejan et al. [6]. We then interpolate linearly between the stability factors for both years, and translate to implicated reliability indices per cross section using the relation from Jongejan et al. [6]. Time-variant factors are typically limited and included in the two inner slope instability assessment computations.

B.2. Modelling of reinforcement measures

In our study we consider 4 different reinforcement measures. In the following subsections we briefly describe for each type of measure how we computed costs and how changes in reliability were determined for each failure mode. It has to be noted that we also consider combinations of measures (e.g., a berm with a Vertical Sandtight Geotextile). In such a case we use the envelope of the reliability by taking the highest of the two measures over time, for each failure mode. This is illustrated in Figure B.2

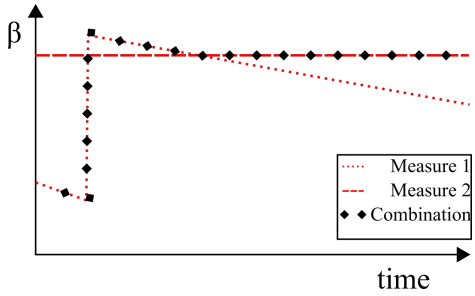


Figure B.2.: *Approach for the reliability when combining two different measures.*

Table B.1.: *Input values for the different cost parameters.*

	Unit	Value
C_{init}	€	200,000
C_{soil}	€/m³	40
C_{buil}	€/building	500,000
C_{DW}	€/m'	20,000
C_{SS}	€/m²	350
C_{VSG}	€/m'	1,000

B.2.1. Reinforcement with soil

For a reinforcement with soil two parameters are given: the change in crest height and the additional berm length. The increased crest height translates into a direct increase of h_{crest} in Equation (B.1). For piping erosion the additional berm length translates into an increase in effective seepage length. For inner slope instability additional berm length translates into an increase in stability factor. However, no additional computations were made to exactly determine this increase, but a general relation between additional berm length (ΔL_{berm}) and stability factor (SF) was derived based on a set of cases in the same area. This relation depends on the thickness of the cover layer (i.e., the top part of the soil with low permeability), and is given in Figure B.3.

Cost for soil reinforcement are based on simplified cost formulations used in cost estimation software KOSwat [111]. Based on the additional berm length and increase in crest height the additional area per cross section is computed and then translated to a soil volume (V) using the section length. The cost is the determined using the following formula:

$$C = C_{\text{init}} + C_{\text{soil}} \cdot V + C_{\text{buil}} \cdot B, \tag{B.3}$$

Where C_{init} are the starting costs in €, C_{soil} is the cost per m³ of soil to be added and C_{buil} is the cost for each building that has to be removed due to widening of the berm. B is the number of objects which has been determined based on a geospatial analysis of the location of all buildings with a footprint larger than 50 square meters. The values for the cost parameters are given in Table B.1.

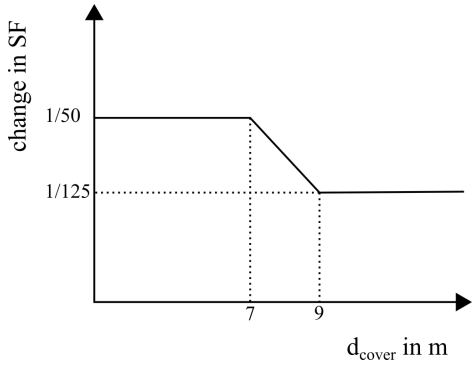


Figure B.3.: *Relation between thickness of the cover layer and increase in stability factor per meter berm.*

B.2.2. Reinforcement with diaphragm wall

Diaphragm walls are a type of structures that are typically constructed such that they are self-retaining, for instance by pouring a 1 m thick reinforced concrete wall that reaches the Holocene sand layer. In our case we assume that the diaphragm wall ensures a failure probability of 10^{-8} for both piping erosion and inner slope instability. For overtopping the failure probability is altered such that it satisfies the target reliability for overtopping in 2075. The cost of a diaphragm wall is computed based on experiences with diaphragm walls in the same area. The unit cost C_{DW} for $1\,\mathrm{m}$ of diaphragm wall is $20000\,€$.

B.2.3. Reinforcement with stability screen

The reinforcement with a stability screen concerns the construction of a sheet pile in the inner toe of the dike. It only impacts inner slope instability. It reaches through the (typically) weaker top layers into the underlying sand layer (1 meter overlap), thus eliminating slip circles through the weaker top layers and increasing reliability. A simple assumption for the strength is that the stability factor increases by a factor 0.2. It has to be noticed that this can differ significantly in practice. The cost of a stability screen depends on the thickness of the cover layer (d_{cover}) and is computed as follows:

$$C = C_{\mathrm{SS}} \cdot (d_{\mathrm{cover}} + 1) \cdot L, \tag{B.4}$$

where C_{SS} is the unit cost for a square meter of sheet pile, and L is the length of the dike section.

B.2.4. Reinforcement with Vertical Sandtight Geotextile

A Vertical Sandtight Geotextile (VSG) is a vertically installed geotextile that prevents sand particles from eroding [65]. Thus it can prevent backwards piping erosion. However, it only functions if the VSG is installed correctly. We compute the failure

probability for piping erosion as follows:

$$P_{\mathrm{f}}(t) = P_{\mathrm{f}|\overline{\mathrm{VSG}}} \cdot P(\overline{VSG}) + P_{\mathrm{f}|\mathrm{VSG}} \cdot P(VSG), \tag{B.5}$$

where $\overline{VSG}$ denotes failure of the VSG, and f denotes failure of the dike. we assume that $P_{\mathrm{f}|\overline{\mathrm{VSG}}}$ is equal to the failure probability without VSG, $P(\overline{VSG}) = 1/1000$ and $P_{\mathrm{f}|\mathrm{VSG}} = 10^{-8}$. The unit cost for a VSG is 1,000 € per meter, this is based on experiences in other projects.

C

Additional data from inspection field test

This appendix presents some additional data from, and methods for the field test described in Chapter 6. Appendix C.1 describes the questionnaires and influential factors used for the Bayesian Parameter Estimation for different influential variables. Appendix C.2 describes the assumptions for the Bayesian Parameter Estimation. Appendix C.3 presents some additional figures and results.

C.1. Questionnaires and influential factors

Table C.1 displays the questions posed to the individual inspectors at different times before and during the field test in Chapter 6. Table C.2 displays the factors considered in the analysis, including the questions these relate to.

Table C.1.: *Overview of all questions in the questionnaires during the test. 1-5 scales represent ranges between very bad and very good or disagree and agree.*

	Question	Category	Type of answer	Additional remarks
Before the day of the field test				
1	Age	Personal characteristics	Number	
2	Color blindness	Personal characteristics	yes/no	
3	Wearing glasses or contact lenses?	Personal characteristics	yes/no	

Table C.1.: *Overview of all questions in the questionnaires during the test. 1-5 scales represent ranges between very bad and very good or disagree and agree.*

	Question	Category	Type of answer	Additional remarks
4	How well do you see at short distance? (<2 meters, incl. glasses/lenses)	Personal characteristics	1-5	
5	How well do you see at long distance? (>10 meters, incl. glasses/lenses)	Personal characteristics	1-5	
6	Level of education	Training	Multiple choice (1 answer)	Levels of Dutch education
7	What % of your daily work is inspecting flood defences?	Experience	Number	
8	How many years have you worked on flood defences in general?	Experience	Number	
9	How many years have you been inspecting flood defences?	Experience	Number	
10	What types of flood defence inspections have you been involved with?	Experience	Multiple choice	Common types of inspection
11	Are you or were you ever involved in flood defence maintenance?	Experience	Multiple choice	Possible answers: Planning maintenance Executing maintenance No involvement
12	Do you have experience catching musk rats?	Experience	yes/no	
13	How many years?	Experience	Multiple choice (1 answer)	<5, 5-10 and >10 years
14	Did you do inspections of other types of infrastructure?	Experience	yes/no	
15	How many years of experience inspecting other infrastructure?	Experience	Multiple choice (1 answer)	<2, 2-5, 5-10 and >10 years
16	Which types of other infrastructure?	Experience	Multiple choice	Possible answers: Industrial systems, roads, bridges, sluices, banks, quay walls, other
17	Did you follow the regular course on flood defence inspections?	Training	yes/no	yes/no
18	Years since following the course	Training	Number	
19	Did you follow any other courses on inspections?	Training	Text	
20	Wat is your most common method for doing a spring inspection?	Inspection approach	Multiple choice	By car, walking, alone, in a team.

Table C.1.: *Overview of all questions in the questionnaires during the test. 1-5 scales represent ranges between very bad and very good or disagree and agree.*

	Question	Category	Type of answer	Additional remarks
21	During an inspection I register …	Inspection approach	Multiple choice (1 answer)	Possible answers: All damages All damages in need of maintenance Global damages per section Condition of elements Other
22	If I inspected the same section in a previous spring inspection, I know what to expect	Bias	1-5	
23	If I inspected the same section less than 3 months ago, I know what to expect	Bias	1-5	
24	If I see a flood defence, I know what damages to expect	Bias	1-5	
25	In a spring inspection I only register damage of which I think it is important	Bias	1-5	
26	If I do a spring inspection with a colleague, the colleague is more stringent	Bias	1-5	
27	If I have been inspecting for a few hours, I register differently than when I just started	Bias	1-5	
28	As preparation for an inspection, I check which damages have been observed in previous inspections	Inspection approach	Multiple choice (1 answer)	Yes, no, sometimes
29	At the end of a long day the pace of inspections is lower	Bias	1-5	
30	Explanation	Bias	Text	
31	My physical condition is … for doing spring inspections	Personal characteristics	1-5	
32	Inspecting flood defences is of large societal importance	Motivation	1-5	
33	I feel appreciated for my efforts as inspector/asset manager	Motivation	1-5	
34	I'm proud that I'm allowed to and capable of inspecting flood defences	Motivation	1-5	
35	Inspecting flood defences is interesting work	Motivation	1-5	
36	If I could improve one aspect of spring inspections I would …			
Before the field test started				
37	I register using a ...	Inspection approach	Multiple choice (1 answer)	tablet/smartphone

Table C.1.: *Overview of all questions in the questionnaires during the test. 1-5 scales represent ranges between very bad and very good or disagree and agree.*

	Question	Category	Type of answer	Additional remarks
38	Are there physical limitations that might impact your inspection work?	Personal characteristics	Yes/no + explanation	
39	How do you feel today?	Personal characteristics	1-5	
40	How many days ago was your last inspection of a primary flood defence?	Experience	Number	
After the field test				
41	The inspections done were similar to what I usually do in a spring inspection	Field test evaluation	1-5	
42	The field test set-up was suitable to measure inspection quality	Field test evaluation	1-5	
43	The tasks given were clear	Field test evaluation	1-5	
44	Even though people were watching, this had no impact on my inspection and how I behaved	Field test evaluation	1-5	
45	I tried doing even better than in a normal inspection	Field test evaluation	1-5	
46	I did not miss any major damages	Field test evaluation	1-5	
47	Weather conditions impacted the inspection	Field test evaluation	1-5	
48	The equipment I used was comparable to what I commonly use	Inspection approach	1-5	

Table C.2.: *All variables and their relation to the questionnaires. In some cases variables have not been considered in the analysis. The main reasons are either due to the fact that there is no reason for factors to have an influence on the inspection performance (e.g., evaluation questions), or if visual inspection of the data clearly shows that no relationship can be derived (i.e., unbalanced among groups or very bad fit of regression).*

Variable	Kind	Based on question	Category	Considered in analysis
Expectation after previous spring inspection	Numerical	22	Bias	No, not likely to have influence on field test result
Expectation if inspected less than 3 months before	Numerical	23	Bias	No, not likely to have influence on field test result
Expectation of damages	Numerical	24	Bias	No, linear regression has bad fit
Registration approach 2	Numerical	25	Bias	No, linear regression has bad fit
Stringency compared to colleagues	Numerical	26	Bias	Yes
Drift in inspection assessments over time	Numerical	27	Bias	No, not likely to have influence on field test result
Inspections are slower at the end of the day	Numerical	29	Bias	No, not likely to have influence on field test result
Other objects inspected	Categorical	14	Experience	Yes
Involved in daily inspection	Categorical	10	Experience	Yes
Involved in emergency inspection	Categorical	10	Experience	Yes
Involved in summer and autumn assessment	Categorical	10	Experience	Yes
Involved in maintenance planning	Categorical	11	Experience	Yes
Involved in maintenance execution	Categorical	11	Experience	Yes
Was previously involved in maintenance planning	Categorical	11	Experience	No, covered by related factors on maintenance
Not involved in maintenance	Categorical	11	Experience	Yes
Inspected bridges	Categorical	16	Experience	Yes
Inspected quay walls	Categorical	16	Experience	Yes
Inspected river banks	Categorical	16	Experience	Yes
Inspected sluices	Categorical	16	Experience	Yes
Time since last inspections	Categorical	40	Experience	Yes
Percentage inspecting flood defences	Numerical	7	Experience	Yes
Years of experience with flood defences	Numerical	8	Experience	Yes
Years of experience in flood defence inspection	Numerical	9	Experience	Yes
Field test was like common spring inspection	Numerical	41	Field test evaluation	No, evaluation question
Field test is capable of measuring inspection performance	Numerical	42	Field test evaluation	No, evaluation question
Not influenced by people watching me	Numerical	44	Field test evaluation	No, evaluation question

Table C.2.: *All variables and their relation to the questionnaires. In some cases variables have not been considered in the analysis. The main reasons are either due to the fact that there is no reason for factors to have an influence on the inspection performance (e.g., evaluation questions), or if visual inspection of the data clearly shows that no relationship can be derived (i.e., unbalanced among groups or very bad fit of regression).*

Variable	Kind	Based on question	Category	Considered in analysis
Put in an extra effort during field test	Numerical	45	Field test evaluation	No, evaluation question
Device used for registration	Categorical	37	Test circumstances	Yes
Societal importance of inspections	Numerical	32	Motivation	No, limited variation among inspectors
Appreciation of inspection work	Numerical	33	Motivation	No, unbalanced answers dominated by 1 inspector
Proud of being a flood defence inspector	Numerical	34	Motivation	No, limited variation among inspectors
Inspections are interesting	Numerical	35	Motivation	No, unbalanced answers dominated by 1 inspector
Wearing glasses or contact lenses	Categorical	3	Personal characteristics	Yes
Asset manager	Categorical	Other	Personal characteristics	Yes
Age	Numerical	1	Personal characteristics	Yes
Visual acuity short	Numerical	4	Personal characteristics	Yes
Visual acuity long	Numerical	5	Personal characteristics	Yes
Self-assessed physical state	Numerical	31	Personal characteristics	Yes
Self-assessed feeling during field test	Numerical	39	Personal characteristics	Yes
Session A	Categorical	Other	Test circumstances	Yes
Session C	Categorical	Other	Test circumstances	Yes
Session B	Categorical	Other	Test circumstances	Yes

Table C.2.: *All variables and their relation to the questionnaires. In some cases variables have not been considered in the analysis. The main reasons are either due to the fact that there is no reason for factors to have an influence on the inspection performance (e.g., evaluation questions), or if visual inspection of the data clearly shows that no relationship can be derived (i.e., unbalanced among groups or very bad fit of regression).*

Variable	**Kind**	**Based on question**	**Category**	**Considered in analysis**
Solo at section 1	Categorical	Other	Test circum-stances	Yes
Solo at section 2	Categorical	Other	Test circum-stances	Yes
Solo at section 3	Categorical	Other	Test circum-stances	Yes
Solo at section 4	Categorical	Other	Test circum-stances	Yes
Total solo rounds	Numerical	Other	Test circum-stances	Yes
Regular course on flood de-fence inspections	Categorical	17	Training	Yes
Years since inspection course numeric	Numerical	18	Training	Yes

C.2. Bayesian Parameter Estimation

This appendix outlines the assumptions for the Bayesian Parameter Estimation as applied in Chapter 6 in more detail. The main goal is to obtain posterior distributions of the number of detected damages for different categorical and numerical variables.

For categorical variables we compare 2 groups at a time ($i = 1, 2$). The goal is to determine for each group i the mean μ, the precision λ and the degrees-of-freedom ν of a three-parameter Student-t distribution of the number of detected damage points d given by:

$$f(d|\mu, \lambda, \nu) \sim \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}(\frac{\lambda}{\pi\nu})^{1/2}\left[1 + \frac{\lambda(x-\mu)^2}{\nu}\right]^{-(\nu+1)/2} \tag{C.1}$$

For both groups we define the prior mean μ_i as a normal prior:

$$\mu_i \sim N(\overline{x}, 2s), \tag{C.2}$$

where $\overline{x}$ and s are the pooled empirical mean and standard deviation of the input data. For the standard deviation σ_i for both groups we use a uniform prior:

$$\sigma_i \sim U(0.1, 5), \tag{C.3}$$

5 is a relatively high limit, given that the variation of detected damage points is typically much smaller. For the degrees-of-freedom ν we use an exponential prior with a relatively low mean of 7, as typically 1 of the groups is quite small, meaning that the data is typically heavy-tailed (note that we shift to exclude $\nu = 0$):

$$\nu \sim Exp(\frac{1}{7-1}) + 1. \tag{C.4}$$

Next we can determine the difference of the means of the groups $(\mu_1 - \mu_2)$, standard deviations $(\sigma_1 - \sigma_2)$ and the effect size $(\mu_1 - \mu_2)/\sqrt{(\sigma_1^2 + \sigma_2^2)/2}$.

For numerical variables we use a similar approach, but here we estimate the parameters of a linear regression between the number of damage points detected d and the numerical parameter y such that $d = a + b \cdot y + \epsilon$. We assume normal priors for intercept a, slope b and a half-normal prior for noise ϵ:

$$a \sim N(10, 10) \tag{C.5}$$

$$b \sim N(0, 10) \tag{C.6}$$

$$|\epsilon| \sim N(0, 10) \tag{C.7}$$

Within the context of the input data (detected damage points in the order of 10-20 points per inspector), these priors are relatively uninformed.

Based on the available data on the number of registered damage points for inspectors in different groups, and with different numerical variables we can obtain estimates for the posterior distributions. We do this using Markov Chain Monte Carlo (MCMC) sampling in PyMC3 [210] which yields posterior distributions for all uncertain parameters for each categorical and numerical variable.

C.3. Additional figures and results

In this appendix some additional figures and results are given for the field test in Chapter 6. Figure C.1 presents results for damage points for single and pairs of inspectors. Figure C.2 presents the same results for section damages.

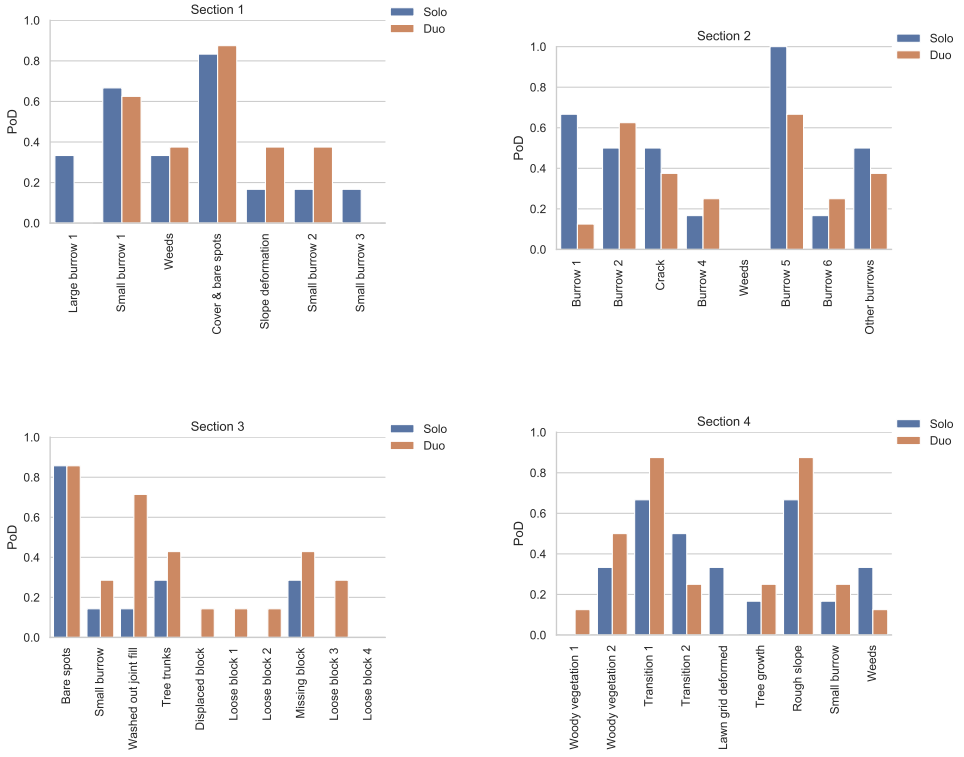


Figure C.1.: *Comparison of PoD for damage points by single and pairs of inspectors.*

Figure C.2.: *Comparison of PoD for section damages by single and pairs of inspectors.*

D

Computations for reliability and cost

This appendix presents the reliability models used for external erosion (Appendix D.1), wave overtopping (Appendix D.1.1), as well as assumptions regarding costs of inspection, maintenance and reinforcement (Appendix D.2) as made in Chapter 7.

D.1. Reliability models for external erosion

This appendix outlines the limit state functions for external erosion of the inner and outer slope. We focus on flood defence segments in river systems such as the Rhine river. For most of such segments it holds that wave loads and water levels are uncorrelated. This means that the probability that large waves hit the top part of the slope is relatively small. If we have marginal distributions for the exceedence probability of the significant wave height $p(H_{m0} > \overline{H_{m0}})$ and water levels $p(h > \overline{h})$ the probability that a statistically uncorrelated high water level and high waves coincide is:

$$P(h(t) > \overline{h} \cap H_{m0}(t) > \overline{H_{m0}})) = p(h > \overline{h}) \cdot p(H_{m0} > \overline{H_{m0}}) \cdot \frac{t_h / t_{\text{storm}}}{f}, \qquad \text{(D.1)}$$

where t_h is the elementary interval of an extreme water level, and t_{storm} is the elementary interval of a storm event that causes large waves in accordance with Borges & Castanheta [211]. We assume $t_h = 96$h (approximately the peak width of the discharge of the Rhine river) and $t_{\text{storm}} = 12$h (typical duration of a storm) in line with assumptions in the Dutch safety assessment. The factor f represents the period of the year during which extremes might occur. Here we use $f = 365/2$, as we assume that such combinations only occur during the hydrological winter. This also implies that the failure probability with and without damage during the hydrological summer is approximately 0. These values can be different for each river system.

The next subsections describe the models used for determining the failure probability of the inner and outer slope due to external erosion. An overview of input parameters and distributions is given in Table D.1.

D.1.1. Failure due to wave impact at the outer slope

For failure due to wave impact at the outer slope we consider the failure paths displayed in Figure 7.2a. We evaluate these failure paths along the slope of section n which yields the failure probability for points x along the slope $P(F(x)|S_{\mathrm{i}}, n)$, where a slope zone z is the area between vertical coordinates $[y_i, y_j]$. If the revetment is intact, we first determine the erosion of the grass cover using resistance-duration curves [200]. A resistance-duration curve relates the wave height to the duration of the load and whether a grass revetment can withstand it such that:

$$H_{\mathrm{m0}} = a \cdot e^{b \cdot t_{\mathrm{grass}}} + c, \tag{D.2}$$

where a, b and c are values based on erosion tests such as described in [187] (values for open and closed sod are given in Table D.1), t_{grass} is the time until failure of the grass cover in hours (i.e., the top 20 cm of the cover layer). Note that this equation is based on the assumption that the waves represent a Rayleigh spectrum similar to the erosion tests. Using this formulation we can determine the time it takes for the grass revetment to fail, after which the underlying part of the clay layer starts to erode. The erosion of the clay layer is modelled using the formulations from Klein Breteler [188]:

$$\begin{aligned} t_{\mathrm{clay}} &= \frac{d_{\mathrm{c,outer}} - 0.2}{R_{\mathrm{e}}} && \text{with} \\ R_{\mathrm{e}} &= c_{\mathrm{d}} \cdot (H_{\mathrm{m0}} - 0.5) && \text{and} \\ c_{\mathrm{d}} &= 0.6 + \max(0; 8 \cdot (F_{\mathrm{sand}} - 0.7)) \end{aligned}$$

where t_{clay} is the time in hours for the clay to erode, $d_{\mathrm{c,outer}}$ is the cover layer thickness in meters, R_{e} and c_{d} are erosion parameters, and F_{sand} is the fraction of sand in the clay layer. Note that the effective clay layer thickness is reduced by 20 centimetres, as this is the part of the cover layer already included in the resistance-duration curve of the grass revetment.

After the clay layer has been eroded we account for the erosion of the sandy core in accordance with the formulations in Klein Breteler et al. [212]. This is a cliff erosion model where the erosion speed is given by:

$$\frac{\delta V_{\mathrm{e}}}{\delta t} = m_{\mathrm{core}} \frac{H_{\mathrm{m0}}^2}{T_{\mathrm{p}}} \left(\frac{0.15}{s^{1.3}} + (\tan\alpha)^{0.8} \cdot (135 - 1500s) \cdot \exp\left(-0.0091 \left(\frac{B_{\mathrm{t}}}{H_{\mathrm{m0}}} \right) \right) \right), \tag{D.3}$$

where V_{e} is the erosion volume in m m^{-3}, T_{p} is the peak wave period in seconds and B_{t} is the width of the terrace in front of the cliff in meters. m_{core} is a model factor for the eroded sand volume for which we assume a uniform distribution between the values suggested by Klein Breteler et al. [212] (see Table D.1).

Table D.1.: *Overview of (uncertain) input parameters for the load, wave impact and wave overtopping models.*

		Unit	Input values	Source
Load parameters				
h	Water level	m +ref	Gumbel($\beta = 0.20, \gamma = 15.146$)	Case-specific
H_{m0}	Significant wave height	m	Gumbel($\beta = 0.139, \gamma = 0.835$)	Case-specific
t_{storm}	Storm duration	h	12	Case-specific
t_h	Representative duration of high water	h	96	Case-specific
s	Wave steepness	-	0.04	Case-specific
Input for wave impact				
a_{closed}	parameter a for closed sod	m	Lognormal($\mu = 1.82, \sigma = 0.62$)	[200]
b_{closed}	parameter b for closed sod	h^{-1}	-0.035	[200]
c_{closed}	parameter c for closed sod	m	0.25	[200]
a_{open}	parameter a for open sod	m	Lognormal($\mu = 1.40, \sigma = 0.50$)	[200]
b_{open}	parameter b for open sod	h^{-1}	-0.035	[200]
c_{open}	parameter c for open sod	m	0.25	[200]
$d_{\mathrm{c,outer}}$	thickness of clay cover at outer slope	m	Lognormal($\mu = 0.5, \sigma = 0.05$)	
m_{core}	model factor for core erosion volume	-	Uniform($a = 0.82, b = 1$)	[212]
F_{sand}	Sand fraction of clay cover layer	-	Lognormal($\mu = 0.35, \sigma = 0.07$)	Case-specific
B_{min}	Minimum crest width at water level	m	4	Case-specific
$\tan\alpha_{\mathrm{t}}$	Slope of the terrace during erosion	-	0.1	[212]
$\tan\alpha$	Slope of the cliff during erosion	-	1	[212]
Input for wave overtopping				
$d_{\mathrm{c,inner}}$	thickness of clay cover at inner slope	m	0.5	-
$U_{\mathrm{c,closed}}$	Critical velocity closed sod	m/s	Lognormal($\mu = 8.0, \sigma = 1.0$)	[199]
$U_{\mathrm{c,open}}$	Critical velocity open sod	$\mathrm{m\,s}^{-1}$	Lognormal($\mu = 6.0, \sigma = 0.75$)	[199]
$U_{\mathrm{c,clay}}$	Critical velocity clay layer	$\mathrm{m\,s}^{-1}$	Lognormal($\mu = 0.85, \sigma = 0.085$)	[177]
C	Advance rate coefficient of head cut erosion	$\mathrm{s}^{-2/3}$	Lognormal($\mu = 1.7e-3, \sigma = 5.0e-4$)	[213]
A_0	Threshold parameter for head cut erosion	$\mathrm{m\,s}^{-1/3}$	Lognormal($\mu = 0.15, \sigma = 0.1$)	[213]
ω	Turbulence parameter	-	2.0	[152]

D

By numerically solving this differential equation for a combination of water level and wave height in time, one can obtain the time required for the width of the dike profile to reach a critical minimum width that is associated with failure (t_{core}). The erosion width E_{x} is the largest horizontal width between the original and eroded profile. For a profile without berm, this width is maximal at the still water level, and can be obtained through:

$$E_{\text{x}} = B_{\text{t}}\left(1 - \frac{\tan\alpha_{\text{t}}}{\tan\alpha}\right), \tag{D.4}$$

where α_{t} is the angle of the terrace slope, and α the original slope angle.

In this study failure is defined as the situation where the width at the water level h is smaller than the minimum crest width $B_{\min}$.

Combining these three failure modes the limit state function for grass erosion at the outer slope is given by:

$$g = t_{\text{storm}} - (t_{\text{grass}} + t_{\text{clay}} + t_{\text{core}}). \tag{D.5}$$

Peters [214] demonstrates that the impact point is typically between the water level and 1 significant wave height H_{s} below, with the most likely point around $0.5H_{\text{s}}$ below the water level. We represent this by a symmetrical triangular distribution with limits at h and $h - H_{\text{s}}$ in order to translate the evaluations of the limit state function g to a probability of failure along the slope. We determine the probability that the slope fails at coordinate y at section n when at state s using the following approach:

$$P_f(y|s) = \int_h \int_{H_{\text{s}}} P(g < 0|s,h,H_{\text{s}}) \cdot p(H_{\text{s}}|h) \cdot p(h) \cdot p(y|h,H_{\text{s}}) dH_{\text{s}} dh \tag{D.6}$$

where $p(h)$ and $p(H_{\text{s}}|h)$ are the probability density functions in accordance with Equation (D.1), $p(y|h,H_{\text{s}})$ is the triangular distribution of wave impact locations for coordinates y and $P(g < 0|s,h,H_{\text{s}}))$ is the fragility surface for state s, dependent on h and H_{s}.

We assume that for each zone z between coordinates y_i and y_j it holds that the failure probability for state s at section n is given by:

$$P_f(z,n,s) = \max P_f(y,n,s) \text{ for } y_i < y < y_j. \tag{D.7}$$

We evaluate this probability for each zone and each state. For state S_1 we assume a closed sod grass revetment, for state S_2 an open sod, for S_3 we only account for the erosion resistance of the clay layer and sandy core, whereas for S_4 we (conservatively) assume the clay layer provides no resistance.

D.1.2. Failure due to wave overtopping and overflow

Similar to wave impact at the outer slope, the erosion of the dike cover by overtopping waves can be described by two phases. During the initiation phase, the grass cover and the upper clay layer with roots is eroded. The upper 20 cm of the

cover is most important for the overall erosion resistance of the dike cover, as the resistance of the soil layer underneath, where less roots are present, is lower [184, 215]. Once the top layer is eroded, a vertical cliff forms at the damaged spot. This cliff will increase the load on the dike cover and start to expand, leading to advancement and deepening of cover layer erosion.

The erosion of the top layer is well described by scour erosion models [152, 215, 216]. In this study, we used the analytical grass-erosion model of van Bergeijk et al. [152] to calculate the failure probability at different locations along the slope. In this model, the load is determined by the flow velocity of the overtopping waves and a turbulence parameter $\omega = 2.0$. The strength of the dike cover is described using a distribution for the critical velocity U_c that depends on cover type and coverage (Table D.1).

Once the upper 20 cm is eroded, the advancement of erosion is calculated using the head-cut erosion model introduced by Natural Resources Conservation Service [217].The head-cut advance is calculated as:

$$\frac{dX}{dt} = C(A - A_0), \tag{D.8}$$

with C the advance rate coefficient [$s^{-2/3}$] and the threshold parameter A_0 [$m\,s^{-1/3}$]. The load A in $m\,s^{-1/3}$ depends on the specific discharge q [$m\,s^{-2}$] and the head cut height H [m]:

$$A = (qH)^{1/3}. \tag{D.9}$$

The specific discharge remains approximately constant along the dike profile and is calculated using the flow velocity u_0 and layer thickness h_0 on the dike crest that depend on the overtopping volume V [216, 218]:

$$q = u_0 \cdot h_0 \quad \text{with } u_0 = 4.5\,V^{0.34} \text{ and } h_0 = 0.133\,V^{0.5}. \tag{D.10}$$

van Hoven [213] calibrated the threshold parameter and advance rate coefficient based on overtopping and overflow experiments in the Netherlands. Based on the test results the values $A_0 = 0.1 m\,s^{-1/3}$ and $C = 1.5 \cdot 10^{-3} s^{-2/3}$ were recommended for plastic, erosion resistant clay as defined in TAW [219]. These are translated to lognormal distributions using a coefficient of variation of 20%, and such that the 5 and 95% values of A_0 respectively C coincide with the recommended values of [213].

The cover layer consists of grass on top of clay layer with a thickness $d_{c,inner}$ of 0.5 m in the considered case. The erosion of the first 20 cm (E_g) is calculated using the analytical grass-erosion model followed by the head-cut erosion model to calculate the horizontal migration of erosion hole E_h. The dike cover fails once the sandy dike core is reached. We do not include the residual strength of the sandy core, as erosion typically proceeds very quickly once overtopping erosion reaches the sandy core. The failure probability is calculated for the 4 states considered. We consider the failure probability for a closed sod (S_1), open sod (S_2), bare clay with good quality (S_3) and an animal burrow schematized as a hole with a depth and width of 30 cm (S_4). As such, in the fourth case we only consider advancement

and deepening of the burrow due to head-cut erosion. As the specific discharge is constant along the slope, so is the failure probability of this head-cut process.

The limit state function for states S_1, S_2 and S_3 is given by:

$$g = 0.2 + \frac{d_{\mathrm{c,inner}} - 0.2}{\sin \varphi} - (E_{\mathrm{g}} + E_{\mathrm{h}}) \tag{D.11}$$

with the thickness of the cover layer $d_{\mathrm{c,inner}}$, the angle of the landward slope φ, the erosion depth of the grass cover E_{h} and the head-cut erosion advancement E_{c}. For state 4, the limit state is given by:

$$g = \frac{d_{\mathrm{c,inner}} - 0.3}{\sin \varphi)} - E_{\mathrm{h}} \tag{D.12}$$

If the water level exceeds the crest level it is assumed, in line with the Dutch statutory safety assessment [52] that the flood defence fails due to overflow. It is assumed that this is independent of the state of the inner slope.

D.2. Costs of inspection, maintenance and structural upgrades

This supplement briefly describes the assumptions for deriving the cost of different inspection, maintenance and structural upgrades considered in the analysis.

D.2.1. Inspection and maintenance cost

For visual inspections the costs contain almost entirely out of personnel costs, which we assume to be 60 €/h in line with general personnel costs in the Netherlands. For repair works we use the same personnel costs and some additional costs for equipment and material. These might vary from case to case, but are typically not very high for repair works to grass revetments and assumed to be 50 €/h.

With the above cost components, we can derive the cost of an inspection I_i per km using the following equation:

$$C(I_i) = C_{\text{pers}} \cdot n_{\text{pers}} \cdot t(I_i) \tag{D.13}$$

where C_{pers} indicates the cost per person, n_{pers} the amount of personnel involved, $t(I_i)$ the time required to inspect 1 km.

For maintenance a similar type of function is used, except that now material and equipment costs (C_{equip}) are included:

$$C(M_i) = (C_{\text{pers}} \cdot n_{\text{pers}} + C_{\text{equip}}) \cdot t(M_i), \tag{D.14}$$

n_{pers} indicates the amount of personnel required, and $t(M_i)$ is the time required to repair 1 km.

Table D.2 shows the assumptions for the different types of inspection and maintenance. Note that all costs are computed per km of slope. The cost of M_3 is 0 as this is assumed to be part of I_1 (see main text). Estimates for the time and personnel required for inspection and maintenance are based on expert estimates and Chapter 6 (for inspections).

Table D.2.: *Assumed costs for inspection and maintenance per kilometre of slope.*

Action	Description	n_{pers} [-]	$t(I_i)$ [km h^{-1}]	$t(M_i)$ [h]	$C(I_i)$ [€/km]	$C(M_i)$ [€/km]
I1	General inspection by car	1	0.5		30	
I2	Condition inspection by foot	2	1		120	
I3	Burrowing inspection by foot	2	1		120	
M1	Repair of damaged spot	2		4		680
M2	Overhaul of all damaged spots	2		16		1920

D.2.2. Structural upgrades

We consider 4 types of structural upgrades: increase of the clay layer to 1 meter, decreasing the inner slope, heightening of the crest and targeted protection against

animal burrowing. For the latter we assume the costs are 100000 €/km. Other costs are estimated using Deltares [111], and indexed based on the price index for general construction works in the Netherlands in January 2021 [220]. For protection against animal burrowing we assume a lifespan of 30 years, for the other measures a lifespan of 100 years. Table D.3 shows how the resulting Equivalent Annual Cost (EAC) of the investment (EAI) were obtained.

Table D.3.: *Assumed Equivalent Annual Cost for investment (EAI) in structural upgrades.*

Intervention	C_A [k€/km]	**Lifespan** [yr]	**Annuity** [-]	EAI [-]
Increase of inner slope clay cover to 1 meter	600	100	49.72	12067
Decrease inner slope to 1:4.5	500	100	49.72	10056
Burrow-preventive measures at outer slope	100	30	23.68	4223
Burrow-preventive measures at inner slope	100	30	23.68	4223
Crest heightening 0.5 m	1400	100	49.72	28157
Crest heightening 1.0 m	1800	100	49.72	36202

List of Acronyms

CoV Coefficient of Variation. 72, 83, 98, 197

CPT Cone Penetration Test. 26, 197

DBN Dynamic Bayesian Network. 172, 197

DMC Dike Monitoring and Conditioning-system. 197

EAC Equivalent Annual Cost. 138, 150, 157, 158, 160–162, 166, 167, 172, 197, 228, 237, 241

EAD Expected Annual Damage. 98, 197

FORM First Order Reliability Method. 99, 197

I&M Inspection & Maintenance. vii, viii, 3, 5–7, 10, 137, 138, 148, 155, 156, 166, 167, 172, 175–177, 197

ILH International Levee Handbook. 14, 15, 112, 140, 197, 231

KDE kernel density estimation. 102–106, 197, 234

LCC Life Cycle Cost. 99, 170, 197

MIP Mixed Integer Programming. 49, 55, 57, 197

NPV Net Present Value. 197

PoD Probability of Detection. viii, 98, 117, 119, 120, 124, 126, 127, 131, 133–135, 149, 154, 157, 160, 162–167, 172, 175, 197, 219, 220, 235, 237, 240

SHM Structural Health Monitoring. 5, 92, 105, 197

TC Total Cost. vii, viii, 55, 74–76, 78, 79, 81–83, 86, 87, 106–109, 150, 170, 172, 175, 197, 233, 234

VoI Value of Information. viii, 59, 62, 66, 74–79, 81–88, 92, 94, 97, 99, 101–109, 163, 170, 171, 174, 175, 197, 233, 234

VSG Vertical Sandtight Geotextile. 39, 51, 82, 84, 86, 87, 170, 197, 206, 208, 209

List of Figures

List of Tables

Curriculum Vitæ

Wouter Johannes Klerk

05-08-1991 Born in Wageningen, the Netherlands.

Education

2011–2013 MSc. Civil Engineering (specialization Hydraulic Structures)
Delft University of Technology

2008–2011 BSc. Civil Engineering
Delft University of Technology

2002–2008 Pre-university Secondary Education (VWO)
Ichthus College, Veenendaal

Experience

2017–2021 PhD researcher
Section of Hydraulic Structures and Flood Risk
Faculty of Civil Engineering and Geosciences
Delft University of Technology, Delft

since 2013 Researcher/consultant,
Inland Water Systems
Deltares, Delft

2013 Graduate intern
Inland Water Systems
Deltares, Delft

2010–2012 Teaching assistant Structural Mechanics
Faculty of Civil Engineering and Geosciences
Delft University of Technology, Delft

List of Publications

Peer-reviewed journal publications

Related to this thesis

Klerk, W. J., Kanning, W., Kok, M., Bronsveld, J. & Wolfert, A. R. M. (2021). Accuracy of visual inspection of flood defences. *Structure and Infrastructure Engineering* doi:10.1080/15732479.2021.2001543.

Klerk, W. J., Kanning, W., Kok, M., & Wolfert, R. (2021). Optimal planning of flood defence system reinforcements using a greedy search algorithm. *Reliability Engineering and System Safety*, 207, 107344. doi:10.1016/j.ress.2020.107344

van der Krogt, M. G., **Klerk, W. J.**, Kanning, W., Schweckendiek, T., & Kok, M. (2020). Value of information of combinations of proof loading and pore pressure monitoring for flood defences. *Structure and Infrastructure Engineering*, 1–16. doi:10.1080/15732479.2020.1857794

Klerk, W. J., Schweckendiek, T., den Heijer, F., & Kok, M. (2019). Value of Information of Structural Health Monitoring in Asset Management of Flood Defences. *Infrastructures*, 4(3), 56. doi:10.3390/infrastructures4030056

Other

Sayers, P., Gersonius, B., den Heijer, F., **Klerk, W. J.**, Fröhle, P., Jordan, P., Ciocan, U. R., Rijke, J., Vonk, B., & Ashley, R. (2021). Towards adaptive asset management in flood risk management: A policy framework. Water Security, 12, 100085. doi:10.1016/j.wasec.2021.100085

Vonk, B., **Klerk, W. J.**, Fröhle, P., Gersonius, B., den Heijer, F., Jordan, P., Radu Ciocan, U., Rijke, J., Sayers, P., & Ashley, R. (2020). Adaptive Asset Management for Flood Protection: The FAIR Framework in Action. Infrastructures, 5(12), 109. doi:10.3390/infrastructures5120109

Gersonius, B., Vonk, B., Ashley, R. M., den Heijer, F., **Klerk, W. J.**, Manojlovic, N., Rijke, J., Sayers, P., & Pathirana, A. (2020). Maturity Improvements in Flood Protection Asset Management across the North Sea Region. Infrastructures, 5(12), 112. doi:10.3390/infrastructures5120112

Curran, A., de Bruijn, K. M., **Klerk, W. J.**, & Kok, M. (2019). Large scale flood hazard analysis by including defence failures on the Dutch river system. Water (Switzerland), 11(8). doi:10.3390/w11081732

Dupuits, E. J. C., **Klerk, W. J.**, Schweckendiek, T., & de Bruijn, K. M. (2019). Impact of including interdependencies between multiple riverine flood defences on the economically optimal flood safety levels. Reliability Engineering and System Safety, 191. doi:10.1016/j.ress.2019.04.028

Jonkman, S. N., Voortman, H. G., **Klerk, W. J.**, & van Vuren, S. (2018). Developments in the management of flood defences and hydraulic infrastructure in the Netherlands. Structure and Infrastructure Engineering, 1–16. doi:10.1080/15732479.2018.1441317

Klerk, W. J., Winsemius, H. C., van Verseveld, W. J., Bakker, A. M. R., & Diermanse, F. L. M. (2015). The co-incidence of storm surges and extreme discharges within the Rhine–Meuse Delta. Environmental Research Letters, 10(3), 035005. doi:10.1088/1748-9326/10/3/035005

Conference contributions

Related to this thesis

Klerk, W. J., & Adhi, R. A. (2021). Degradation of grass revetments: a comparison of field observations and structured expert judgement. FLOODrisk 2020 - 4th European Conference on Flood Risk Management. doi:10.3311/FloodRisk2020.8.2

Klerk, W. J., Kanning, W., Vos, R. J., & Kok, M. (2020). Risk-based maintenance of asphalt revetments on flood defences. In A. Chen, X. Ruan, & D. M. Frangopol (Eds.), Life-Cycle Civil Engineering: Innovation, Theory and Practice: Proceedings of the 7th International Symposium on Life-Cycle Civil Engineering (IALCCE 2020). CRC Press / Balkema - Taylor & Francis Group.

Klerk, W. J., Roscoe, K. L., Tijssen, A., Nicolai, R. P., Sap, J., & Schins, F. (2019). Risk based inspection of flood defence dams: An application to grass revetments. In D. M. Frangopol, R. Caspeele, & L. Taerwe (Eds.), Life-Cycle Analysis and Assessment in Civil Engineering: Towards an Integrated Vision - Proceedings of the 6th International Symposium on Life-Cycle Civil Engineering (IALCCE 2018). CRC Press / Balkema - Taylor & Francis Group.

Klerk, W. J., Kanning, W., van Veen, N.-J., & Kok, M. (2019). Influence of monitoring on investment planning of flood defence systems. In J. Ching, D.-Q. Li, & J. Zhang (Eds.), Proceedings of the 7th International Symposium on Geotechnical

Safety and Risk (ISGSR 2019): State-of-the-Practice in Geotechnical Safety and Risk (pp. 792–797). Research Publishing, Singapore. doi:10.3850/978-981-11-2725-0_IS4-10-cd

Klerk, W. J., Kanning, W., & Kok, M. (2018). Time-dependent reliability in flood protection decision making in the Netherlands. Safety and Reliability – Safe Societies in a Changing World: Proceedings of ESREL 2018, 3167–3174.

Other

Thöns, S., **Klerk, W. J.**, & Köhler, J. (2019). Case studies for quantifying the value of structural health monitoring information: Lessons learnt. IABSE Symposium, Guimaraes 2019: Towards a Resilient Built Environment Risk and Asset Management, 345–352. doi:10.2749/guimaraes.2019.0345

Gersonius, B., Ashley, R., Den Heijer, F., **Klerk, W. J.**, Sayers, P., & Rijke, J. (2019). Asset management maturity for flood protection infrastructure: A baseline across the North Sea Region. Life-Cycle Analysis and Assessment in Civil Engineering: Towards an Integrated Vision - Proceedings of the 6th International Symposium on Life-Cycle Civil Engineering (IALCCE 2018).

Klerk, W. J., Pot, R., van der Hammen, J., & Wojciechowska, K. A. (2019). A framework for assessing information quality in asset management of flood defences. In D. M. Frangopol, R. Caspeele, & L. Taerwe (Eds.), Life-Cycle Analysis and Assessment in Civil Engineering: Towards an Integrated Vision - Proceedings of the 6th International Symposium on Life-Cycle Civil Engineering (IALCCE 2018). CRC Press / Balkema - Taylor & Francis Group.

Jonkman, S. N., Voortman, H. G., **Klerk, W. J.**, & van Vuren, S. (2016). Developments in the management of flood defences and hydraulic infrastructures in the Netherlands. In J. Bakker, D. M. Frangopol, & K. van Breugel (Eds.), Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of the Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016). CRC Press / Balkema - Taylor & Francis Group.

Klerk, W. J., & den Heijer, F. (2016). A framework for life-cycle management of public infrastructure. In J. Bakker, D. M. Frangopol, & K. van Breugel (Eds.), Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of the Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016). CRC Press / Balkema - Taylor & Francis Group.

Klerk, W. J., Kanning, W., van der Meer, M. T., & Nieuwenhuis, J. W. (2016). Structural health monitoring in life-cycle management of dikes: a case study in the

north of the Netherlands. In J. Bakker, D. M. Frangopol, & K. van Breugel (Eds.), Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of the Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016). CRC Press / Balkema - Taylor & Francis Group.

van der Wiel, W. D., Klanker, G., **Klerk, W. J.**, Persoon, E., Wessels, J., & de Wit, A. (2016). A system approach for replacement strategy of hydraulic structures. In J. Bakker, D. M. Frangopol, & K. van Breugel (Eds.), Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of the Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016). CRC Press / Balkema - Taylor & Francis Group.

Klerk, W. J., den Heijer, F., & Schweckendiek, T. (2015). Value of information in life-cycle management of flood defences. In B. Sudret, L. Podofillini, B. Stojadinovic, E. Zio, & W. Kroger (Eds.), Safety and Reliability of Complex Engineered Systems. Proceedings of the ESREL 2015 conference (pp. 931–938). CRC Press.

Klerk, W. J., Kok, M., de Bruijn, K. M., Jonkman, S. N., & van Overloop, P. J. (2014). Influence of load interdependencies of flood defences on probabilities and risks at the Bovenrijn/IJssel area, The Netherlands. 6th International Conference on Flood Management, October. doi:10.13140/2.1.2233.0881

De Bruijn, K. M., Diermanse, F., **Klerk, W. J.**, & der Most, H. (2014). Societal flood risk assessments: An advanced probabilistic method and its application to the Rhine-Meuse delta. 6th International Conference on Flood Management, 1(October). doi:10.13140/2.1.1184.5125