Geohydrology and safety of Dutch canal dikes: threats from above

From monitoring to probabilistic risk assessments

Bart Strijker

Geohydrology and safety of Dutch canal dikes: threats from above

From monitoring to probabilistic risk assessments

Bart Strijker

Geohydrology and safety of Dutch canal dikes: threats from above

From monitoring to probabilistic risk assessments

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus,
Prof.dr.ir. H. Bijl,
chair of the Board for Doctorates
to be defended publicly on
Thursday, 4 June 2026, 12:30

by

Bart STRIJKER

This dissertation has been approved by the promotors.
Composition of the doctoral committee:

Rector Magnificus	chairperson
Em.prof.dr.ir. M. Kok	Delft University of Technology, promotor
Prof.dr.ir. S.N. Jonkman	Delft University of Technology, promotor

Independent members:

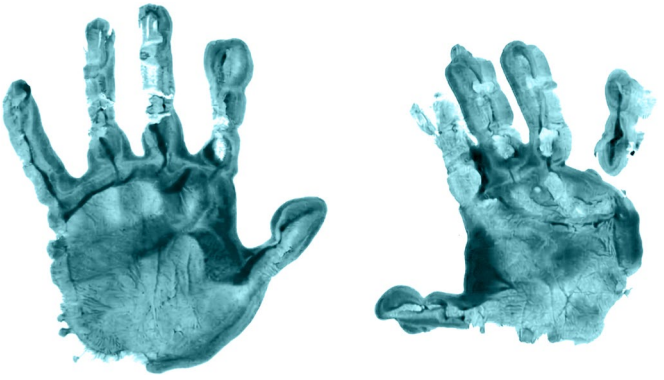
Prof.dr. T.A. Bogaard	Delft University of Technology
Prof.dr.ir. R.D.J.M. Steenbergen	Ghent University / TNO
Prof.dr.ir. M.J. van der Ploeg	Wageningen University & Research
Dr.ir. T. Schweckendiek	Delft University of Technology / Deltares
Dr. O. van Dam	Stichting Toegepast Onderzoek Waterbeheer
Prof.dr.ir. R. Uijlenhoet	Delft University of Technology, *reserve member*

This research was funded by, among others, Stichting Toegepast Onderzoek Waterbeheer

Keywords: Dike safety, stability, monitoring, droughts and floods, geohydrology, head response, time series modelling, reliability updating, credible failure probabilities.
Printed by: Proefschriftspecialist | proefschriftspecialist.nl
Layout and design: Henry Smaal, persoonlijkproefschrift.nl

Cover illustration: Knowledge, illustrated by the candle, throws a shadow of a dike onto the wall. We only see that shadow. Not the dike itself, not reality. Just a projection of what we currently know. Like Plato's cave: we mistake the shadow for truth.

ISBN 978-94-6518-321-3
An electronic version of this dissertation is available at: http://repository.tudelft.nl



Contents

Acknowledgement

Why have I spent four years, and really my entire working life, on dikes and polders? It often comes down to soil and water, nothing more. Yet these small mounds of earth matter deeply to the Netherlands. They have inspired and fascinated me for years: the artificial green landscape, the engineering, the field measurements, and the history. And of course, the people.

I want to thank everyone who contributed over the past years. From quick coffee chats to reviewing a text, from helping in the field to thinking along with scientific or practical questions, your involvement meant a great deal.

In particular, I want to thank Matthijs Kok. You gave me room to do research and the freedom to meander like a river: the banks as the boundaries, but the course shifting each year. Our meetings were irregular and sometimes seemed unstructured, yet I always left with more energy. I also want to thank Bas Jonkman for your sharp comments that pushed me forward. Even though the storyline still felt messy to me, you quickly grasped the core and helped me find a practical way forward.

Beyond academia, I learned a lot from practice. My special thanks go to Douwe Yska, Erik Vastenburg and Henk van Hemert for your time, interest and many conversations. We talked about the risk, or the lack of it, of cracks in dikes, and about the lack of measurements and large spatial variations in groundwater levels, and how to deal with this in practice.

I also want to thank HKV. The culture at HKV is unique, as is its flexibility, engagement, and ability to switch quickly, in which Wouter ter Horst and Jan Stijnen played an important role for me. My colleagues and the projects offered valuable input and inspiration to connect my research to real-world practice. I also want to thank Bas Kolen explicitly, who patiently listened to my surprises, wonders, and at times irritation about the Dutch flood safety sector. And Joost Stenfert, with whom I reflected on work, HKV and much more. I want to mention Chris Geerse as well, who passed away in 2024. He taught me what it means to take time for someone. He spent many hours explaining probabilistic methods to me, and I always admired how he combined a strong theoretical and statistical background with a practical, straight-to-the-point approach to move forward.

I am grateful as well for the involvement of STOWA, Rijkswaterstaat and the Dutch water authorities of Schieland and the Krimpenerwaard, Rijnland, Delfland and Hollands Noorderkwartier. From enabling monitoring efforts to sharing data and providing financial support; this research would not have been possible without you.

I am writing this acknowledgement in the summer of 2025, from our bedroom at Vretgränd 4A in Uppsala. I spent the last two and a half years of my PhD working from Sweden. That was only possible thanks to the flexibility of both the university and HKV. This period abroad did not come from my PhD but from Michelle, my partner, who had the opportunity to work on a project here. Together with our son Lio, and later our daughter Tosca, who was born here, these past years have been fantastic; intense, full of experiences, with highs and lows, and *kan het nou nooit eens normaal zijn*. Sweden has earned a place in my heart, and in ours as a family. The *onverstoorbaarheid* en *oordeelloosheid* of Swedish people and nature, and the time and space to be together: dancing at the Danspaviljongen with overpriced beers at Brasserie21, playing at Stadsträdgården's playground, having fika at Fågelsången, walking in Hammarskogen, and weekend trips to Stockholm and the Archipelago. In December 2025, we will return to the Netherlands, bringing a bit of the Swedish mindset with us and leaving behind, thankfully, the dark winter days with sunsets before 3 p.m.

These were turbulent years, but I always enjoyed working on my PhD. The final months are still ahead. And since my talent lies in finishing things to about 80%, I'm curious how the final stretch of this thesis will unfold. What may help is remembering that it doesn't need to be a masterpiece and that research is always a work in progress. It is something that shifts, tilts, and grows, and is not the most important thing in life:

"Beliefs are hypotheses to be tested and not treasures to be guarded."
– P.E. Tetlock & D. Gardner

"Nothing in life is as important as you think it is, while you are thinking about it."
– D. Kahneman

Bart Strijker
Uppsala, July 2025

Summary

Flood risk is of key interest to societies in low-lying areas such as the Netherlands, where a wide variety of dikes protect people and property from flooding. These dikes can vary in terms of hydraulic loads, subsoil conditions, and dimensions. This thesis focuses on so-called canal dikes, which are located along drainage canals in polders, mainly in the northern and western regions of the Netherlands. Safety assessment outcomes, along with recent dike failures, indicate that inner-slope instability is a key failure mechanism that primarily affects canal-dike safety. This instability can occur under both very dry and very wet conditions. Assessing the inner-slope instability of dikes remains challenging due to large uncertainties in geotechnical and geohydrological parameters.

The aim of this thesis is to enhance the geohydrological understanding of canal dikes through monitoring and observation-based analysis, thereby improving the assessment of their stability and the safety of dike systems as a whole. To this end, this study established two unique datasets and analysed them using data-analysis techniques, time-series modelling, and probabilistic modelling. The datasets include: 1) a multi-year detailed monitoring series of soil moisture levels and hydraulic heads at ten canal dikes, and 2) a nationwide collection of over one hundred hydraulic head time series in canal dikes from about 50 monitoring sites.

The detailed monitoring series included two extremely dry summers (2020 and 2022) and highlighted the impact of meteorological variations on the subsurface water conditions. Hydraulic heads at inner-slope and toe monitoring points showed strong seasonal variation. During winter, conditions were near saturation, with about half of the 37 measurement points measuring heads within 20 cm of the surface. During dry summers, however, these heads could drop down to nearly 2m, despite outside water levels remaining nearly constant. During these summers, non-hydrostatic hydraulic head levels were observed within dike bodies. These conditions can be detrimental to dike stability and are often not accounted for in safety assessments for drought situations. Furthermore, the effectiveness of various meteorological drought indicators for subsurface water conditions was evaluated: the precipitation deficit proved to be the most reliable measure and outperforms the standardised drought indicators (*SPEI* and *SPI*). However, the precipitation deficit remains an indicator and does not account for geohydrological processes such as soil storage capacity.

By analysing the nationwide dataset in combination with time-series models, variations in hydraulic head responses across canal dikes were quantified in relation to precipitation and evaporation. Various model structures were evaluated, and a non-linear model performed best. After calibration, the models were evaluated using reliability criteria, including a minimum goodness-of-fit and a sufficient time-series

length compared to model parameters. This resulted in a set of 35 reliable time-series models for further analysis. Within this model set, four clusters of dikes were identified, each showing similar peak head responses within the cluster and distinct responses between clusters. The differentiating factor was the response times, defined as the time required for 95 % of an impulse's influence to dissipate. As a result, heavy precipitation events lead to different peak head responses: longer response times cause peak heads to be driven by more prolonged precipitation events, resulting in peaks that typically occur later in the winter. The differences and similarities in peak head responses between dikes and identified clusters were related to physical dike characteristics. While subsurface material and dike width appeared to influence head-response variation at the cluster level, their presence across multiple clusters indicates that they are not definitive predictors. Furthermore, peak head statistics across various dikes indicated that extreme and yearly occurring load conditions are relatively close to each other, with a median decimate height (the difference between head levels with 100-year and 10-year return periods) of 15 cm and a range of 5 to 50 cm. With climate change driving higher winter precipitation and summer evaporation, head statistics are changing. It was found that by 2100, extreme peak heads are expected to occur between three times less frequently and eight times more frequently, depending on the climate scenario and the type of canal dike.

Next, the developed time-series models and the variation in head responses were used to hindcast 60 years of hydraulic head levels in Dutch canal dikes, based on historical gridded meteorological data, and estimated spatial dependencies between peak heads at various spatial scales. The spatial dependencies for loads within dike systems were quantified using the length-effect factor of peak heads. This factor indicates how much more frequently peak heads occur at larger spatial scales than at the section level. Because monitoring is limited, the actual head response is unknown for most dike sections, so upper and lower bounds were estimated using two configurations: one assuming similar head responses within a square kilometre (homogeneous), and one assuming varying responses (heterogeneous). The length-effect factors of peak heads were calculated for the entire canal-dike system and three polders with various dike lengths and areas. At the polder scale, the length-effect of peak heads primarily arises from variations in head responses, while spatial variability in heavy precipitation events contributes moderately, though this contribution increases with polder size. Spatial weather variability can double the length-effect factor of peak heads at the polder scale, whereas variation in head responses can increase the factor by a factor of four to five. At the scale of the entire canal-dike system, variations in head responses have similar effects, but the length-effect of peak heads is dominated by spatial weather variability. The factor increases with event extremity, from about 6 for annually occurring heads to around 40 for heads with an exceedance frequency of 1/100 per year. The length-effect

of peak heads also varies across head-response types: for dikes with shorter response times (the time for 95% of the impulse to dissipate), shorter precipitation durations become more relevant and are associated with smaller spatial extents, resulting in larger length-effect factors.

Building on the improved understanding of the geohydrological response of canal dikes, including head statistics, the implications for dike safety were further investigated. The impact of small load variations on reliability updating and on assessing credible failure probabilities was analysed. First, this was examined conceptually using hypothetical dikes with various load and strength configurations. It was found that the impact of reliability updating increases as load variations decrease, regardless of the prior failure probability. This occurs because a large contribution to the failure probability originates from relatively frequent load levels, for which conditional failure probabilities are most strongly reduced through reliability updating. Second, these findings were confirmed by case studies examining the stability of two canal dikes and two river dikes in the Netherlands, which varied in terms of fragility curves and load variations. The two canal dikes, with decimate heights below 10 cm, benefited most from reliability updating, whereas the updated failure probabilities of the river dikes changed only slightly. For the river dikes considered, with decimate heights of the phreatic line up around 30 and 60 cm, even extreme load events with a 1/100 exceedance probability per year that occurred and survived the dike had minimal impact on reducing failure probabilities through reliability updating. Finally, a credibility check was introduced for dikes that have remained stable for decades, without any noticeable damage. For such dikes, load levels with return periods of up to 10 years are not expected to account for more than a limited share of the failure probability. If they do, this suggests that the estimated failure probability is not credible and that there is an imbalance between the load variability and strength uncertainty. This imbalance occurs when the uncertainty in dike strength is more than 1.5 times larger than the variation in loads, for example, an inverse gradient of the fragility curve of 0.15 m for a dike with a load variation of 0.1 m. In these situations, reducing uncertainties in the dike's strength lowers the inverse gradient of the fragility curve, resulting in more credible failure probability estimates. This can be achieved, for example, by monitoring load levels in combination with reliability updating, or by gathering more detailed data on soil strength.

Samenvatting

Overstromingsrisico's zijn van groot belang voor samenlevingen in laaggelegen gebieden zoals Nederland, waar een breed scala aan dijken inwoners en eigendommen beschermt tegen overstromingen. Deze dijken verschillen in hydraulische belasting, ondergrond en afmetingen. Dit proefschrift richt zich op boezem- en polderkaden (hierna: kaden), gelegen langs afwateringskanalen in polders, voornamelijk in Noord- en West-Nederland. De resultaten van veiligheidsbeoordelingen, samen met recente faalgebeurtenissen, wijzen erop dat macro-instabiliteit van het binnentalud het belangrijkste faalmechanisme is voor de veiligheid van kaden. Deze instabiliteit kan optreden onder zowel zeer droge als zeer natte omstandigheden. Het beoordelen ervan blijft uitdagend vanwege grote onzekerheden in geotechnische en geohydrologische parameters.

Het doel van dit proefschrift is om de geohydrologische respons van kaden te kwantificeren op basis van monitoring en waarnemingen, en deze inzichten door te vertalen naar de beoordeling van stabiliteit en faalkansen op dijksectie- en systeemniveau. Daartoe zijn twee unieke datasets opgesteld en geanalyseerd met behulp van data-analysetechnieken, tijdreeksmodellering en probabilistische modellering. De datasets omvatten: 1) een meerjarige gedetailleerde monitoringsreeks van bodemvochtgehaltes en grondwaterstanden bij tien kaden, en 2) een landelijke verzameling van meer dan honderd tijdreeksen van grondwaterstanden in kaden van ongeveer 50 monitoringslocaties.

De gedetailleerde monitoringsreeks omvat twee extreem droge zomers (2020 en 2022) en benadrukt de impact van meteorologische variaties op de grondwaterstanden en bodemvochtgehaltes. Grondwaterstandmetingen in het binnentalud en aan de teen tonen een sterke seizoensvariatie. In de winter zijn kaden vrijwel verzadigd, waarbij op ongeveer de helft van de 37 monitoringspunten de grondwaterstand binnen 20 cm van het maaiveld ligt. Tijdens droge zomers dalen deze tot bijna 2 m, terwijl de buitenwaterstand nagenoeg constant blijft. In deze perioden ontstaan ook niet-hydrostatische drukverlopen in het dijklichaam. Dit kan de stabiliteit negatief beïnvloeden en wordt vaak niet meegenomen in veiligheidsbeoordelingen voor droogtesituaties. Daarnaast is de effectiviteit van verschillende meteorologische droogte-indicatoren voor bodemvochtgehaltes en grondwaterstanden in kaden geëvalueerd: het neerslagtekort blijkt de meest betrouwbare maatstaf en presteert beter dan de gestandaardiseerde droogte-indicatoren (SPEI en SPI). Het blijft echter een meteorologische indicator en houdt geen rekening met geohydrologische processen, zoals de bergingscapaciteit van de bodem.

Door de landelijke dataset te analyseren met tijdreeksmodellen, zijn variaties in de grondwaterrespons in kaden gekwantificeerd in relatie tot neerslag en verdamping.

Verschillende modelstructuren zijn geëvalueerd, en een niet-lineair model presteert het best. Na kalibratie zijn de modellen geëvalueerd aan de hand van betrouwbaarheidscriteria, waaronder een minimale 'goodness-of-fit' en een voldoende tijdreekslengte in relatie tot de gevonden modelparameters. Dit leidt tot een set van 35 betrouwbare tijdreeksmodellen voor verdere analyse. Binnen deze modelset zijn vier clusters van kaden geïdentificeerd met vergelijkbare responsen binnen elk cluster en duidelijke verschillen tussen clusters. De responstijd vormt de belangrijkste onderscheidende factor, gedefinieerd als de tijd die nodig is voordat 95% van de invloed van een impuls op de grondwaterstand is uitgewerkt. Lange responstijden leiden tot piekgrondwaterstanden die worden bepaald door langdurige neerslag en later in de winter optreden. De verschillen en overeenkomsten in grondwaterrespons tussen dijken en geïdentificeerde clusters zijn gerelateerd aan fysieke dijkkenmerken. Terwijl de ondergrond en dijkbreedte de variatie in grondwaterrespons op clusterniveau lijken te beïnvloeden, duidt hun aanwezigheid in meerdere clusters erop dat zij geen bepalende voorspellers zijn. Verder gaven de afgeleide grondwaterstandstatistieken van de verschillende kaden aan dat extreme en jaarlijks voorkomende belastingcondities relatief dicht bij elkaar liggen, met een mediane decimeringshoogte (het verschil tussen grondwaterstand met herhalingstijden van 100 en 10 jaar) van 15 cm en een bereik van 5 tot 50 cm. Met klimaatverandering die zorgt voor hogere winterneerslag en zomerse verdamping, veranderen de statistieken van grondwaterstanden. Tegen 2100 zullen extreme grondwaterstanden naar verwachting drie keer minder vaak tot acht keer vaker voorkomen, afhankelijk van het klimaatscenario en het type kade.

De ontwikkelde tijdreeksmodellen maken het mogelijk om 60 jaar aan grondwaterstanden te reconstrueren op basis van historische meteorologische rasterdata, en daarmee de ruimtelijke afhankelijkheden tussen piekgrondwaterstanden op verschillende ruimtelijke schalen in te schatten. Deze ruimtelijke afhankelijkheden zijn gekwantificeerd met behulp van de lengte-effectfactor van piekgrondwaterstanden. Deze factor geeft aan hoeveel vaker piekgrondwaterstanden voorkomen op grotere ruimtelijke schalen dan op vakniveau. Omdat metingen schaars zijn, blijft de werkelijke respons voor de meeste dijken onbekend. Daarom zijn boven- en ondergrenzen geschat met twee configuraties: een homogene respons binnen één vierkante kilometer en een heterogene, variërende respons. De lengte-effectfactoren van piekgrondwaterstanden zijn berekend voor het volledige kadensysteem en voor drie polders met verschillende afmetingen. Op polderschaal wordt het lengte-effect van piekgrondwaterstanden vooral bepaald door variatie in grondwaterresponsen. De ruimtelijke variabiliteit van hevige neerslaggebeurtenissen speelt een kleinere rol, hoewel deze rol toeneemt met de poldergrootte. Ruimtelijke weersvariabiliteit kan de lengte-effectfactor van piekgrondwaterstanden op polderschaal verdubbelen, terwijl variatie in de grondwaterrespons deze een factor vier tot vijf kan verhogen. Op de schaal van het volledige

kadensysteem heeft de variatie in de grondwaterrespons een vergelijkbaar effect, maar domineert de ruimtelijke weersvariabiliteit het lengte-effect van piekgrondwaterstanden. De lengte-effectfactor neemt toe met de extremiteit van de gebeurtenis, van ongeveer 6 voor jaarlijks voorkomende grondwaterstanden tot ongeveer 40 voor grondwaterstanden met een overschrijdingsfrequentie van 1/100 per jaar. Het lengte-effect van piekgrondwaterstanden varieert ook tussen typen grondwaterrespons: kaden met korte responstijden reageren vooral op kortdurende buien met kleinere ruimtelijke omvang, wat leidt tot grotere lengte-effectfactoren.

Voortbouwend op het verbeterde begrip van de geohydrologische respons van kaden, inclusief grondwaterstandstatistieken, zijn de implicaties voor dijkveiligheid verder onderzocht. De impact van kleine belastingvariaties op faalkans updating op basis van een overleefde belastinggebeurtenis en op het beoordelen van geloofwaardige faalkansen is geanalyseerd. Conceptuele analyse met hypothetische dijken laat zien dat het effect van faalkans updating toeneemt bij kleinere belastingvariaties, ongeacht de a-priori faalkans. Een groot deel van de faalkans wordt namelijk bepaald door relatief frequente belastingen, waarvan de conditionele faalkansen het sterkst worden gereduceerd door faalkans updating. Casestudies van twee kaden en twee rivierdijken in Nederland bevestigen dit beeld, met variatie in sterkte-onzekerheid (fragility curves) en belastingvariaties. Kaden met een kleine variatie in de belasting (decimeringshoogte < 10 cm) profiteren het meest van faalkans updating. Bij rivierdijken met grotere belastingvariatie (decimeringshoogte van circa 30 en 60cm) verandert de faalkans nauwelijks, zelfs na extreme belastinggebeurtenissen met een overschrijdingsfrequentie van 1/100 per jaar die optraden en door de dijk zijn overleefd. Tot slot is een geloofwaardigheidscheck ontwikkeld voor dijken die decennialang stabiel zijn gebleven, zonder noemenswaardige schade. Voor dergelijke dijken zouden belastingen met een herhalingstijd tot 10 jaar slechts voor een beperkt deel moeten bijdragen aan de berekende faalkans. Is de bijdrage van frequente belastingen groot, dan wijst dit op een onbalans tussen belastingvariabiliteit en sterkte-onzekerheid. Deze onbalans ontstaat wanneer de onzekerheid in sterkte meer dan 1,5 keer groter is dan de variatie in belasting, bijvoorbeeld een fragility curve waarvan de inverse helling 1,5 keer groter is dan de decimeringshoogte van de belasting. Het verkleinen van de onzekerheid in de sterkte leidt tot steilere fragility curves en daarmee tot geloofwaardigere faalkansschattingen. Dit kan bijvoorbeeld door belastingparameters te monitoren in combinatie met faalkans-updating, of door aanvullende gegevens over bodemopbouw en bodemsterkte te verzamelen.

1

Introduction

1.1 Flood management

Flooding is one of the leading climatic threats to people's livelihoods, affecting development prospects worldwide (IPCC, 2022; Jevrejeva et al., 2018). According to the EM-DAT disaster database, it is the most frequent type of natural disaster (Guha-Sapir et al., 2016). Over recent decades, the average number of fatalities per flood event at the global scale decreased (Jonkman et al., 2024), in contrast to the rising trend in economic damages caused by floods (Coronese et al. 2019). The decreasing trend in fatalities is attributed to improved flood risk management practices over the last decades, such as including better warning, forecasting and early warning communication. In contrast, the rising trend of economic damages is largely driven by socio-economic developments, like increasing economic value in flood-prone areas, as well as the impacts of climate change, including sea level rise and more frequent extreme rainfall events (Bouwer, 2018; Merz et al., 2021). In response to floods, societies apply a broad range of flood risk reduction and adaptation strategies. These include both non-structural approaches, such as spatial planning, insurance, and the flood-proofing of assets (Green et al., 2000; Poussin et al., 2012), and structural measures, such as dikes[1] and dams. In highly developed low-lying regions with large concentrations of people and economic activity, structural protection often forms the foundation of flood risk management (O'Dell et al., 2021; Rentschler et al., 2022). As a result, many flood-prone areas around the world are protected by extensive systems of dikes, for example, the Netherlands.

The Netherlands is a highly developed low-lying region which relies heavily on structural flood defenses as over 50% of Dutch land is flood-prone and protected by an extensive system of dikes, dunes, and dams (PBL Netherlands Environmental Assessment Agency, n.d.). For these flood defenses, a distinction can be made between primary and regional flood defenses. While primary flood defenses protect the hinterland from floods caused by major water bodies such as the sea, lakes, and large rivers, regional flood defenses are located along smaller inland water bodies, see Figure 1.1. Both primary and regional flood defenses have safety standards that are based on a risk-based approach and are legally embedded in, respectively, the Dutch Water Act and provincial regulations (Kok et al., 2017).

Assessing dike safety requires a clear understanding of both the loading conditions and the resistance of dikes and is essential for making decisions about management and maintenance, dike reinforcement and emergency measures during hazardous events (Schweckendiek et al., 2015; Vrijling et al., 2011). Therefore, the safety assessments should be credible. This means that the estimated safety, often in terms of failure

1 Dike is used throughout this thesis for earthen flood defense structures that retain water; equivalent terms are levee or (flood) embankments

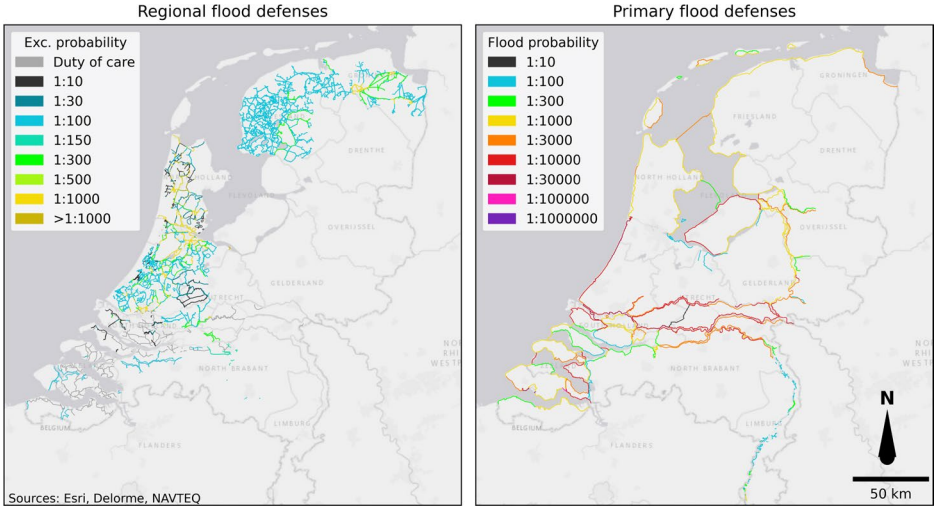


Figure 1.1 Safety standards for the Dutch regional[2] (left) and primary (right) flood defenses. For regional flood defenses, the standard is expressed as an exceedance probability of the water level that must be safely withstood. In contrast, for the primary flood defenses, it is the probability of flooding.

probability per year, is realistic and justifiable, and does not contradict the (historical) observed behaviour and performance of the dike. Safety assessments that are credible prevent dikes from being unnecessarily estimated as unsafe, or unjustifiably considered safe. It supports investment decisions aimed to protect societies in embanked areas against floods by providing a well-founded safety profile both now and in the future.

With a changing climate and without adaptation and measures taken, flood risks can be amplified worldwide in several ways. Rising global temperatures have accelerated sea-level rise, increasing flood risks in coastal areas due to, among others, more extreme high water levels. In addition, higher sea levels can reduce the drainage capacity of surface water and groundwater in low-lying urban areas, thereby increasing the risk of pluvial flooding (IPCC, 2022; Nicholls & Cazenave, 2010). At the same time, warmer air holds more moisture, resulting in more frequent and intense extreme precipitation events that can increase extreme discharge levels in river catchments and drainage systems (Merz et al., 2021; Bouwer, 2018). Furthermore, these changing precipitation and temperatures can also impact the performance of dikes, as their strength and stability can be significantly impacted by climate change through several thermo-

2 The geodata of the regional flood defenses was compiled from multiple sources, as no single dataset was fully comprehensive. It is known that some features are missing, but it provides an overview of safety standards variations.

hydro-mechanical processes (Elia et al., 2017; Vardon, 2015). In low-lying delta areas, such as the Netherlands, these changes in the climate, combined with land subsidence and land-use pressures, can further increase flood risks without additional adaptation measures. To guarantee dike safety, reinforcement measures may be required to meet the prescribed safety levels. Understanding how climate change influences the dike's performance is essential for evaluating the effectiveness of these measures.

1.2 Background on Dutch canal dikes

The Dutch landscape, like many delta areas in the world, was shaped by the dynamics of water from the rivers and seas. This also formed a threat to the inhabitants of the lowlands. They protected themselves against floods with dikes and dams. This engineering interfered with the natural drainage patterns of the lowlands and resulted in challenges for water management. For example, how to deal with excessive rainfall? This led to more measures to regulate the water and slowly led to the creation of polders (Schultz, 2024).

Polder areas are characterized by their unique water management, as they lie below sea or river levels and are separated from the surrounding hydrological regime by dikes. Therefore, rainfall must be actively pumped out of the polder to prevent flooding, as shown in the upper panel in Figure 1.2. This is managed through an inland drainage system comprising weirs, pumping stations, and canals. Rainfall in the polder is pumped into higher-level drainage canals, known as the *boezem*, and from there, it is discharged into rivers, lakes or the sea. The canal water levels are regulated with only minor fluctuations, typically within a range of decimeters. These water levels can rise above the polder elevation, because of the formation or construction of dikes along the canals. They can be formed by the subsidence of the polder or constructed to reclaim land from lakes, and are called canal dikes[3] (Den Haan & Kruse, 2007).

In total, there are more than 10,000 km of regional flood defenses in the Netherlands, of which the bulk consists of canal dikes (Pleijster et al., 2015), which are mainly located in the northwestern part of the country, as shown in Figure 1.3. Canal dikes consist mainly of soft soils, such as clay and peat, which were naturally deposited in the Netherlands and still present in today's topsoil. These dikes can be over a century old and have been reinforced over the years, resulting in heterogeneous dike bodies with varying soil compositions, including gravel and debris, like potsherds and bricks. Furthermore, the dikes are grass-covered, with often a road running along the crest, and the crests of the canal dikes generally rise only a few decimeters above the normal daily

3 In Dutch, these dikes are called *polder- and boezemkaden*.

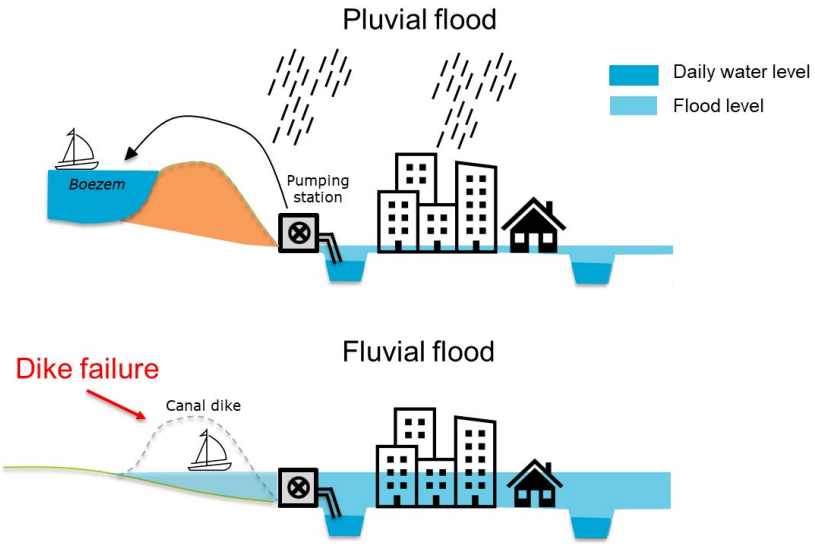


Figure 1.2 Cross-sectional schematic of a Dutch polder system. The top panel illustrates a pluvial flood, where extreme rainfall exceeds the drainage capacity of the pumping stations in a polder. The bottom panel depicts a fluvial flood resulting from a canal dike breach, allowing water from the higher boezem to inundate the polder.

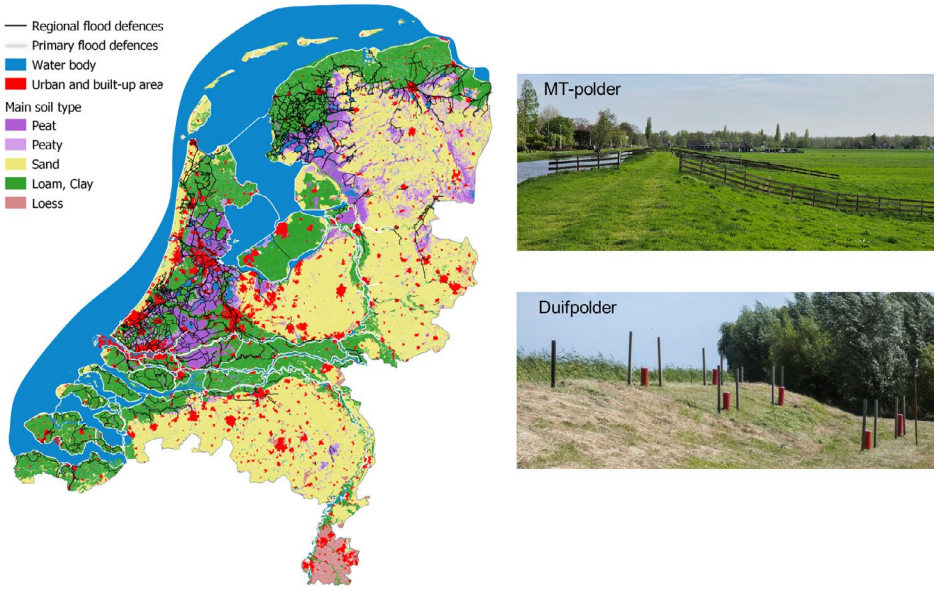


Figure 1.3 Left: Map of the Netherlands showing the primary and regional flood defenses, water bodies and urban areas, including the main soil types from the Dutch soil map, providing an indication of the soil composition up to approximately one meter below the surface (Data source: Wageningen Environmental Research, 2023). Right: two photographs of canal dikes along two different polders.

canal water level, as the canal levels are actively controlled by pumping stations. This means that they are exposed to "high" water levels on a daily basis. Safety assessments of these dikes are done to regulate flood risk levels and are typically conducted once every 12 years by the water authorities. Within these assessments, various failure mechanisms are considered, including insufficient crest height, inner-or outer-slope instability and backward erosion piping. Previous safety assessments have shown that hundreds of kilometers of canal dikes either currently do not meet the required safety standards or have recently been reinforced to comply (Province of South Holland, 2023; Hoogheemraadschap Hollands Noorderkwartier, 2025).

1.2.1 Historical canal dike failures

Catastrophic dike failures have occurred throughout Dutch history, triggered by storm surge, ice drift and heavy rainfall and drought (Özer et al.,2019a; Van Baars & Van Kempen, 2009). These naturally varying forces have led to failures through mechanisms, such as overflow and overtopping causing soil erosion, external erosion, piping and instability of the inner slope. Between 1900 and 2005, there were 32 recorded canal dike breaches (Van Etten, 2006). Of these, 12 occurred at peat dikes, while another 13 breaches occurred at dikes where the type of subsurface material was not identified, as shown in Table 1.1. The majority of these breaches were likely caused by high water levels, performing maintenance activities causing failures and droughts. Over the last decades, the number of canal dike failures has declined. Only a few breaches have been recorded in the last 50 years, namely at Aarlanderveen in 1995, Wilnis and Terbregge in 2003, and Reeuwijk in 2021. In all cases, these breaches happened when there were no high water levels or heavy rainfall. This decrease in the number of breaches over time can likely be attributed to regular and systematic safety assessments and reinforcement efforts, as well as advancements in pumping technology that enhance the regulation of water levels in drainage canals.

It should be noted that in Table 1.1 only possible causes were identified, and no clear distinction was made between different failure mechanisms. For instance, it remains uncertain whether high water levels led to overtopping and overflow, or whether they caused, in combination with infiltration of heavy rainfall, increased pore-water pressures that triggered slope instability. Distinguishing between these mechanisms in the field is challenging, especially since failures like slope instability mainly occur below the surface. Additionally, these historical failures cannot be directly compared to current conditions, as drainage and water management systems have improved, and water levels are now regulated more than in the past.

Table 1.1 Subsurface material and possible causes of canal dike breaches from 1900 until 2005 (Van Etten, 2006)

	Possible causes of dike breaches								
Subsurface material	Unknown	High water level	Maintenance	Strange elements	Drought	Seepage	Bombing	Animal burrow	**Total**
Unknown	3	5	3	1			1		**13**
Peat	2	3	2	1	2	1	1		**12**
Clay	3	1			1				**5**
Sand								2	**2**
Total	**8**	**9**	**5**	**2**	**3**	**1**	**2**	**2**	

1.3 Threats from above

Dikes can become unstable in various ways, each with different underlying failure mechanisms. One of the most important failure mechanisms identified in canal dike assessments is inner slope instability, due to too wet conditions (Lendering et al., 2018; Van der Meer et al., 2009). This occurs when the driving forces on the dike exceed the resisting forces, causing sliding along a shear surface within the dike or its foundation, leading to a dike breach and flooding of a polder. The driving forces typically include high phreatic head levels within the dike body combined with a horizontal water pressure exerted from the canal and traffic loads on top of the dike. High head levels can increase the driving moment and decrease the effective stresses and subsequently decrease the soil shear strength (Frank et al., 2004; Sharp et al., 2013). These high head levels result from water infiltrating the dike body due to high water levels and/or heavy rainfall.

In addition to failures caused by water infiltration and high head levels, excessively dry conditions can also induce instabilities. In general, droughts can affect the stability of dikes in multiple ways, both positive and negative. Drought can decrease the driving moment and increase the effective stresses, as pore-water pressures are lower. This improves the dike stability. In contrast, it can also result in weight loss and soil shrinkage, which can cause the horizontal translation of a dike body. This mechanism is likely to happen primarily at peat dikes, as it did at two sites in the Netherlands during the summer of 2003 following an extended dry period, namely near Wilnis and Ter Bregge (Bezuijen et al., 2005; Van Baars, 2005). Additionally, the presence of cracks in canal dikes, caused by for example swell and shrinkage, can reduce the shear resistance and may lead to the formation of shallow slip planes (Zhang et al., 2021). At the same

time, these cracks can enhance the infiltration of water when heavy precipitation events occur, potentially leading to higher hydraulic head levels and localised saturation. However, cracks can also facilitate drainage, reducing hydraulic head buildup over time.

Next to the impact of extreme conditions, cyclic wetting and drying also affect the dikes' resistance. For example, unsaturated soils may lose shear strength due to swelling and shrinkage (Stirling et al., 2021), while also the hydraulic conductivity and water retention capacity can change over time (Briggs et al., 2023; Azizi et al., 2020). In contrast, the dikes' resistance can also improve over time because of the effects of ageing, depending on the soil type (Schmertmann, 1991). Soils can undergo restructuring and consolidation, which can increase their shear strength over time (Terzaghi et al., 1996; Ladd & Foott, 1974). Depending on the subsoil and site-specific conditions, extreme head levels and dry-wet cycling can improve or worsen dike stability.

In conclusion, the subsurface water conditions affect the dike's stability in multiple ways, both during extreme high-water and drought situations, and through cyclic wetting and drying, as illustrated in Figure 1.4. Therefore, the geohydrological response of a dike, the way and to what extent hydraulic heads and soil moisture levels vary over time, plays a key role in its stability and safety.

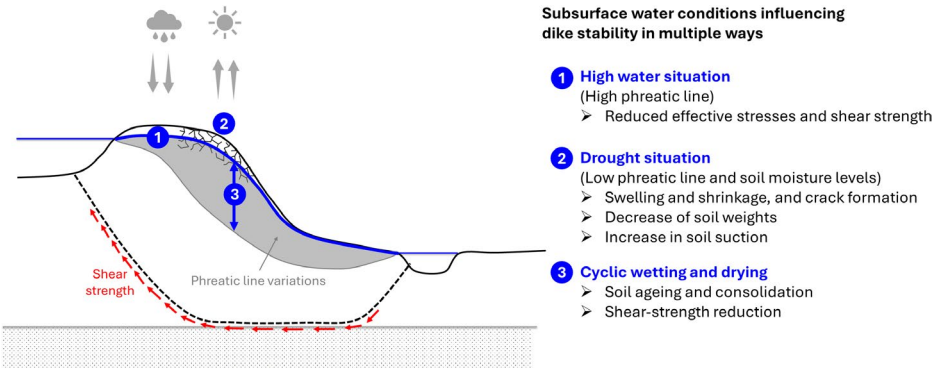


Figure 1.4 A conceptual figure illustrating how subsurface water conditions in canal dikes can affect the dike's stability.

The geohydrological response will differ for every dike, depending on, for example, the outside water level fluctuations, climate regime (precipitation and evaporation characteristics), geometry (slope and width) and subsurface material. In general, it is driven by fluctuations in the outside water levels, precipitation and evaporation. For canal dikes, precipitation typically plays a more significant role compared to the water levels (Rikkert, 2022; Jamalinia et al., 2020; Lendering et al., 2018). This is because

the water levels are regulated with only minor fluctuations, typically within a range of decimeters. When water levels do rise, they remain high for just a few days, and only a small amount of water infiltrates into the dike, because clay and peat have relatively low hydraulic conductivities. In contrast, the hydraulic head levels fluctuate from decimeters up to meters, driven by precipitation and evaporation. This is illustrated by head and water level observations at two canal dikes in Figure 1.5. The head levels are generally highest during the winter (November until February) and lowest at the end of the summer (August/September). During the summer months, drying takes place as water drains mainly toward the polder ditch, driven by the horizontal hydraulic gradient, while water can also be extracted by the vegetation through transpiration. The two dikes illustrated in Figure 1.5 show already a very different geohydrological response. The Hennipslootkade is more responsive to rainfall events, with sharper peaks in its course, and it begins drying out earlier than the dike at the MT-polder. Understanding the geohydrological response for various canal dikes and the impact on dike safety and flood risk in polders is the main interest of this dissertation.

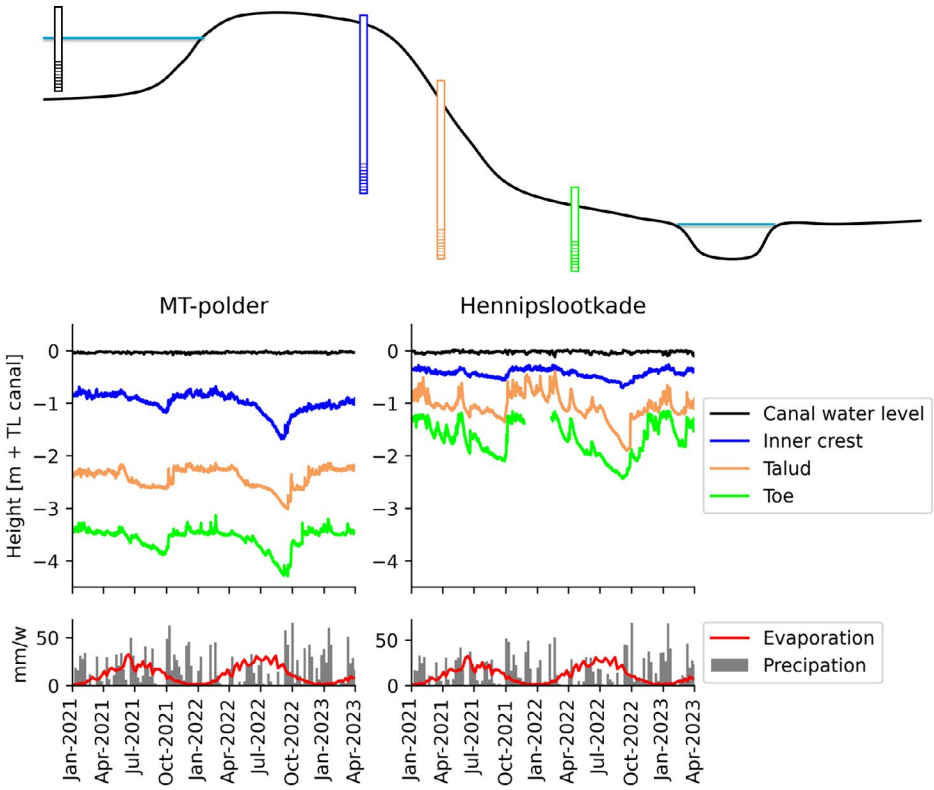


Figure 1.5 Measured water levels and hydraulic heads at different locations within the dike profile of two canal dikes, as also illustrated in the dike profile, as well as the meteorological conditions (precipitation and evaporation).

1.4 Quantifying flood risk in polders

1.4.1 Quantitative probability and the role of observations

To understand and manage the threat of flooding, flood risk assessment is an often-used framework (UNISDR, 2011; Jonkman et al., 2008; Vrijling, 2001). This comprises quantifying the probability of flooding and its consequences. In polders, the flood probability strongly depends on the probability of failure of any dike section within the surrounding dike system. Estimating the failure probability of a dike requires quantifying failure processes, which not only enables us to assess dike safety or to predict weak spots, but also informs proper design and enables appropriate measures. Failure can occur through various mechanisms, with failure mechanisms often described by a limit state function Z, which can be assessed by comparing the resistance R and the load S:

$$Z = R - S \tag{1.1}$$

In probabilistic safety assessments, engineers consider the probability that the limit state function is exceeded (Z<0) or that the load exceeds the resistance (*R*<*S*), which is known as the probability of failure:

$$P_f = P(Z < 0) = \int_{S=-\infty}^{S=\infty} f_S(s)\, F_R(s)\, ds \tag{1.2}$$

In which $f_S(s)$ is the probability density function of load levels and $F_R(s)$ is the cumulative density function that gives the probability of failure conditional on the load *S*, also known as the fragility curve. These distributions can be described by a set of variables. For example, the load can comprise multiple variables, such as water level and waves, and the resistance can depend on various soil layers described using multiple soil parameters. These variables are often schematized or quantified using data. This data may include model outputs, such as hydraulic load models like Hydra-NL that estimate water level exceedance probabilities at specific sites in the Netherlands (Geerse, 2011), or observational data, including the dike's geometric profile and material properties. However, the underlying models often rely on observational data. Depending on how uncertain a variable is and how strongly it influences the limit state function, it can be treated as either stochastic or deterministic. Based on the schematization of load and strength in the limit state function, the dike stability can be calculated using both fully probabilistic and semi-probabilistic calculations. Both calculations rely on the same limit state functions and the same probability distributions of the stochastic variables. The only difference is that a semi-probabilistic approach relies on several simplifications and approximations, which give it the appearance of a deterministic method (Steenbergen et al., 2004; Jongejan et al., 2020). It applies characteristic values and partial factors chosen to match a target safety level, and the result is expressed as a factor of safety. This result is indirectly linked to the failure probability. The partial

factors are calibrated through fully probabilistic studies, ensuring that when the required factor of safety is met, the corresponding probability of failure meets the target level. In contrast, the outcome of a fully probabilistic calculation is a failure probability directly. The process from observational data to failure probability estimates is illustrated in Figure 1.6. Once the failure mechanism's limit state function is identified, the variables within it need to be quantified (step 1). By using observational data in combination with, among others, models and statistical analyses, a schematization of load and strength variables in the limit state function can be obtained. Based on this schematization, the failure probability can be calculated using probabilistic methods, such as numerical integration or Monte Carlo simulations (step 2).

Traditionally, engineers deal with uncertainties in failure probability analyses by starting with initial estimates based on available information and refining these estimates through additional field observations. Therefore, the outcome and model schematization should be reviewed, possibly by collecting more observational data to refine the schematization. Since the uncertainty of stochastic variables relates to the outcomes of potential observations, it can be reduced or eliminated through the collection of more observations, depending on the type of uncertainty (Bedford & Cooke, 2001).

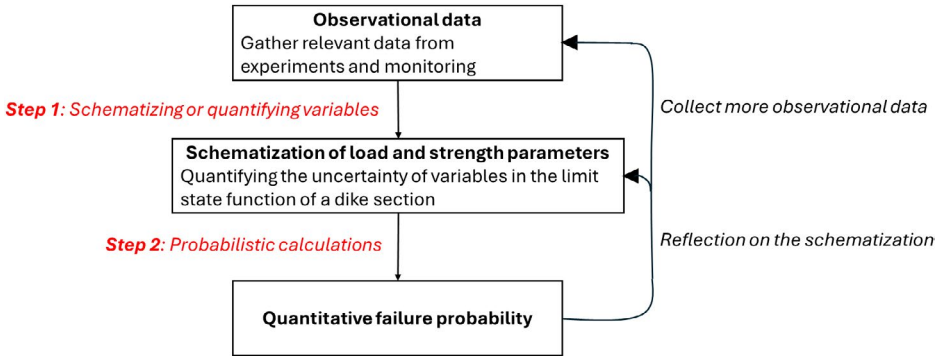


Figure 1.6 The process from observational data to failure probability estimates involves, in general, two steps: 1) schematizing or quantifying the load and strength parameters in the limit state function, and 2) calculating the failure probability using probabilistic methods. Observational data play a key role and form the basis of failure probability estimates.

1.4.2 Types of uncertainties

Understanding and quantifying uncertainties is fundamental in probabilistic dike safety assessments, as every estimate of failure probability depends on how these uncertainties are schematized. Some uncertainties can be reduced or eliminated, while others are nearly impossible to reduce. The various sources and types of uncertainty are illustrated here using the example of the sea dike protecting the city of Rotterdam.

The sea dike protecting the city of Rotterdam could fail on the 30th of September 2026, but how likely is that to happen? Estimating this probability involves dealing with numerous uncertainties. What is the probability that a certain water level will be exceeded on that day? What is the dike's resistance to withstand that high water level on that day? How do the water level and resistance vary along the sea dike? The water levels at Rotterdam naturally and randomly vary over time and are difficult to predict more than one month in advance. In contrast, the resistance is more static and does not vary randomly over time. It can change over time due to changes in soil properties, vegetation or animal burrowing, but this is not a random process. The resistance can be seen as an epistemological or knowledge uncertainty, while the water level is inherently uncertain. This distinction between types of uncertainty is often made in scientific literature (Der Kiureghian & Ditlevsen, 2009; Vrouwenvelder, 2003; van Gelder, 2000; Walker et al., 2003). Although the inherent uncertainties of water levels and waves cause floods in the outside world and are the primary reason why flood risk analyses are conducted, knowledge uncertainties can also play a significant role in dike safety. For example, if we are very uncertain about the resistance of a dike, the failure probability can be primarily attributed to knowledge uncertainties rather than the inherent uncertainties. This can result in uncredible results, as illustrated below with an example of a hypothetical dike.

The failure probability of a hypothetical dike can be calculated using a so-called fragility curve, which represents the uncertainty of the dike's resistance. The curve describes the failure probability conditional on a load variable, for example, the water level. That water level is also uncertain and has a probability distribution. Multiplying this fragility curve and the probability distribution of the water level produces the failure probability distributions, after which the dike's overall failure probability can be calculated by integrating that distribution, as described in equation 2. Suppose the theoretical case that everything is known about the dike's resistance and its failure process, and the dike fails if the water level exceeds NAP+5m. In this case, the fragility curve approaches a step function, as shown in Figure 1.7, and the failure probability is equal to the exceedance probability of a water level of NAP+5m, which is in this hypothetical case 1/14,900 per year. If we are less certain about the dike's resistance, say we expect the dike to fail at NAP+5m, but there is also a 10% probability of failure

at NAP+4.6m and a 10% probability that it withstands a water level of NAP+5.4m. The fragility curve becomes more S-shaped. Although the uncertainty is symmetric, the failure probability increases because there are more scenarios in which failure can occur. Accounting for this knowledge uncertainty results in a failure probability of 1/200 per year, which is more than a factor of 10 higher than without knowledge uncertainties. Often, any uncertainty added to a failure probability analysis increases the failure probability. Therefore, to produce credible failure probability estimates, one of the major challenges is to estimate the resistance properly, which mainly consists of epistemologically uncertain variables.

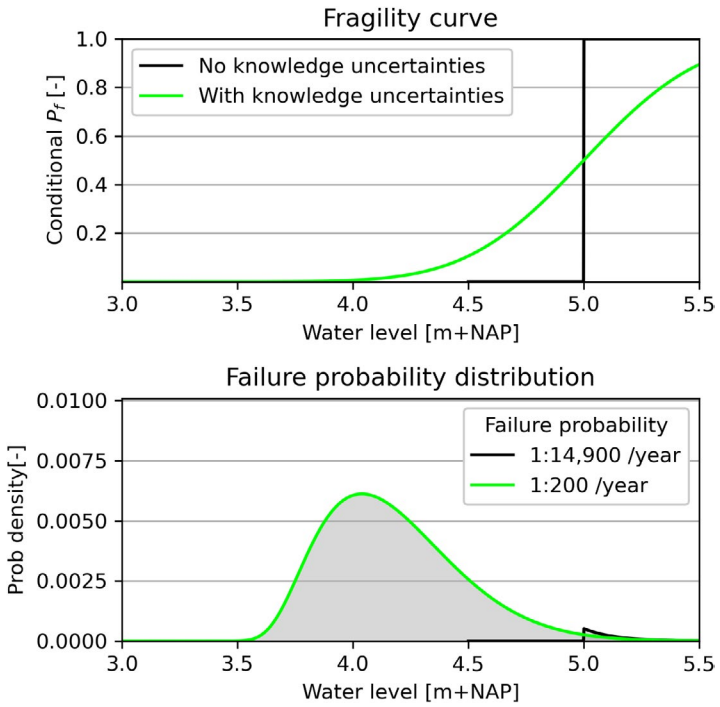


Figure 1.7 Illustrative example of two fragility curves and their corresponding failure probability distributions: one with knowledge uncertainties and the other without knowledge uncertainties.

1.4.3 Dike systems and the length-effect

For effective flood risk management in polders, the reliability of the surrounding canal dike system is essential. The probability of flooding in the polder equals the probability that a dike fails somewhere within the dike system. This can be caused by several mechanisms, and this study focuses on one particular failure mechanism, namely, the failure of inner-slope instability, further referred to as stability.

To estimate the flooding probability of a dike ring or polder, the dike system is often regarded as a series system of individual elements, also known as dike sections, see Figure 1.8. Sections are defined as stretches in which load and resistance properties are statistically homogeneous. The failure probability of a dike section is estimated by calculating the failure probability at the cross-sectional level using two-dimensional reliability analyses. This is a common engineering practice, as 3D analyses are complex and the available spatial data are often limited. To obtain a failure probability of the section, the length effect of the dike section is accounted for. The length effect refers to the phenomenon that the longer the dike, the higher the probability that it will have a weak spot and fail somewhere. Speculations about the length effect can be traced back to Leonardo da Vinci (1500's), who observed that "among cords of equal thickness the longest is the least strong," and proposed that "a cord is so much stronger … as it is shorter," (Williams, 1957). For dikes, even at the section scale, soil properties exhibit spatial fluctuations, resulting in spatial variabilities within such a statistically homogeneous section (TAW, 1989; Vrouwenvelder & Calle, 2003). This results in failure probabilities of dike sections that can be several times larger than those of cross-sections, depending on the length of the dike section, fraction of the total length of the dike section that is sensitive to the failure mechanism and the spatial variability of the strength parameters (Jongejan et al., 2020).

After calculating the failure probabilities of all dike sections, the probability of any dike failure along a polder $P(F_{polder})$ can be estimated. This probability will lie between the maximum of the individual section failure probabilities and approximately their sum,

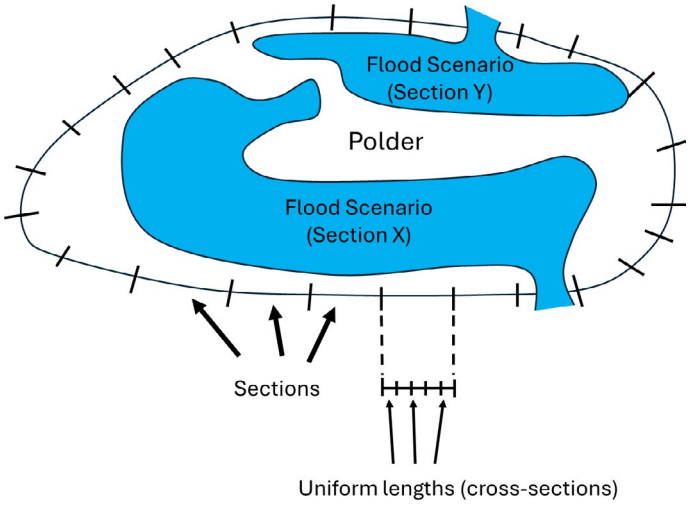


Figure 1.8 A polder with a along it a dike ring divided into sections with cross-sections (Adapted from Jongejan et al., 2020)

$max(P_i) \leq P(F_{seg}) \leq \Sigma P_i$, depending on the spatial correlations. These spatial correlations do not only arise from spatial variations in the dike's resistance, but also from spatial variations in the load. For example, spatial variations in water levels or high phreatic lines, caused by different head responses in each section, can make heavy-rainfall events of varying durations relevant to the dike's stability. Therefore, an extreme load event that occurs once every 100 years on the section level can occur once every 25 years on the polder level when there are four independent dike sections $P(F_{seg}) = \sum_{n=1}^{4} 1/100 = 1/25$. If one dike section stands out as the weakest link, its failure probability dominates the failure probability of the polder and the spatial dependencies between sections do not matter much. The spatial dependencies become important for the failure probability of the polder when there are several sections with similar failure probabilities, since only then does the outcome of $max(P_i)$ and ΣP_i can differ substantially.

1.4.4 Dike stability and the role of the geohydrology

Evaluating inner-slope instability of dikes requires appropriate assessments of the subsurface materials and their soil properties, such as unit weight and shear strength, and pore water pressures. Since these variables can be time-variant and spatially uncertain, this requires an analysis of the uncertainties of soil parameters that determine the strength of dikes, as well as the probability that certain pore water pressures occur. Next to pore-water pressures, the dike's load can also include traffic loads and high water levels. Common practice is to calculate the probability of instability, conditional on a certain load scenario, using two-dimensional analytical Limit Equilibrium Methods (LEM), such as Bishop, Spencer or Uplift-Van (Bishop, 1955; Sharp et al., 2013). There are also other more advanced numerical methods to analyse slope stability, like Finite Element Methods (FEM) or Finite Difference Method (FDM). These numerical methods can capture the behaviour of soils better, as they are capable of modelling, for example, deformations, stress paths, and transient effects. However, these methods are computationally expensive and demand considerably more model parameters and input data than LEM. Therefore, the LEM becomes more attractive for uncertainty analyses. This method is implemented in software like D-stability (Van der Meij, 2023), which can calculate the probability of instability of a dike, considering many possible slip planes, soil properties and strength uncertainties of a dike, as shown in Figure 1.9. By calculating the probability of instability under various load scenarios, a fragility curve can be constructed. Combining the fragility curves and the probability of load scenarios gives the overall probability of instability, see section 1.4. As with every model, the quality of the input to the model determines the quality of the output. Therefore, the schematization of subsurface materials and pore-water pressures is of great importance, where this thesis focuses on the pore-water pressures in dikes and the geohydrological response.

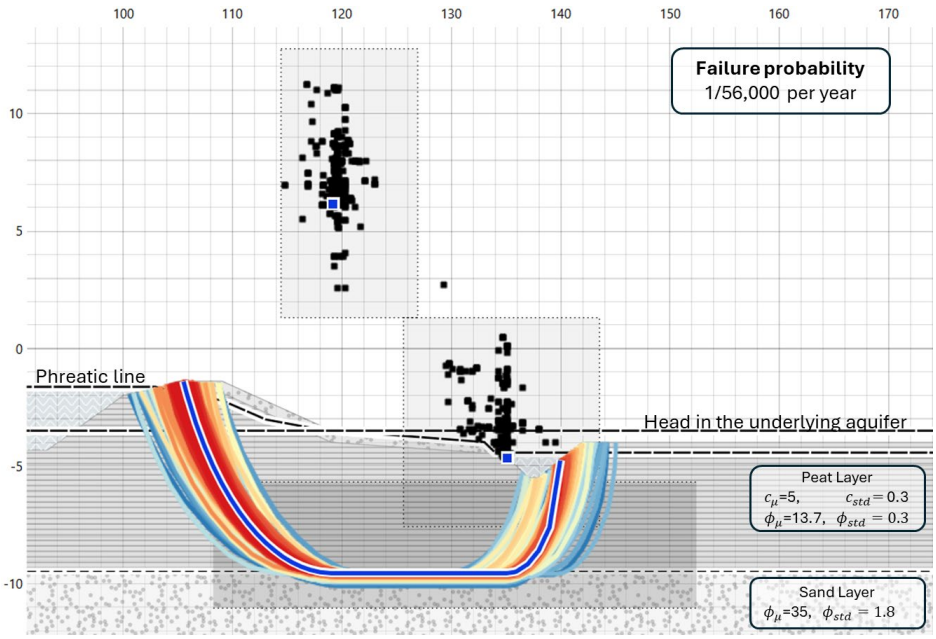


Figure 1.9 An example of a slope stability calculation with the LEM as implemented in D-Stability, where the software considers multiple slip planes where the bright blue indicates the normative slip plane. The failure probability is calculated conditional on one load scenario, in which the soil strength of various layers is described using Mohr-Coulomb and considered stochastically.

Schematizing the spatially and temporally varying pore-water pressure that governs dike stability is challenging. It can be captured by two-dimensional numerical groundwater models, but in practice, data on key input parameters, such as saturated hydraulic conductivity and water retention curves, are often not available. Additionally, long-term pore-water pressure monitoring points are often not available to adequately calibrate these models. Therefore, in practice, the pore-water pressure fields are simplified using headlines. These lines mark the pore-water pressures at designated soil layers, such as the phreatic line or the head in the underlying aquifer. For canal dikes, the phreatic line proved to be a dominant variable for the probability of instability (Lendering et al., 2018; Rikkert et al., 2022). Although this is an important parameter, the way it is schematized in failure probability analyses is often based on simplified prescribed guidelines, while it is known that the head response can vary strongly from dike to dike. These prescribed guidelines are often seen as a first conservative estimate of the phreatic line and can be sufficient if the dike appears to be stable enough. However, assessing a credible probability of instability requires local head measurements using piezometers. When measurements are available, challenges remain in how they are used to schematize the phreatic line, as will be shown below.

Each dike has a different geohydrological response, so its phreatic line behaves differently. The extreme high and low positions of this line are determined by the meteorological events relevant for that location and their characteristic duration. For example, a high phreatic line may result from a single day of heavy rainfall in one case, or from a prolonged eight-day rainfall in another. In general, estimating the probability of the phreatic line in earthen structures during extreme precipitation events is commonly estimated using groundwater flow models in combination with Intensity-Duration–Frequency (IDF) curves for precipitation (Cheng & AghaKouchak, 2014). These curves represent the intensity and frequency of precipitation for various durations, which serve as boundary conditions for the groundwater flow models. Jasim et al., (2017) analysed the performance of a dike in California under extreme precipitation scenarios using numerical model simulations and found that especially rainfall events with a duration of more than three days can have a significant effect on the stability. Robinson et al., (2016) analysed the stability of a slope under current and futuristic climate scenarios, using a rainfall duration of 7 days. In general, both extremely short- and long-term durations can be relevant, depending on the geohydrological response of dikes. Currently, there is limited understanding of the variation in geohydrological response along canal dikes and the associated rainfall durations.

1.4.5 Potential impact of climate change on earth structures

Climate change has increased temperatures and changed precipitation patterns, resulting in changes in the magnitude and frequency of extreme events (Trenberth, 2011). Warmer temperatures enhance evaporation, which can increase the probability of more prolonged periods of drought (IPCC, 2011). On the other hand, precipitation is increasing because warmer air can hold more moisture, which can increase wetter winters and more intense showers. Because of uncertainties caused by incomplete scientific knowledge about how the physical climate system responds to changes in atmospheric composition and land use, climate change is often studied using scenarios. In the Netherlands, there are so-called KNMI'23 climate scenarios (Van der Wiel et al., 2024). These scenarios provide plausible and physically consistent storylines or pathways of climate change in the Netherlands, considering uncertainty for regional climate change and global temperature changes driving it. The projections show that for all scenarios, temperatures in all seasons and winter precipitation will increase, while summer precipitation will decrease. Associated with these changes are increases in extreme precipitation during the winter and increases in meteorological droughts. The projected change in future climate contributes to the stability of canal dikes in at least the following ways (Briggs et al.,2023; Stirling et al., 2021; Azizi et al., 2020; Vardon, 2015):

1) The resistance of dikes can be altered, due to changing soil strength properties by:

   a. The occurrence of stress levels that exceed historic values, caused by changes in the magnitude and frequency of extreme pore-water pressures, both saturated and unsaturated, driven by increasing precipitation and droughts.
   b. The amplification of dry-wet cycles that can induce weakening processes, like soil-strength reduction and desiccation cracking, driven by a more intensified water cycle with more intense precipitation and droughts.
2) The loading conditions can be altered, due to changes in the magnitude or frequency of extreme pore-water pressures in dikes, by:
   a. Changes in extreme precipitation events and droughts
   b. Changes in soil properties, like hydraulic conductivity and soil suction generation capacity, caused by micro-structural changes to the soil fabric and more extensive desiccation cracking.

Some of these effects on earth structures are already demonstrated, for example, for railway embankments in the London area (Briggs et al., 2023; Huang et al., 2024). Seepage analyses were conducted, and the influence of climate change on pore-water pressures within earthwork slopes was quantified. The analyses showed that forecasted hotter, drier summers will increase the water storage capacity of earthworks, leading to a longer time being required to saturate the soil pores and bring the water table back to the slope surface. The effect of more extreme precipitation on dikes was studied with a fully coupled 2D hydromechanical transient variably saturated flow model, which showed a 3–12% increase in failure probability for an earthen levee in California, USA (Jasim et al., 2017). The effects of these processes on canal dikes, with different characteristics and boundary conditions, remain insufficiently understood.

1.5 Knowledge gaps tackled in this research

Quantifying current and future flood risk in polders and estimating the failure probabilities of canal dikes requires knowledge of physical processes and accurate estimates of extreme conditions. This research contributes to the understanding of these processes by tackling the following knowledge gaps. These were identified by reviewing existing studies and talking and discussing with researchers, experts and practitioners in dike safety:

- Multi-year measurements of hydraulic head levels in combination with soil moisture in dikes, especially canal dikes, are often lacking, which makes groundwater modelling exercises difficult to validate. Therefore, **the geohydrological response of canal dikes to heavy precipitation and drought periods remains insufficiently understood**, while these factors have been proven to be important for the stability of canal dikes (Lendering et al., 2018; Rikkert et al., 2022; Van Baars, 2005). Furthermore, proxy variables, such as meteorological indicators, are used as

indicators of subsurface water conditions and inform decision-making regarding inspections (Özer et al., 2019b). However, little is known about the performance and reliability of these proxies in the case of canal dikes. (*Chapter 2*)

- Lack of measurements and the extensive computation time required for groundwater models make it challenging to quantify loading conditions and assess dike safety. There is **limited understanding regarding the variability in peak head responses and head statistics across canal dikes, including the effects of climate change,** as discussed in Chapter 3. Climate change is expected to cause heavy precipitation to increase in both magnitude and frequency, while dry summers are also likely to become more intense and prolonged. What is the variation in head statistics? What is the impact of more extreme weather events on head statistics, and how does this vary from dike to dike? (*Chapter 3*)
- The variability in head responses across canal dikes, together with the natural spatial variability of weather events, determines the spatial dependencies in peak heads along dike sections. These spatial dependencies are relevant for estimating flood probabilities in polders, as well as at national levels, as discussed in Section 1.4. Additionally, while failure observations on the section level can be rare, at the system level, more failures are expected, given the presence of more independent dike sections. By understanding spatial dependencies, failure observations in the entire dike system, or the lack of it, can be used to verify failure probability estimates at the section levels. However, at this moment, **little is known about the spatial dependencies of peak heads in canal dikes and how they affect the occurrence of extreme events across different spatial scales**. How many independent dike sections are there on the polder level and in the entire dike system? This depends on spatial dependencies in both load and strength, where this thesis emphasises the spatial variability of peak heads in the canal dikes. (*Chapter 4*)
- Over a thousand kilometres of regional and primary flood defences are considered unsafe in the Netherlands (Nationaal Georegister, 2024). Many of the Dutch dikes are centuries old and have often been reinforced and upgraded over the years, resulting in heterogeneous dikes with varying subsurface materials (van Woerkom et al., 2022). This heterogeneity, combined with limited data on the subsurface buildup and the uncertain shear strength of the soil, makes the overall strength of dikes uncertain. Therefore, it is unknown whether the dikes are truly unsafe or whether it is challenging to demonstrate adequate safety with the available data and models. **When do failure probability estimates become uncredible, meaning that probabilities are unrealistic and unjustifiable, and contradict the (historical) behaviour of the dike? What is the role of time-varying inherent load uncertainties, meaning natural fluctuations in loads such as water levels and hydraulic head levels, in assessing credible failure probabilities?** (*Chapter 5*)

1.6 Aim and research questions

This thesis aims to

> Enhance geohydrological understanding of canal dikes and improve the reliability assessment of their stability and the dike system through observation-based analysis.

This aim is pursued with the following research questions that form the core of this thesis:

1. What is the geohydrological response of canal dikes with a focus on drought periods, and what are effective meteorological indicators that reflect this response? (Chapter 2)
2. What is the variability in peak head responses and peak head statistics across Dutch canal dikes and the impact of climate change, and can these variations be explained by dike characteristics? (Chapter 3)
3. How much more frequently do extreme hydraulic head peaks occur at larger spatial scales than at the section level, for the Dutch canal dikes? (Chapter 4)
4. What is the role of load variations in assessing credible failure probabilities? (Chapter 5)

1.7 Outline and methodology

The way the research questions and chapters in this thesis are related to each other is illustrated in Figure 1.10. After the introduction chapter, there are four chapters containing the main content of the thesis, where each chapter elaborates on one research question. To answer the research questions, two unique datasets are used that are both publicly available at 4TU.ResearchData, which is an international data repository for science, engineering and design:

1. A dataset that contains detailed subsurface water measurements at ten canal dikes with different characteristics, such as clay and peat dikes with various crest widths and slope angles, in the western Netherlands. This network has been operational since the spring of 2020, and the measurements are ongoing in 2026. These measurements include soil moisture measurements at various depths and hydraulic head levels, and is used in Chapters 2 and 5 of this thesis.
   https://doi.org/10.4121/136aa5df-1907-43ac-a5b0-e0ea8f2dedf3
2. A dataset that contains head observations of piezometers in Dutch canal dikes covering 108 head time series at 48 different monitoring sites/canal dikes. Multiple head observations can be located at one monitoring site, where piezometers are aligned within a dike cross-section, for example, at the crest, inner slope, and inner toe of the dike. This dataset is used in Chapter 3.
   https://doi.org/10.4121/4004f445-b71b-4996-bd7d-10b1fafbc86b

These datasets are used in combination with various methods to reveal patterns in the data and explain why those patterns occur. By capturing the underlying cause-

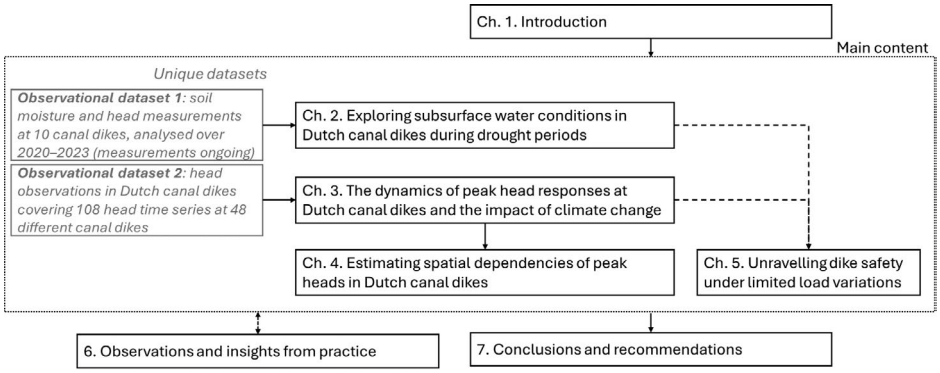


Figure 1.10 Graphical representation of the thesis's structure.

and-effect relationships, we can also estimate how dikes may behave in new or future situations. This thesis uses various methods to answer the research questions, including data-analysis techniques, time series modelling and probabilistic modelling.

Chapter 2 explores the subsurface water conditions in Dutch canal dikes during droughts by analysing three years of head level and soil moisture data to characterise geohydrological behaviour. Next, Chapter 3 applies time-series models on a larger dataset of head observations to understand the head response and quantify how higher winter precipitation and summer evaporation, influenced by climate change, affect extreme high head levels. Chapter 4 takes the response variations identified in Chapter 3 and uses them to evaluate the spatial dependencies in peak heads at various spatial levels, such as polder level and national level. Chapter 5 investigates how load variations affect failure probability analyses using fragility curves and stability analyses, where phreatic lines in dikes are based on the observations and time series models developed in the earlier chapters.

After the main content of this thesis, Chapter 6 describes some observations and insights from practice. During my research activities, I actively applied scientific insights to practical projects for water authorities, and in turn, real-world challenges inspired and refined my research. These analyses are less in-depth, however, they can offer valuable insights for dike safety and could not have been possible without the scientific insights gained from this thesis. Finally, Chapter 7 concludes the thesis by summarising the key findings and offering recommendations.

This research, as well as installing and maintaining the monitoring sites, was supported by the Stichting Toegepast Onderzoek Waterbeheer (STOWA), Rijkswaterstaat and the regional water authorities Rijnland, Schieland en de Krimpenerwaard and Delfland.

2

Exploring subsurface water conditions in Dutch canal dikes during drought periods: insights from multiyear monitoring

This chapter analysed a novel multiyear data set of soil moisture and hydraulic heads (from February 2020 until March 2023) from a monitoring network covering various canal dikes with different characteristics in the western Netherlands. The data, including two extremely dry summers, highlight the impact of meteorological variations on the subsurface water conditions. Non-hydrostatic hydraulic head levels were observed during droughts that can be detrimental to dike stability and that are often not accounted for in safety assessments for drought situations. The effectiveness of various meteorological drought indicators applied to subsurface water conditions was evaluated: the precipitation deficit is the most reliable measure and outperforms the standardized drought indicators (*SPEI* and *SPI*). The drought recovery of dikes was analyzed to understand seasonal transitions and the sequence of different failure mechanisms, during dry and wet situations. This analysis also reveals differences between meteorological, soil moisture, and groundwater droughts, highlighting soil's storage capacity after drought and the limitations of meteorological drought indicators as proxies for soil moisture and groundwater. The insights from this study enhance assessments, inspection procedures and the identification of weak spots of dikes and other earthworks of infrastructure.

This chapter is based on Strijker, B., Heimovaara, T. J., Jonkman, S. N., & Kok, M. (2024). Exploring subsurface water conditions in Dutch canal dikes during drought periods: Insights from multiyear monitoring. Water Resources Research, 60*(9), e2023WR036046. https://doi.org/10.1029/2023WR036046*

2.1 Introduction

Flooding is among the leading climatic threats to people's livelihoods, affecting development prospects worldwide (Jevrejeva et al., 2018). Extensive systems of flood defenses protect flood prone areas worldwide (O'Dell et al., 2021) and the top three countries with the highest relative population exposed to flood risk have to a certain degree flood protection systems (Rentschler et al., 2022). To manage flood risks in an embanked area, the performance of dikes plays a crucial role and several failure mechanisms can induce dike breaching and flooding of the hinterland (Özer, van Damme, & Jonkman, 2019). Key inputs to the assessment of dike performances are the loading conditions and the resistance of dikes. Important factors contributing to the loading conditions are subsurface water conditions, such as hydraulic head levels and soil moisture content, which can trigger failure (Sharp et al., 2013). At the same time, changes in subsurface water conditions also induce weakening processes that reduce the resistance, such as soil-strength reduction, and desiccation cracking (Robinson & Vahedifard, 2016; Stirling et al., 2021; Vahedifard et al., 2016; Vardon, 2015). Therefore, understanding subsurface water conditions in dikes is essential for flood protection in embanked areas.

A special case of an embanked area is a polder, which is a low-lying area enclosed by embankments with an internal drainage system. Polders are present in many parts of the world, like the Netherlands, Bangladesh, Vietnam and China (Martín-Antón et al., 2016; Morton & Olson, 2018; Lendering et al., 2018; Tran & Weger, 2018; Triet et al., 2017; Manh et al., 2013; Warner et al., 2018). Drainage canals in polders drain excess water from the polder to the main water bodies (or vice versa) to mitigate the flood risk and facilitate irrigation. Water levels in the canal, regulated using weirs, sluices and pumping stations, can lie several meters above the surrounding polder area. Dikes along the canals, known as canal dikes, are vital infrastructure in low-lying polders and their performance can either prevent or trigger internal flooding if a breach occurs. The Netherlands has more than 10,000 km of canal dikes (Pleijster et al., 2015), which mainly consist of soft soils (clay and peat) covered with grasses (Lendering, 2018). Important driving forces of the changing subsurface water conditions in these dikes are the meteorological conditions since the outside water levels are almost constant, in contrast to many other types of dikes where the outside water levels are leading (Rikkert, 2022; Van Baars & Van Kempen, 2009). Both droughts and wet situations can trigger instabilities, as well as the cascading impacts of droughts. In the past, many Dutch canal dike failures occurred due to inner-slope instability (Van Baars & Van Kempen, 2009), like the failures during summers near Wilnis and Terbregge in 2003 and Reeuwijk in 2021. However, this is not limited to the Netherlands and several historic dike failures worldwide highlight the impact of drought conditions on dike integrity, such as failures along the Murray Riverbank in Australia from 2008 to 2010 and at Edenburry in Ireland in 1989 (Bezuijen

et al., 2005; Hubble et al., 2014; Pigott et al., 1992; Van Baars, 2005). These events underscore the vulnerability of dikes under extended dry periods worldwide.

Soil-atmosphere interaction can significantly impact the stability of soil structures, making it important for safety assessments (Elia et al., 2017; Vardon, 2015). However, modeling soil-atmosphere processes can be challenging and complex. Several attempts have been made to model the effects of rainfall and evaporation on the phreatic surface in dikes, the level within the dike where the soil is fully saturated with water, using a range of models from simple conceptual ones to coupled agro-meteorological and 2D numerical groundwater flow models (Jamalinia et al., 2019; Rikkert, 2022; Van Esch, 2012). Multi-year measurements of hydraulic head levels in combination with soil moisture in dikes, especially canal dikes, are often lacking, which makes modeling exercises difficult to validate. Therefore, knowledge about the geohydrological behavior and the importance of the soil-atmosphere interaction of canal dikes is currently limited. This makes it challenging to quantify loading conditions and assess flood risk levels properly, not to mention accounting for processes that weaken the dike. Furthermore, Dutch canal dikes that are prone to droughts, whose subsoils consist of peat or organic clay, are inspected during dry summers to detect weak spots and take measures to prevent failures. In practice, water managers have to decide when to start inspections. Various meteorological drought indicators are used as an indicator of dike safety, because measurements within dikes are not widely available. The reliability of meteorological indicators for accurately representing the actual drought conditions in canal dikes (such as soil moisture and hydraulic head levels) is unknown, making it difficult to justify decisions made by water managers. Understanding the geohydrological response of various canal dikes leads to more accurate safety and reliability assessments of canal dikes and flood defenses in general. Although the focus of this dissertation is more on canal dikes throughout the Netherlands, the geohydrological response of these dikes can still be very relevant for similar dikes in low-lying polders and for dikes worldwide. The continuous exposure of canal dikes to high water levels and rainfall, and evaporation makes them valuable sources of potential insights applicable to any dike.

2.1.1 Objective and outline

This research provides insights into the geohydrological response of dikes by analyzing multiyear monitoring data with a focus on droughts, based on measurement data from 10 monitoring sites with different characteristics. The applicability of various meteorological indicators, to provide information about the development of drought in dikes, is investigated by comparing these with the observed groundwater and soil moisture levels. Lastly, the recovery from droughts is analyzed, which characterizes the transition between different failure mechanisms, during dry and wet situations.

This chapter starts by providing relevant background on drought failure and drought-induced weakening processes and discussing the Dutch situation. Then, the monitoring sites are described, including the dike characteristics and monitoring equipment, after which different concepts of droughts and several drought indicators are introduced. Next, the data analyses of the geohydrological response of dikes are presented. In the discussion, the limitations and relevance of these insights for practical application are discussed, followed by concluding remarks and recommendations.

2.2 Background and methodology

2.2.1 Dike failure and drought induced weakening

Catastrophic dike failures occurred throughout history with various causes, like storm surge, ice drift and rainfall or drought, and failure mechanisms, such as overflow and overtopping causing soil erosion, external erosion, piping and inner slope instability (Özer, van Damme, & Jonkman, 2019; Van Baars & Van Kempen, 2009). In general, inner slope instability of predominately earthen dikes occurs when the loading conditions exceed the resistance, which is determined by the soil shear strength. For many dikes along rivers and coasts, inner slope instability occurs due to the infiltration of water into the dike body and its foundation, leading to high pore-water pressures, decreasing effective stresses and decreasing the soil shear strength (Frank et al., 2004; Sharp et al., 2013). The infiltration of water into the dike body can be caused by high water levels and heavy rainfall (Rikkert, 2022; Van Baars & Van Kempen, 2009). In addition to failures caused by water infiltration and high pore-water pressures, excessively dry conditions can also induce instabilities. Droughts can affect the stability of dikes in multiple ways, both positive and negative.

Soil moisture levels in the unsaturated zone directly affect the soil suction and weight. While higher soil suctions during droughts can enhance the soil strength along the slip plane, lighter soils can reduce effective stresses and worsen stability. However, lighter soils can also contribute to a smaller thriving moment, thereby enhancing stability. For hydraulic head levels, the relationship seems to be more straightforward: higher head levels result in less stable dikes, because of higher pore-water pressures and less effective stresses (Ridley et al., 2004). Next to the direct impact on effective stresses, droughts can also induce weakening processes that decrease the soil resistance. It is shown that the hydraulic material properties, including permeability and water retention, are affected by cyclic wetting and drying and are fundamental properties for slope stability (Stirling et al., 2021). Two failed Dutch canal dikes in the summer of 2003 were primarily caused by weight loss and soil shrinkage, due to the dry weather conditions (Bezuijen et al., 2005; Van Baars, 2005). Cracks in canal dikes can also reduce the shear resistance and may lead to the formation of shallow slip planes (Zhang et al.,

2021). In the past, dike failures were also attributed to seepage through cracks, caused by the settlement of the dike and due to shrinkage cracks on the dike body caused by exposures (Pigott et al., 1992).

The combination of a higher water level just after a drought can also result in unsafe situations. Deyer et al. (2009) suggest that desiccation can create an interconnected network of cracks that increases the permeability of the fill material and hence allows rapid seepage of flood water through the surface layer of the embankment. When the water level increases after droughts, high rates of seepage may cause localized uplifting of soil blocks, leading to progressive slope failure and successive breaching. Another cascading interaction between dry and wet situations impacting stability is an increased infiltration during extreme precipitation after prolonged drought caused by desiccation cracks in dikes (AghaKouchak et al., 2023; Jamalinia et al., 2020; Shao et al., 2015; Vahedifard et al., 2016; Vardon, 2015). Figure 2.1 conceptual illustrates the subsurface water conditions within dikes, the influence of various factors on groundwater flow and highlighting the impact of changing soil moisture and head levels on dike stability. The impact of changing subsurface water conditions on stability varies depending on the case. Nonetheless, studying how meteorological conditions affect these subsurface water conditions is the first step in understanding the impact of droughts on dike stability.

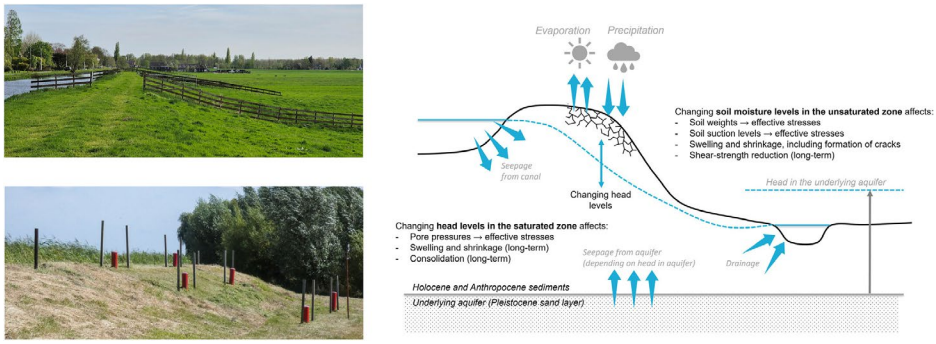


Figure 2.1 Right side: Conceptual figure that illustrates the subsurface water conditions in canal dikes and the way changing soil moisture and head levels can affect the stability. Left side: Photographs of two canal dikes in the Netherlands, with the MT-polder (top image) and the Duifpolder (bottom image).

The Dutch Situation and Safety Assessments

In the Netherlands, peat dikes consisting predominantly of original in situ peat with a clay cover, spanning about 3,000 km of the canal dikes (Bezuijen et al., 2005), are especially vulnerable to droughts. Variations in the weights of peat soils, an organic

soil made up of partially decomposed plant material with a high water content, can influence the stability, and shrinkage of peat can have large indirect effects on stability, such as causing internal seepage and erosion, and enabling the efficient transfer of high water pressures. Over the past decades, layers of clay have been added to these dikes to maintain an adequate crest height, increase their weight and help prevent the underlying peat soils from drying out. The recent Dutch landscape is formed by Holocene deposits, which are rich in organic matter, on top of Pleistocene sediments (Berendsen & Stouthamer, 2002). Drainage and excavation of peatlands lowered the surface locally and created deep polders with canal dikes situated on top of Holocene deposits. The hydraulic head in the underlying Pleistocene aquifer can exceed the surface (water) levels (Oude Essink et al., 2010), as illustrated in Figure 2.2. This geohydrological situation can result in marginally stable canal dikes, due to the lightweight peat soils and low effective stresses in soils. The drying of peat soils can be critical for dike stability, and a reduction of 300 mm of water subtracted from the unsaturated zone can lead to dike failures, as demonstrated by Van Baars (2005).

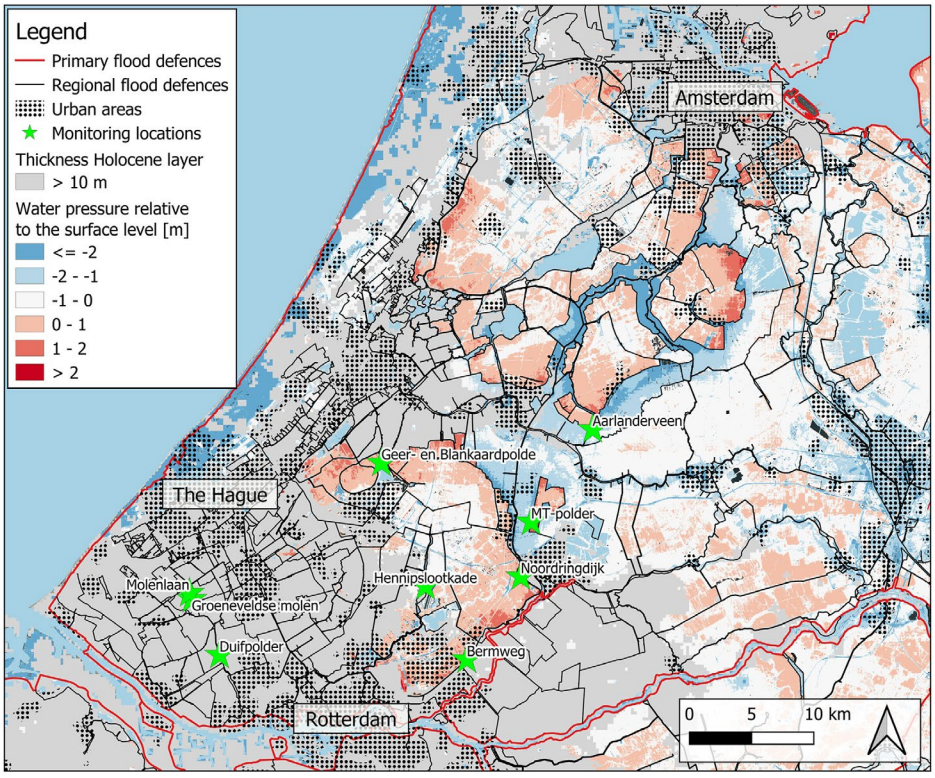


Figure 2.2 Locations of primary and regional flood defenses in the Western part of the Netherlands are indicated by red and black solid lines. Colors blue to red show the water pressure in the Pleictocene sand layer with regard to the local ground surface elevation

Water authorities assess the safety of canal dikes periodically (every 6 years) to ensure that they meet the required protection levels, which are often based on the acceptable risk of flooding (Vrijling, 2001). In line with the International Levee Handbook (Sharp et al., 2013), the Dutch guidelines prescribe to assess, among others, the failure mechanism of inner-slope instability, where two different loading events are considered: the wet and dry situation. In wet situations, the phreatic surface is close to the surface and soils are fully saturated, while under dry situations the phreatic surface is assumed to be several meters lower and the soil weights in the unsaturated zone are reduced. In both situations, the Mohr-Coulomb constitutive model is used to govern the soil mechanical behavior, assuming drained conditions. In practice, safety assessments indicate that most canal dikes are less stable during wet situations compared to dry situations, because lowering the phreatic surface increases effective stresses and enhances the dike stability, which outweighs the decrease in soil weight. The current safety assessment does account for the effects of fluctuating soil suction levels, swelling and shrinkage and shear-strength reduction, partly due to insufficient data availability. As a result, it may not accurately estimate risks during droughts realistically.

Drought recovery of dikes helps to understand seasonal transitions and the sequence of different failure mechanisms, during dry and wet situations. The geohydrological response to heavy precipitation varies depending on the hydraulic state of the dike. For example, in saturated soils, head levels will increase more than in dried-out soils, where more water can be stored in the unsaturated layers, Furthermore, as droughts are expected to become more frequent (Philip et al., 2020) and the time between droughts may become shorter than the recovery time for certain (ground)water systems, it is important to evaluate the recovery time of droughts. The recovery time is the duration required for a system to return to its pre-drought functional state. This metric is commonly used for ecosystems (Liu et al., 2019; Schwalm et al., 2017), but is also useful for the assessments of dikes. In this study, the drought recovery period is defined as the period between the moment of maximum drought and the moment that the average winter situation, between the beginning of October and the end of March, is reached. The average winter situation is characterized by nearly fully saturated soils, indicating the transition to a different season.

2.2.2 Drought Indicators

Droughts can typically be classified into three types: meteorological drought (resulting from rainfall deficit), soil moisture drought and hydrological drought (surface and groundwater water deficit) (Hisdal et al., 2001; Lloyd-Hughes, 2014; Van Loon, 2015). These categories can provide a useful framework for understanding the impact of droughts on dike stability. Meteorological drought indicators are often used for drought analysis, as they are easy to calculate and require less data and information compared

to other indicators. A meteorological drought can translate into a soil moisture and groundwater drought, as a lack of precipitation and ongoing evaporation can lead to decreasing soil moisture levels and lowering hydraulic head levels. Additionally, there is an interplay between soil moisture in the unsaturated zone and hydraulic head levels, because of the capillary rise of groundwater (when hydraulic head levels are shallow) and soil moisture content levels influence the ability of recharge (Hillel, 2003). Soil moisture droughts and groundwater droughts together form the drought in a dike and influence the stability by means of soil weight, soil suctions, hydraulic head levels and crack formation.

Meteorological Drought Indicators

In order to identify drought periods and compare drought severity between locations, meteorological-based (e.g. rainfall and evaporation) drought indicators are often used, for example, to decide whether inspections of dikes have to start. Meteorological data, such as satellite- and radar-derived precipitation amounts combined with climate and weather models (Muñoz-Sabater et al., 2021), are widely available and cover a long period, making them practical for decision-making purposes. Many meteorological drought indicators are available (Kchouk et al., 2021; Van Loon, 2015) and assess the meteorological and hydrological droughts and their relationships (Senatilleke et al., 2023). Two commonly used meteorological drought indicators are the Standardized Precipitation Index (McKee et al. 1993) and the more extensive Standardized Precipitation and Evaporation Index (Vicente-Serrano et al., 2010), referred to as the *SPI* and *SPEI*, respectively. These indicators measure deviations in rainfall (and evaporation) over a given period (e.g. 2, 6, or 12 months depending on the time scale of interest) compared to the long-term average over the same period. These indicators can be interpreted as the number of standard deviations by which the observed anomaly deviates from the long-term mean. In this study, the *SPI* and *SPEI* were calculated using a gamma and normal distribution transformation, respectively.

Additionally, this study looked at the PD, which is a commonly used meteorological drought indicator in the Netherlands (Van der Wiel et al., 2024). The *PD* is defined as the cumulative difference between precipitation and grass reference evaporation from April 1 onward (the start of the hydrological summer or growing season in the Netherlands). When the *PD* becomes negative, it is reset to zero (Beersma & Buishand, 2004, 2007). It is primarily used for agriculture purposes and similar measures, such as the Soil Moisture Deficit in UK, can be found in other countries (Clark, 2002; Schulte et al., 2005). In this study, the continuous *PD* at time t (PD_t) is defined as follows, where excess rainfall is fully drained (capped at zero):

$$PD_t = \max\left(PD_{t-1} + E_t - P_t, 0\right) \quad [2.1]$$

where *E* is the reference evaporation in mm based on the Makkink formula (De Bruin & Stickler, 2000), which is appropriate for Dutch climate conditions, and *P* is the amount of precipitation in mm.

Meteorological drought indicators can be locally estimated by considering local precipitation and evaporation levels. In this study, RADAR-derived precipitation amounts (Wolters et al., 2013), which have a spatial resolution of 1 km, and triangular interpolation between the estimated grass reference evaporation at KNMI stations are used. The radar-derived product is not available for long-term periods of historical meteorological data to derive standardized indicators. Therefore, the measurements at the KNMI station Rotterdam are used for historical data prior to 1 January 2020, thereby extending the time series back to 1990. The grass reference evaporation, which assumes a well-watered and maintained grass, is used rather than the actual evaporation, which is influenced by vegetation type and plant stresses. Some models can take into account factors like drainage, soil type and vegetation (Schulte et al., 2005), but additional information about these factors is often not present and for dike safety purposes barely used.

Figure 2.3 illustrates the development of the considered meteorological drought indicators by averaging values over all monitoring sites. In the summers of 2020 and 2022, the average maximum precipitation deficits were approximately 250 and 350 mm, corresponding to an exceedance probability of 1/15 and 1/60 per year (Beersma & Buishand, 2007). The *SPEI* aggregated over three months indicates extremely dry summers (*SPEI* < –2) for 2020 and 2022, while aggregating over 6-month period results in extremely dry summers in 2022 and a moderate drought (–1.5 < *SPEI* < –1) in the summer of 2020. The thresholds to identify droughts of dikes and prompt inspections depend on the susceptibility of the dikes to droughts, which is dike-specific and can range from minimal to substantial impact on dike stability. Commonly used thresholds for identifying droughts and initiating inspections include a threshold of –1 for *SPI*/*SPEI* (McKee et al., 1993) and a *PD* of approximately 175 mm, following common practices in the Netherlands, which is typically exceeded every 3 years (Beersma & Buishand, 2007). This study connects meteorological drought indicators with subsurface water conditions, and a subsequent step could involve quantifying their impact on stability to enhance the substantiation of thresholds.

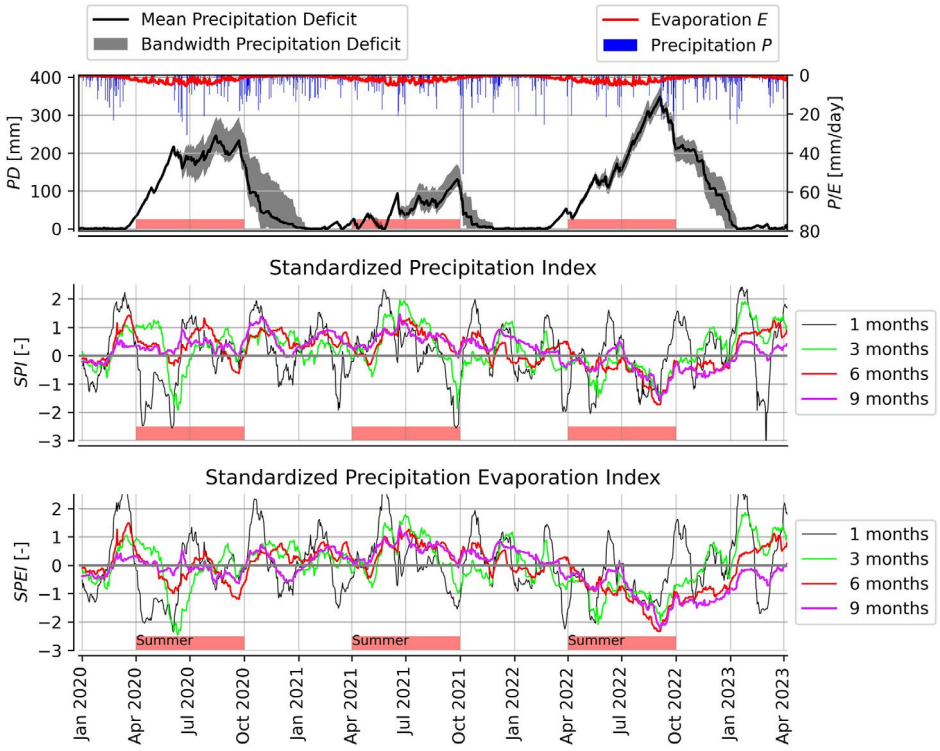


Figure 2.3 Different meteorological drought indicators from January 2020 to November 2022 (measurement period) by averaging values over all monitoring sites. The upper graph shows on the left y-axis the precipitation deficit and on the right y-axis the precipitation P (blue) and potential evaporation E (red). The middle and lower graphs show the SPI and SPEI for different time periods (1, 3, 6 and 9 months), where negative values indicate below-average conditions, suggesting drier or more severe drought conditions compared to the long-term average and positive values indicate above-average conditions.

2.3 Monitoring sites

2.3.1 Dike characteristics

The monitoring network consists of 10 locations where subsurface water conditions of canal dikes are measured (see Figure 2.2) from February 2020 to March 2023 (Strijker, 2023). These dikes are the result of historic human activities and primarily consist of various clays (silty and sandy clays with different levels of organic matter) and peats, with occasional thin layers of sand and gravel. Additionally, remnants of human activities, such as fragments of pots and glass, as well as metal objects, were found in the dikes. This makes the composition of the dikes rather heterogeneous, making it challenging to model and predict flow paths.

Under normal conditions, where excessive precipitation events are absent, the water levels in the canals remain constant, but the levels can increase by several tens of centimeters during heavy precipitation. The characteristics of these dikes differ, like head difference, dimensions and subsoil characteristics. The cross-sectional profiles and classified soil properties of the boreholes are shown in Figure 2.4. The degree of organic matter (moderate or high) in clay is high for all locations (peat contains by definition a large amount of organic matter). Based on soil properties, a distinction can be made between dikes where the dike body mainly consists of clay material (nr 1–4) and peat material (nr 5–10). Note that the peat dikes are not solely composed of peat; clay layers are often present, including a clay cover on top. The dike bodies at all the monitoring sites are located on Dutch Holocene deposits. The thickness of the Holocene deposits, which consist of clay, silt or peat, varies and shows a west- and northward thickening in the Netherlands (van der Meulen et al., 2007), resulting in an approximately 5m thick Holocene layer at Aarlanderveen and the MT-polder and about 15m at the Duifpolder (Table 2.2 shows the top of the Pleistocene layer w.r.t. The national reference level NAP).

Table 2.1 Characteristics of the monitoring sites (see Figure 2 for geographic locations). The reference level is NAP: Normaal Amsterdams Peil. TL = Target water level, ΔH= head difference between canal and polder, L= distance between inner crest and toe, β = orientation of the dike relative to the North, D_{PS}= the depth of the Pleistocene sand layer. The symbols are conceptually illustrated in Figure 2.5.

Nr	Location name	*TL* canal [m+NAP]	*TL* polder [m+NAP]	*ΔH* [m]	*L* [m]	*ΔH/L* [-]	*β* [º]	D_{PS} [m+NAP]	Distance first piezometer to canal [m]	Distance last piezometer to ditch [m]
1	Duifpolder Noord	-0.43	-3.10	2.67	12.5	0.21	255	-19.5	4.1	2.2
2	Duifpolder Zuid	-0.43	-3.10	2.67	12.5	0.21	255	-19.5	4.1	2.3
3	Groeneveldse-Molen	-0.43	-2.32	1.89	31	0.06	245	-21.5	3.3	18.0
4	Molenlaan	-0.43	-2.60	2.17	-	-	65	-21.0	3.7	-
5	Bermweg	-2.15	-4.55	2.4	37.0	0.06	320	-12.5	3.0	7.0
6	Hennipslootkade	-2.15	-4.70	2.55	36.5	0.07	130	-12.5	3.0	9.9
7	Noordringdijk	-2.15	-6.32	4.17	40.5	0.10	225	-13.5	3.9	13.2
8	MT-polder	-2.29	-5.84	3.55	44.5	0.08	105	-11.0	7.2	6.0
9	Aarlanderveen	-2.25	-4.44	2.19	25.5	0.09	295	-9.5	3.1	7.7
10	Geer-&Blankaardpolder	-0.61	-4.86	4.25	50.5	0.08	325	-13.0	4.6	12.5

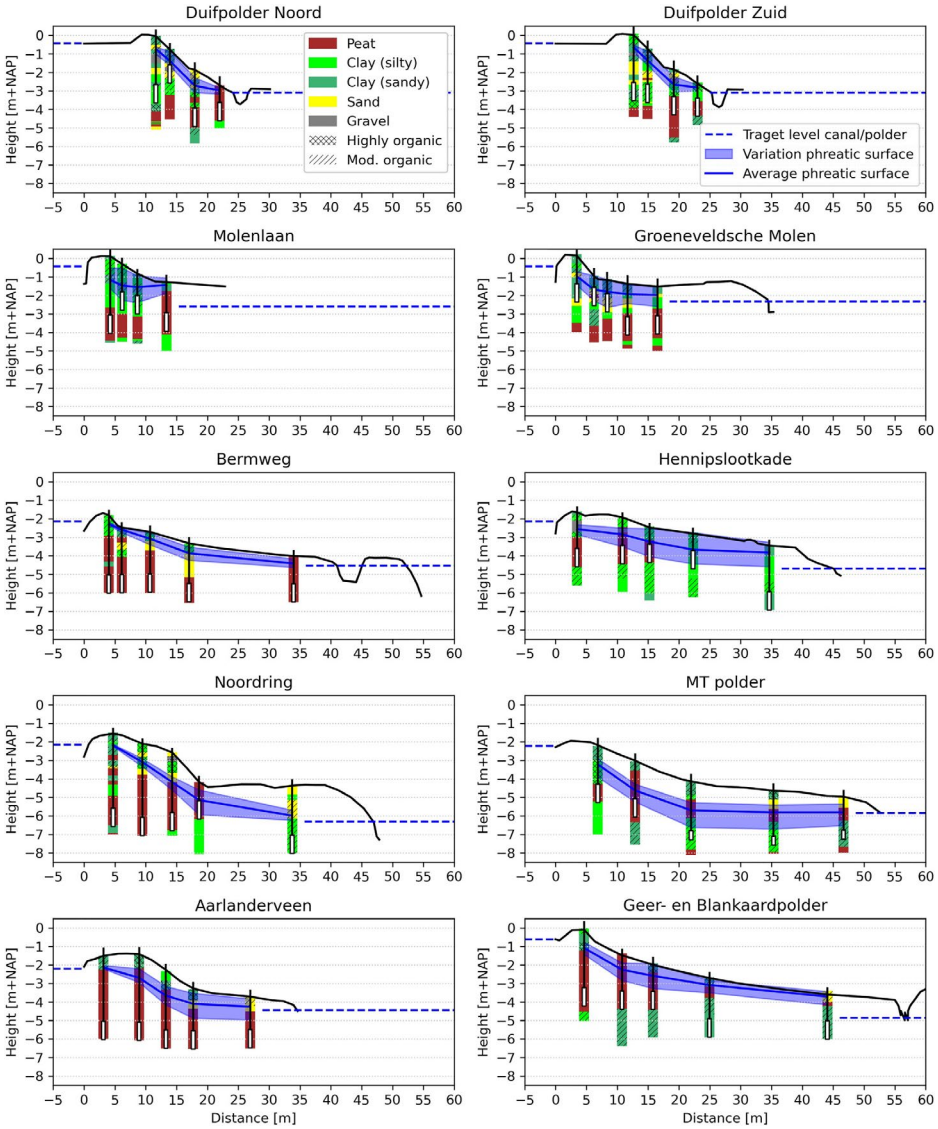


Figure 2.4 Cross-sectional profiles of monitoring sites. Solid black lines indicate the ground surface level and the dashed blue lines indicate the target water level in the canal and polder or ditch. Within every cross-sectional profile, boreholes are shown where colors indicate soil types and the vertical black line ending with a white rectangular box indicates the standpipes. The blue areas show the variation of the hydraulic head levels based on the highest and lowest measured hydraulic heads at the standpipes and represent the phreatic surface through the dike.

The monitoring network consists of 10 monitoring sites, each with four or five monitoring points. These points are positioned in a cross-sectional profile of the dike at which soil moisture levels (at different depths) and/or the hydraulic head level is measured. The dikes are all vegetation-covered with varying species that can have plant material above the ground surface up to 50 cm. The Noordringdijk includes a road construction near the toe of the dike, while the Hennipslootkade features a small asphalt cycle path along its crest. Other locations do not have any road constructions.

Soil samples from different monitoring points and depths are taken and analyzed in the laboratory to determine the in-situ and dry volumetric weight, in-situ water content, soil particle distribution by sieving (>63 µm) and sedimentation techniques (<63 µm) and organic matter. The soil samples were mainly taken up to 2 m below the ground surface, as they were used to interpret the soils where soil moisture sensors were installed (next paragraph). These data are available in Strijker (2023).

2.3.2 Monitoring equipment

The measured variables are the volumetric soil moisture content levels and the hydraulic head levels. The general set-up of the monitoring equipment can be seen in Figure 2.5, where five standpipes are installed within a cross-sectional profile and next to them the soil moisture sensors are installed at several depths. There are no soil moisture sensors installed next to standpipe P5, since fluctuations of the soil moisture are low due to (expected) high hydraulic head levels. At several locations, a conscious decision has been made to deviate from the general set-up, because there was not enough space to install five standpipes or the soil moisture sensors could not be placed because of the subsoil conditions (e.g., gravel layers). Table 2.2 shows the monitoring equipment used.

Table 2.2 Overview of Sensors That Were Used to Measure Certain Variables. Note: Numbers between brackets behind variables that are measured by the TEROS sensors indicate the model; soil electrical conductivity was only measured by TEROS-12 sensor, while soil moisture content and soil temperature were measured by both TEROS-11 and TEROS-12.

Sensor	Variable	Unit
TEROS-11/12	Soil moisture content (11/12) Soil temperature (11/12) Electrical conductivity (12)	[m^3/m^3] [degrees Celsius] [mS/cm]
TD-Diver	Hydraulic head level	[cm H_2O]

Soil moisture content

TEROS 11 and TEROS 12 sensors from METER Group were used to measure the soil moisture and soil temperature (TEROS 11/12) and the electrical conductivity (TEROS12 only). The majority of sensors utilized were TEROS11, with TEROS12 sensors specifically

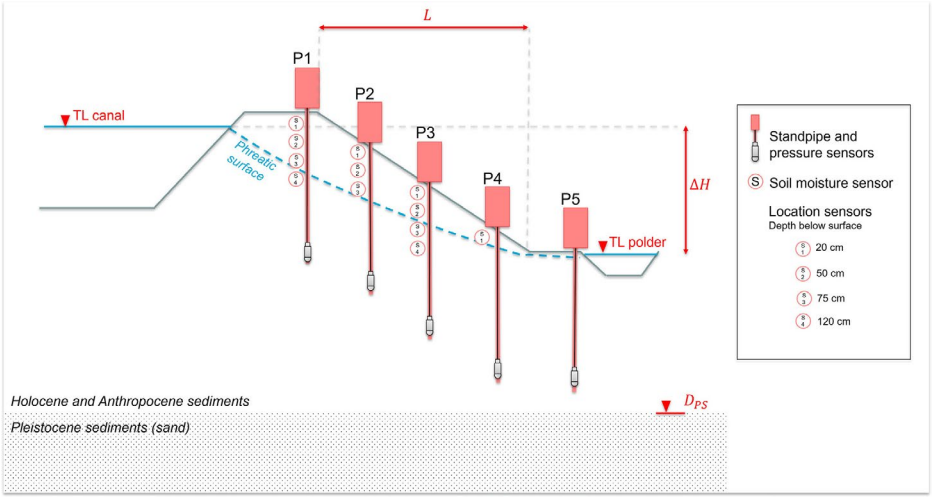


Figure 2.5 Schematic overview of the general set-up of the monitoring equipment. This cross-section also includes various definitions (in red) of the dike characteristics, as stated in Table 2.1.

placed in the dike crest. This placement was chosen due to the greater depth of the water table compared to the surface level, where cracks are more likely to occur. The formation of cracks can affect the evolution of the electrical conductivity (Kong et al., 2012), potentially providing valuable insights when electrical conductivity was measured. However, no significant cracks were observed at the monitoring sites. The TEROS 11/12 determines the soil moisture content levels using capacitance/frequency-domain technology. It uses an electromagnetic field between two metal electrodes (probes or needles) to measure the dielectric permittivity of the surrounding medium. The measurement sensitivity is contained within a 1.010-mL volume and the sensor uses a 70-MHz frequency that minimizes the salinity and textural effects. These sensors were not individually calibrated and the standard calibration curve is used that gives an absolute accuracy of ±0.03 m3/m3 for most mineral soils (Meter Group, 2021). The standard calibration curve was used for efficiency reasons, as individually calibrating all 120 sensors would be time-consuming and costly, while ensuring consistency in measurements across the locations. The sensor readings were recorded at 60-min intervals. The timezone was aligned with the local timezone and differs for the summer (UTC+2) and winter (UTC+1). These sensors were connected to a ZENTRA (ZL6) data logger by Meter group and sent the data every day to the data platform ZENTRA Cloud.

Hydraulic head levels

Hydraulic head levels were measured in standpipes, also known as open standpipe piezometers, that were placed at several locations within the dike body. The diameter of the installed standpipes and filters are both 32 mm. Filter gravels were placed around

the filters and on top a bentonite cement grout of 100 cm (and at some locations 50 cm) was used to prevent groundwater flow from other soil layers. The standpipes were sealed around with swelling clay. The standpipes were equipped with TD-divers (DI801) that measure the pressure. The accuracy of the TD-diver used is 0.5 cm H2O water column and can measure the water column up to 10 m (Van Essen, 2016). The sensor readings were collected at hourly intervals.

In low hydraulic conductivity materials, like heavy clay soils with a clay percentage of more than 50%, there is often disequilibrium between the head in the pipe and the head in the surrounding soil layer (Neuzil, 1986; Wolff & Olsen, 1968). Groundwater flow through low permeability layers can be slow and it may take hours to weeks for enough water to flow through the medium near the pervious section to establish equivalent heads in the pipe and the medium. Therefore, the measurements of the water level in the piezometer pipe may be subject to attenuation and time lags. An alternative is to use a pressure transducer that needs a smaller water quantity to obtain a measurement and gives a faster response time for measuring head level variations. For the purpose of this study, 18 pressure transducers were deployed at four distinct monitoring sites. However, it was determined that a majority (90%) of these sensors provided inaccurate and unrealistic data. As a result, standpipes were utilized as an alternative method for obtaining long-term monitoring data. The reason for this discrepancy remains unknown and requires further investigation. Nevertheless, the observed head levels demonstrate a flashy and reactive behavior, suggesting limited attenuation and damping, see Figure 7. The head response following a heavy rainfall event exhibits rapid peak reactions the day after. Moreover, given the focus of this study on droughts, where head level changes are generally slower, the influence of standpipes can be anticipated to be minimal.

2.3.3 Subsurface water measures for dikes

The volumetric soil moisture content is the volume of water within a total soil volume and it varies at different depths. For dike safety purposes, the total amount of soil moisture in a soil column is relevant (Van Baars, 2005) rather than a point measurement at one depth. Therefore, a measure that represents the integrated soil moisture in a soil column during droughts is introduced, called the Water Storage Capability (*WSC*). This measure gives the amount of water per unit surface area (in mm) that the soil surface absorbs before further precipitation cannot be stored in the profile (i.e., the soil has reached saturation level). The higher the number, the dryer the soil. While a lysimeter can measure the total moisture extraction of a soil column, in many cases, this must be estimated using point measurements that are integrated over the depth. Depending on the thickness of the unsaturated zone and available measurements, different soil column heights can be considered for the *WSC*. The saturation levels are not always known, unless soil samples are analyzed in the laboratory. However, with long-term

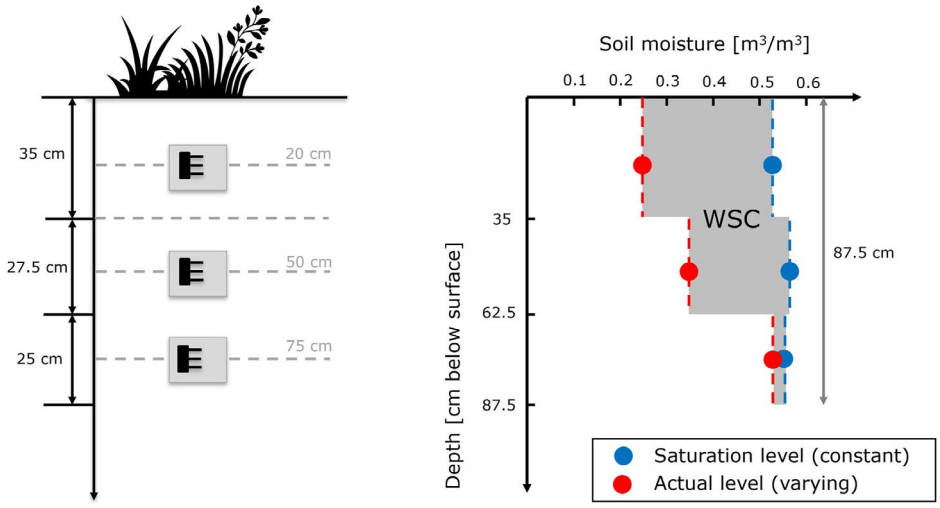


Figure 2.6 Conceptual illustration of how the water storage capacity (WSC) is calculated. Left drawing shows the different sensor depths and representative soil columns of every sensor. Right drawing shows how the WSC is calculated by showing actual soil moisture levels that can vary in time (red) and the saturation level that can be seen as a reference level which is constant and based on the maximum measured soil moisture levels. The volumetric water content does not necessarily increase with depth due to variations in soil types.

measurements and information from piezometers indicating shallow head levels close to the surface in winter, it can be assumed that soil moisture levels in winter will approximate saturation levels. Consequently, the maximum measured soil moisture content at various depths can be regarded as an approximation of the saturation level.

In our case, the *WSC* was estimated up to about 1 m below the surface (87.5 cm) by making use of the soil moisture sensors at 20, 50 and 75 cm depth. This integration depth captures most of the dynamics within the unsaturated zone of the dikes, as the majority of soil moisture fluctuations in the dikes occur within the upper meter of the dike, as can be seen in Figures A3, A4 and A5 in Appendix A. Figure 2.6 illustrates how the WSC can be calculated from the in-situ volumetric water content measurements. A simple depth integration is applied and the use of water retention curves may give more accurate estimates of soil moisture profiles and absolute WSC levels. Given the application of dike safety and the unknown matric suction profile, the WSC measure seems to be reasonable. The time series of the WSC for each location is calculated by averaging the WSC at the monitoring points where soil moisture levels are measured at depths of 20, 50, and 75 cm. This approach is adopted because the focus is primarily on the drying-out processes of the entire dike, rather than at a specific position in the cross-sectional profile. By averaging data from multiple sensors at different positions in the cross-sectional profile, sensitivity to uncertainties due to spatial variability

(resulting from different soil types) and measurement errors linked to individual sensors is reduced. The sensors are uncalibrated, which may result in consistent overestimations or underestimations of actual soil moisture levels, thereby impacting the WSC calculation. However, these systematic uncertainties do not affect the relative development of the WSC and, consequently, the drought development.

2.4 Results

2.4.1 Hydraulic head observations

Time series of measured hydraulic heads at all monitoring sites, as shown in Figure 2.7, illustrate the seasonal pattern with high head levels in the winter (October until March) and lower head levels in summer (April until September). The absolute hydraulic head levels, the manner in which head levels decrease during droughts and the responsiveness during heavy rainfall events vary by location. For example, a high responsiveness of heads to a heavy rainfall event in February 2022 with head increases of up to 50 cm at Hennipslootkade or Noordring (upper right graph in Figure 2.7), while heads at Aarlanderveen, Bermweg and Geer-en Blankaardpolder show limited increases (daily local rainfall amounts ranged from 30 to 50 mm between 19 and 21 February). This can be attributed to the presence of more peaty material in the subsoil at these locations with lower vertical hydraulic conductivity causing different geohydrological responses. During the summer of 2022, a gradual decrease in head levels of approximately 10 cm per week is observed (lower right graph in Figure 2.7). Despite these fluctuations during drought and heavy rainfall events, the water levels in the canals remain relatively constant, with fluctuations on the order of centimeters. This substantiates that the response of groundwater levels in dikes is primarily driven by meteorological variations.

By analyzing the frequency distributions of measured hydraulic heads using violin plots, some basic characteristics of the groundwater behavior can be extracted (see Figure A1 in Appendix A). At Duifpolder and Bermweg, the tail of the frequency distribution describing the lower head levels is more or less bounded indicating a limit to low head levels. At other locations, like MT-polder, Hennipslootkade and Geer-& Blankaardpolder the tail of lower head levels is long, indicating dropping head levels during the measurement period. Furthermore, upper boundedness can be noticed, for example, at MT polder and Bermweg, where the tail of high head levels is small. At 50% of the monitoring points near the inner slope and toe (position 2 and higher), the maximum hydraulic head levels exceed 0.2 m below the surface within the 3 years measuring period. In these situations, infiltration can be limited during more severe rainfall events as the dike is almost fully saturated, leading to increased surface runoff. This indicates an upper limit to the loading conditions of dikes, which can be relevant for the stability and failure probability of these dikes.

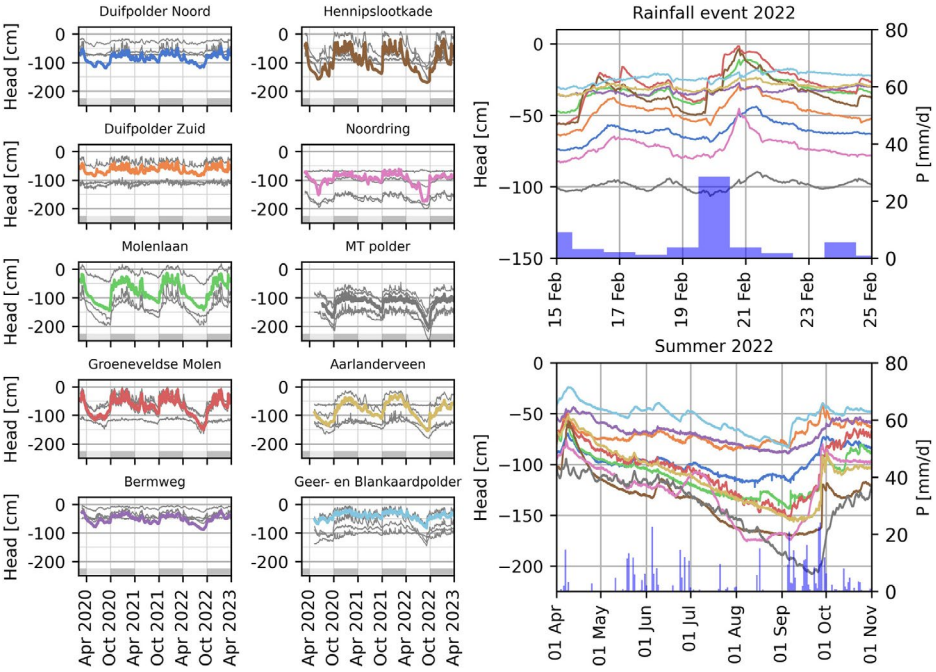


Figure 2.7 Small graphs on the left show time series of hydraulic head levels at the monitoring sites, where colored lines indicate the representative monitoring point at every location and the gray lines are the monitoring points at other positions within the cross-sectional profile. The representative monitoring point is number three in the cross-sectional profile where the standpipe is located in the inner slope. This monitoring point represents the fluctuations occurring well and is relevant for stability issues. Right: Head levels zoomed in at two events, namely a heavy rainfall event in 2022 (top) and the dry summer in 2022 (below), where the average precipitation and precipitation deficit of all sites is shown.

The measured head ranges are linked to the distance between the measurement point and the nearest drainage (canal or ditch), with higher fluctuations observed at greater distances (see Figure S2 in Appendix A). Therefore, wider dikes with a greater distance between the canal and ditch are expected to experience more head fluctuations than smaller dikes, due to a smaller head gradient and reduced seepage through the dike, coupled with decreased drainage of rainfall. However, this relationship remains ambiguous, and cannot be solely explained by soil type or geometry, even though these factors are expected to be significant, for example, according to Hooghoudt's equation (Hooghoudt, 1940). The heterogeneity of dikes and the resulting unknown flow paths and field hydraulic conductivities may contribute to the unpredictability of the geohydrological response.

2.4.2 Non-hydrostatic hydraulic head levels

In most safety assessments of dikes, the primary focus is on high-water events that increase hydraulic head levels or pore-water pressures, including non-hydrostatic heads within the dike, which can lead to failure (TAW, 2004; U.S. Army Corps of Engineers, 2003). For canal dikes, the hydraulic head levels in the dike body during droughts vary with meteorological conditions, and the way this is schematized in safety assessments for drought situations can be crucial for their stability. These head levels within the canal dike body are often assumed to be hydrostatic (Lendering, 2018), or they are linearly interpolated between the phreatic surface and the head in the underlying aquifer. This assumption is often made due to a lack of information about the vertical head profile. Additionally, hydrostatic pressure is relatively easy to incorporate in stability analyses. However, during dry summers, high hydraulic head levels in deeper layers and drying top layers (decreasing soil weights) can lead to severe stability issues for the dike. Time-dependent processes can result in non-hydrostatic head levels and unfavorable conditions for dike stability. At some monitoring points, two piezometers were installed in different soil layers, about 1.5 and 2.0 m vertically apart. The hydraulic heads in the dike body are non-hydrostatic during summers (see Figure 2.8), causing an upward flow that replenishes the soil moisture depleted by evaporation and root water uptake. In the summer, the head levels in deeper soil layers are up to 50 cm higher compared to shallow layers, whereas in the winter, the head levels tend to approach hydrostatic conditions, where the head levels approach each other. This could have implications for dike stability, because, in this scenario, the effective stresses within the dike can be lower than initially accounted for, due to the remaining high head levels and a reduced soil weight as a result of drying out.

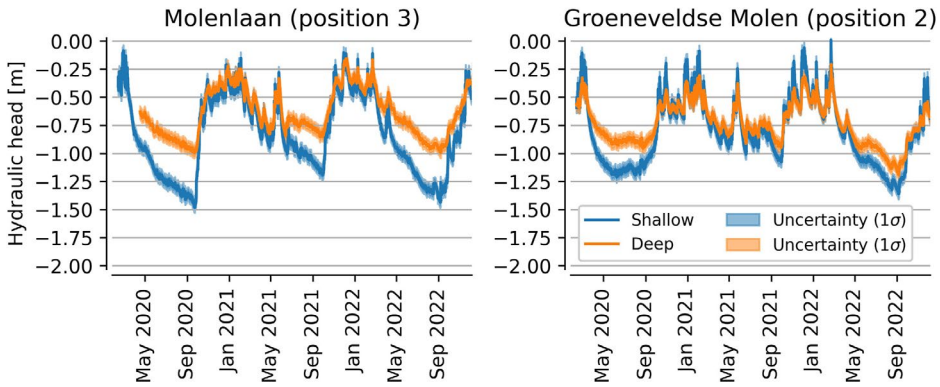


Figure 2.8 The hydraulic head levels at standpipes that are located at different depths for two monitoring sites. The filter depths (bottom of the filter) in meters below surface level are 2.1/3.45 m (shallow/deep) at Molenlaan and 2.55/4.65 m (shallow/deep) at Groeneveldse Molen.

2.4.3 *Drying of unsaturated soils*

During winter, the unsaturated soils are close to saturation with high groundwater levels and drying occurs during spring and summer, as can be seen by the development of the WSC in Figure 2.9. The decrease in soil moisture levels in the canal dikes is caused by water extraction from the soil by evaporation and transpiration and dropping groundwater levels, but can be replenished by upward capillary flow from the groundwater. The balance between these processes differs for every location. For example, the WSC at the Bermweg remains relatively low during all three summers, caused by high groundwater levels that are capable of supplying the unsaturated soils opposing drying from the surface.

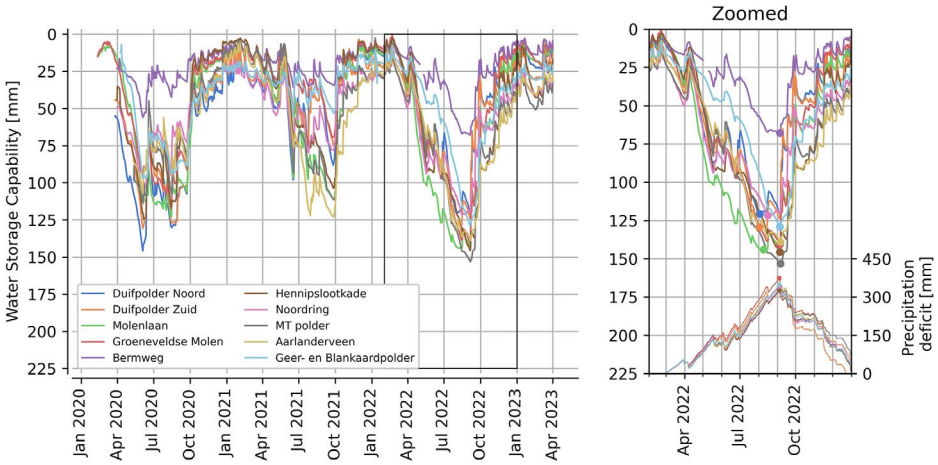


Figure 2.9 WSC during the period 2020–2023. The black box in the left graph indicates the zoomed period in the right graph, where the precipitation deficits are also indicated. The points in the right graph mark the maximum WSC and PD.

The dry summer of 2022 led to dropping soil moisture levels with maximum WSC ranging from 65 mm at Bermweg to 153 mm at MT-polder. The summer of 2021 was meteorologically wetter (with a three times smaller average *PD* than in 2022) and resulted in a larger variation in maximum WSC among the monitoring sites compared to 2022. Despite 2022 being meteorologically wetter, the differences in maximum WSC between the summers were small at Aarlanderveen and MT-polder, two peat dikes (Aarlanderveen: 122 mm in 2021 and 138 mm in 2022; MT-polder: 111 mm in 2021 and 153 mm in 2022). This demonstrates the significance of groundwater flow in dikes and its impact on infiltration, evaporation, and water distribution within the dike, which can result in different severities of drought from meteorological and soil moisture perspectives. An explanation of the two deviating peat dikes can be the hydrophobic property of peaty materials that can block infiltration (Dekker & Ritsema, 2000) or limited groundwater flow through dry, unsaturated soils with very low hydraulic conductivities (Hillel, 2003).

2.4.4 Relationship between soil moisture and hydraulic head

Soil moisture is strongly linked to the water table in the dike due to capillary rise. The relationship between the soil moisture in the unsaturated zone and hydraulic head levels deeper in the dike is crucial to understand the joint impact of both states on dike stability. Moreover, these relationships are also useful for dike monitoring; soil moisture levels can be measured during inspections, and remote sensing techniques are available that observe soil moisture levels in dikes (Aanstoos et al., 2011; Cundill, 2016). The conditions of soil moisture and water table, and their interactions, are not static but dynamic and influenced by time-dependent changes in environmental conditions (e.g., rainfall and the evaporative demand of the atmosphere) and the storage capacity and transmissivity of soils. On top of this, there is also the effect of hysteresis where the water content—hydraulic head relationship follow state dependent different paths for the wetting and drying cycles. In general, unsaturated soils dry out at the surface due to evapotranspiration, and when the water table is shallow, upward capillary flow can replenish the drying soils when the hydraulic conductivity of the soil is large enough. However, the water table in canal dikes itself is also dynamic and affected by the hydraulic resistance of the dike and seepage from deeper soil layers. This complexity adds challenges to data analysis when studying the interactions with soil moisture. Whether a soil profile is dominated by capillary rise can be investigated by considering the soil moisture levels at a specific depth in relation to the height above the water table (see Figure 2.10). In cases where capillary rise governs the soil profile, the moisture levels tend to exhibit a relatively small range of variation in relation to the water table height. At Aarlanderveen, the range of soil moisture is narrow at small heights above the water table and widens as the height increases. This suggests the presence of sufficient capillary rise near the water table, while evaporative demand becomes more dominant closer to the surface. Influences of the capillary rise can also be noticed at Noordring, Bermweg and Duifpolder Noord. At other locations, the range of soil moisture is wide, indicating a dominant influence of the evaporative demand at the surface, for example, at Molenlaan and MT-polder. Saturated soil moisture levels (moisture values below the water table) exhibit variations, such as Geer&Blankaard, Bermweg, and Groeneveldse molen. These variations can be explained by changes in soil volume resulting from soil compaction caused by changes in effective stresses (see also Figures A3–A5 in Appendix A) or non-hydrostatic hydraulic head levels, which can cause the actual water table to be lower and the soils to be unsaturated.

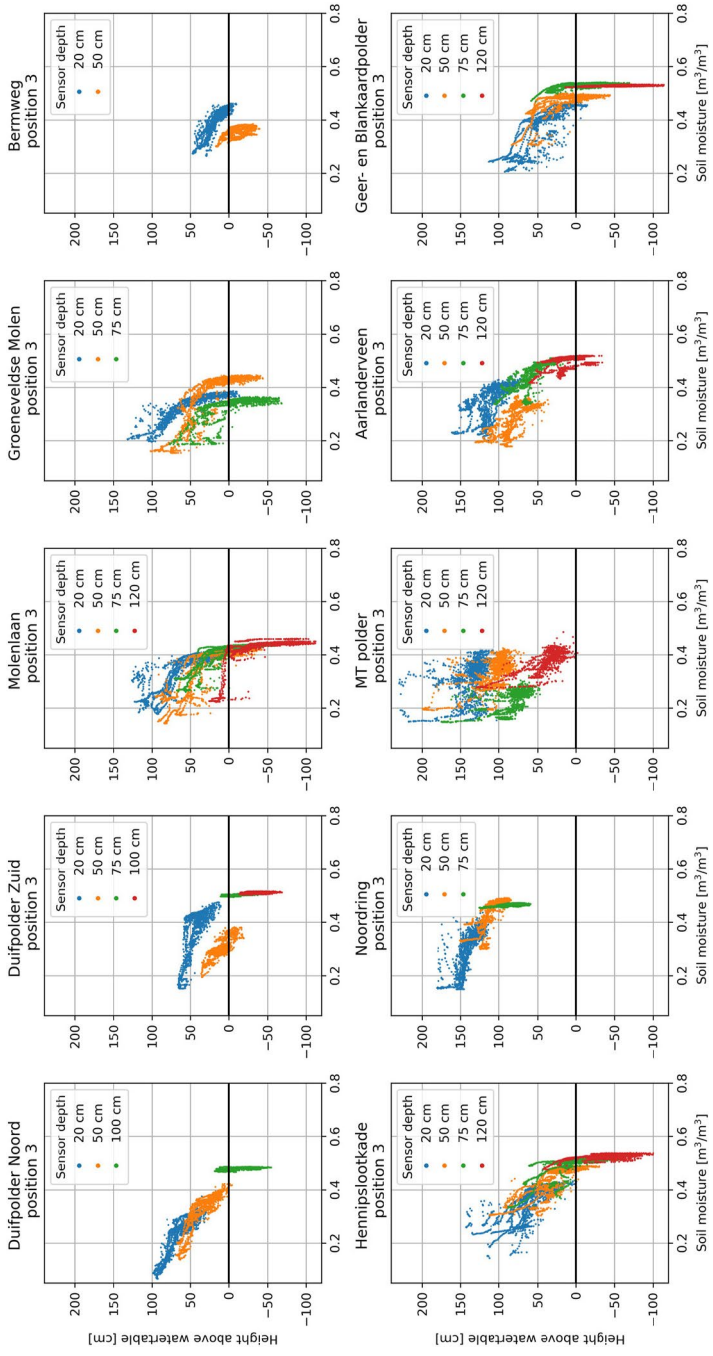


Figure 2.10 Measured moisture content levels plotted (x-axis) against the sensor height relative to the groundwater level (y-axis) for different soil moisture sensors at 3rd monitoring point in the cross-sections.

2.4.5 *Applicability of meteorological drought indicators for dikes*

For dike stability, the development of the absolute level of hydraulic head and soil moisture are of main interest and are used as target measures in this analysis. Three types of meteorological indicators were assessed: the *PD* and the standardized indicators, *SPEI* and *SPI*, for different timescales (1, 2, 3, 6 and 9 months). The relationships between these meteorological drought indicators and subsurface water measures (hydraulic head level at the inner slope and *WSC*) are evaluated using the adjusted *R*-squared of a second-order polynomial regression and the Pearson correlation coefficient and the Spearman rank correlation (Wright, 1921). The relationships are assessed for every individual location for the summer period, from April until September. In this section, the results presented using the adjusted *R*-squared, with the Pearson correlation coefficient and the Spearman rank correlation are only reported, Figures A6 and A7 in Appendix A, and yield similar findings.

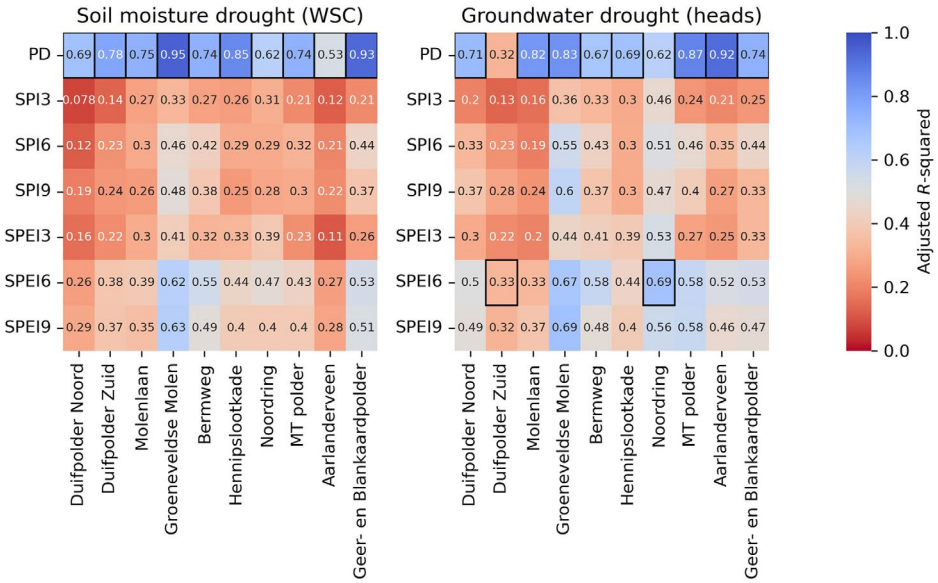


Figure 2.11 Relationship between different meteorological drought indicators and subsurface water measurements in dikes was evaluated using the adjusted R-squared. Values in the heatmap are the adjusted R-squared of the polynomial regression: higher values indicate a better fit than lower values.

The *PD* performs as the best meteorological drought indicator to assess the dryness of subsurface water conditions in nearly all canal dikes considered, see Figure 2.11. In general, it provides significantly better estimates of subsurface water conditions during droughts than the standardized indicators. Precipitation (or its absence) and evaporation are the primary atmospheric processes directly linked to soil moisture fluctuations, which makes the *PD* emerges as the top performer among the indicators

for the *WSC*. However, certain factors such as percolation, capillary rise, surface runoff, and soil moisture redistribution are not considered by the *PD* and could contribute to discrepancies in its performance across different dikes. Since groundwater levels in dikes are shallow and react quickly to changes, often within a day, this rapid response contributes to the effectiveness of the PD. At two locations, the *SPEI-6* proved to be a more effective indicator for groundwater drought. In one location, head level fluctuations are generally minimal, likely due to significant influence from horizontal seepage from the canal. An advantage of the *PD*, in relation to standardized indices, is that you do not need to select a specific time frame, making the application of this drought indicator easier. The standardized indicators have limitations in estimating the absolute subsurface water conditions, as they filter out seasonality and express droughts in relative terms. Furthermore, it can be seen that the evaporation term is crucial for the subsurface water conditions, both for soil moisture deficit and groundwater levels, as the *SPEI* generally shows a better fit than the *SPI*. Figures A8–A11 in Appendix A provide a more detailed view of the relationship between the two most effective meteorological drought indicators (*PD* and *SPEI-6*) and the subsurface water measures.

Although the *PD* is the most effective meteorological drought index, the absolute severity of droughts in dikes varies strongly. Therefore, these indicators provide information about the dryness at each location in relative terms, while the absolute subsurface water conditions depend on other local dike characteristics. Furthermore, the *WSC* measure is limited to the soil moisture in the upper meter of the dike. For a high PD, the amount of soil moisture extracted may extend to deeper soil layers, which are not included in the *WSC*. As a result, the *WSC* may reach a maximum value contributing to the fitted polynomial model that is concave down at eight locations. However, another factor that can also impact this trend is the prolonged duration of drought. During prolonged droughts, the evaporation rate can be limited or dictated by the rate at which the gradually drying soil profile can supply moisture to the evaporation zone, as the hydraulic conductivity of dried soils decreases (Hillel, 2003). At the locations Bermweg and Geer&-Blankaardpolder, the fitted polynomial model is concave up, which can be explained by the interaction with the head levels: at the beginning of the dry periods, high heads kept the *WSC* small, but later on, as head levels dropped, the soils still dry out.

2.4.6 Drought recovery

The moment of maximum drought (start time of recovery period) and recovery moment (end time of recovery period) were calculated for different concepts of drought (meteorological, soil moisture and groundwater). These moments vary by location and drought concept, as shown in Figure 2.12. For most locations, the recovery moment for the soil moisture drought occurs later than for the groundwater drought, as can be seen during summers in 2020 and 2022. This is different compared to other

systems, like river catchments (Van Loon, 2015), and can be caused by different storage characteristics. The soil moisture and groundwater recovery following the exceptionally dry summer of 2022 took up to 4.5 months across various locations and drought recovery occurred in January next year. This exceeded the onset of the subsequent storm/winter season. Despite heavy precipitation events in the winter, the previously dried soils were able to effectively store the water, thus mitigating potential high hydraulic head levels and instabilities. Furthermore, the maximum and recovery moment of the PD, representing the meteorological drought, are indicated in Figure 2.12, which differ from the subsurface water conditions. These differences can be caused by physical elements affecting the infiltration, total evaporation and water redistribution within the dike that are not accounted for in the PD.

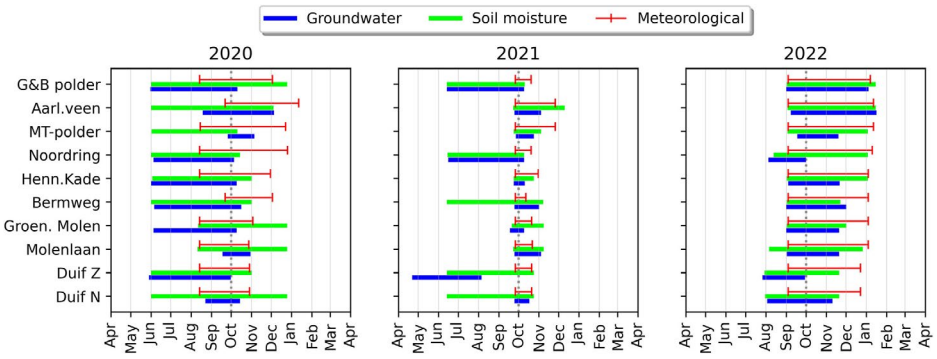


Figure 2.12 Horizontal bars indicate the recovery period, which is defined as the difference between the moment that maximum drought (highest WSC or lowest hydraulic head level) occurs and the moment of recovery. Vertical red markers indicate the moment of maximum drought and recovery moment based on the continuous rainfall deficit (meteorological drought). The gray dotted line indicates the start of the storm/winter season. During this period there are high probabilities of hydraulic loads.

2.5 Discussion

2.5.1 Limitations

The data set comprises 10 monitoring sites, which is unique in the Netherlands, but statistically insignificant given the large heterogeneity of canal dikes. While the data set includes some typical canal dikes, such as peat dikes with gentle slopes that are common in the Netherlands, its representativeness for the geohydrological response remains uncertain, primarily due to limited multi-year observations in general. The accuracy of the soil moisture measurements requires attention since the factory calibration equation is used, which can be improved by a custom calibration. Furthermore, attenuation and time lags can occur when measuring head levels in standpipes that are placed in soils with low hydraulic conductivity levels, like clay and peat.

In addition, it is important to consider that it is assumed that several point measurements of soil moisture levels together represent the state of a certain dike. However, uneven wetting and preferential flow paths, caused by burrowing animals, plant roots or cracks, can make the point measurements unrepresentative. Next, the applicability of meteorological drought indicators for dikes has been assessed based on absolute values of soil moisture and groundwater levels. In this case, standardized indicators may not perform as well, but when one is interested in how dry a dike is at a certain moment compared to previous years, standardized meteorological indicators may be more suitable. Lastly, the definition of drought recovery is an assumption in this study, considering the period between the maximum drought and the average winter situation. This choice is based on the application to dikes with two failure mechanisms in summer and winter. However, drought recovery can also be defined differently which can give different outcomes.

2.5.2 Implications

Currently, nearly all inspections by Dutch water authorities are conducted visually, with other methods used occasionally or in the early development phase, like deformation data using satellite radar interferometry and ground penetrating radar (Chlaib et al., 2014; Özer, Rikkert, et al., 2019). Research has shown that the accuracy of visual inspections is limited (Klerk et al., 2023). Additionally, visual inspections provide information about surface processes, but it is expected that the dominant failure mechanisms are driven by subsurface processes that are not always visible at the surface, such as the two Dutch canal dike failures in 2003. To properly manage risks, water managers should not only focus on visual inspections, but also on monitoring or estimating local subsurface water conditions and associated failure probabilities. This study showed that the *PD* is the best meteorological drought indicator for both soil moisture and groundwater droughts and can be used to organize inspections and substantiate decision-making. Moreover, this indicator can provide valuable insights into subsurface water conditions across various dikes, not only in the Netherlands but also in other countries with similar types of dikes. It can serve as a proxy for subsurface water conditions, allowing for a more substantiated quantification of wetting and drying over the years, when no measurements of the subsurface water conditions are available. This research highlighted the presence of non-hydrostatic pore water pressures in dikes during droughts, which can negatively influence dike safety; the top soil layers dry out, reducing soil weights, while the reduction of deeper pore-water pressures is delayed. This can result in situations where effective stresses are lower than those assumed under hydrostatic conditions. To accurately evaluate dike stability during droughts, safety assessments must take into account these non-hydrostatic conditions.

2.6 Conclusions and recommendations

Data analyses were carried out that improved the geohydrological understanding of canal dikes, particularly concerning the development of droughts, and also in wet situations. Observations indicated that hydraulic head levels in canal dikes were very shallow during the winter, with approximately 50% of the monitoring points located on the inner slope and toe of the dike experienced head levels above 20 cm below the surface. These levels are close to the physical maximum, since these are phreatic heads within the dike body and further saturation of soils limits the infiltration of rainfall. During the monitoring period, two extremely dry summers occurred (2020 and 2022) giving valuable insights into the geohydrological extremes, where the minimum and maximum observed groundwater levels can lie almost 2 m apart. These variations occurred while the canal water levels were more or less constant, highlighting the importance of meteorological conditions. The head levels in several dikes were non-hydrostatic during dry summers, which are often not accounted for in dike safety assessments and can worsen the stability during droughts.

Performances of *PD* and standardized drought indicators (*SPI* and *SPEI*) to estimate subsurface water conditions in dikes were evaluated. The *PD* appeared to be the most reliable meteorological drought indicator for droughts in canal dikes, for both soil moisture and groundwater droughts. The standardized indicators have limitations for dike safety applications, because they filter out seasonality and express the drought in relative terms, while for dike safety the absolute subsurface water levels are of main interest. Last, the drought recovery was analyzed, which indicates the transition between seasons and the corresponding causes of failure, initiated or affected by varying subsurface water conditions. The soil moisture and groundwater levels following the exceptionally dry summer of 2022 took up to 4.5 months to recover across various locations. This exceeded the onset of the subsequent storm/winter season. Despite heavy precipitation events during the winter season (in January 2023), the previously dried-out soils are able to effectively store the water, thus mitigating potential high hydraulic head levels that are relevant for failure mechanisms in the winter.

Although we showed that meteorological drought indicators can provide information about subsurface water conditions, the use of soil moisture and hydraulic head data directly, or in combination with groundwater models, can help to identify weak spots even more accurately. We encourage taking and sharing multi-year measurements of subsurface water conditions in dikes to create a larger data set to improve our geohydrological understanding and make results and conclusions statistically more significant. Furthermore, to accurately assess the safety of dikes in terms of failure probabilities, the extreme loading conditions (head levels, soil moisture content and suction levels) have to be estimated in terms of probabilities. Multi-year observations

are often insufficient to estimate extreme conditions (e.g., the head level with a probability of 1/100 per year), therefore methods to estimate these extreme conditions are necessary. Furthermore, these methods can help to assess how failure probabilities change with changing environmental conditions, like land subsidence and climate change and assess the effectiveness of measures. For example, canal dikes that are prone to droughts are commonly strengthened in practice by covering the dike with a sandy clay layer to (a) increase the weight and height of the dike and (b) to make the underlying peat or organic clay layers less vulnerable to droughts. The impact of this new top layer on the subsurface water conditions is often not incorporated in the design, because lack of reliable measurements and models, despite its potential to affect the effectiveness of measures aimed to reduce failure probabilities.

Data availability statement

The following data are available at the 4TU.ResearchData at https://doi.org/10.4121/136aa5df-1907-43ac-a5b0-e0ea8f2dedf3 (Strijker, 2023):

- Soil moisture, soil temperature and soil salinity data (Excel-files)
- Hydraulic head data from standpipes (CSV-files)
- Cross-sectional dike profiles (Excel-files)
- Bore profiles (pdf-files in Dutch) and soil characteristics data (Excel-file)
- Rainfall and potential evaporation data (CSV-file)

Readme files are added and describe the data files. The data source has a CC0 license, which entails the waiver of all copyright and related rights, enabling unrestricted use of the data for any purpose. Authors appreciate being informed when using the data by contacting the corresponding author. New data has been added in the summer of 2025, and will be added in the future.

3

Dynamics of peak head responses at Dutch canal dikes and the impact of climate change

This chapter examined the dynamics of peak heads in canal dikes on a national scale using time series models calibrated on a unique dataset of head observations across the dike system. Various model structures were evaluated, and a non-linear model performed the best. These models were used to simulate 30 years of head time series representing current and future climate scenarios. Subsequently, dike clusters were identified based on the coincidence of peak heads, allowing for the identification of dikes where peaks are caused by (dis)similar types of rainfall events. The differences and similarities in peak head response between dikes and identified clusters were related to physical dike characteristics. While the subsurface material and dike width appeared to influence the head response variation of clusters, their presence across multiple clusters indicates that they do not yield a definitive outcome. Moreover, peak head statistics across various dikes indicated that extreme and yearly occurring load conditions are relatively close to each other, with a median decimate height of only 15 cm. With climate change driving higher winter precipitation and summer evaporation, head statistics are changing. By 2100, extreme peak heads are expected to occur between 3 times less and 8 times more frequently, depending on the climate scenario and the type of canal dike.

This chapter is based on Strijker, B. and Kok, M. (2025). The dynamics of peak head responses at Dutch canal dikes and the impact of climate change, Nat. Hazards Earth Syst. Sci., 25, 3355–3379, https://doi.org/10.5194/nhess-25-3355-2025

3.1 Introduction

Catastrophic dike failures have occurred throughout history due to various causes, such as storm surges, extreme river discharges, ice drifts, and extreme weather conditions like heavy rainfall or drought. Several failure mechanisms were involved, including overflow and overtopping, external erosion, piping, and inner slope instability (Van Baars and Van Kempen, 2009; Özer et al., 2019). For many dikes along rivers and coasts, inner slope instability occurs due to the infiltration of water into the dike body and its foundation, leading to higher head levels and pore-water pressures, reducing effective stresses and shear strength of the soil (Frank et al., 2004; Ridley et al., 2004; Sharp et al., 2013; van Woerkom et al., 2021; Van der Krogt et al., 2022). The infiltration of water can be caused by high water levels against the dike as well as heavy rainfall (Rikkert, 2022; Van Baars and Van Kempen, 2009). For dikes with controlled water levels that show little fluctuations, such as canal dikes, the infiltration of water caused by heavy rainfall can be significant and is considered a primary mechanism of dike failure. Canal dikes are among others present in polders, which can be found in coastal and alluvial lowlands all over the world, like the Netherlands, Bangladesh, Vietnam, England and China (Martín-Antón et al., 2016; Morton and Olson, 2018; Lendering et al., 2018; Tran and Weger, 2018; Triet et al., 2017; Manh et al., 2013; Warner et al., 2018). The water levels in these reclaimed areas are artificially regulated by an internal drainage system with canals, where water levels can reach several metres above the surrounding terrain (see Figure 3.1). This makes these low-lying areas vulnerable to floods in the event of a canal dike breach. These canal dikes are found not only in polders but also along internal waterways and irrigation canals worldwide, where major dike failures have occurred throughout history (Gildeh et al., 2019).

Figure 3.1 Two examples of canal dikes in the Netherlands: Duifpolder is shown on the left and the Drooggemaakte Geer- en Kleine Blankaardpolder is depicted on the right. In both images, the canal has permanent high water levels close to the crest level (on the left side) and the low-lying polder is mainly used for agriculture (on the right side). The head differences between the canal water levels and water levels in the polders are 2.7 m (Duifpolder) and 4.3 m (Drooggemaakte Geer- en Kleine Blankaardpolder). The red tubes protect the measuring equipment and the piezometers used to measure the hydraulic head. Photos by EURECO/Cyril Liebrand (2022).

To manage flood risks in an embanked area, the performance of dikes plays a crucial role. The failure probabilities of the individual elements or dike stretches contribute to the reliability of the canal dike system and the flood risk level in an area (Vanmarcke, 1977; Kanning, 2012; Jongejan et al., 2020). The Netherlands has an extensive system of canal dikes critical for managing its low-lying land and protecting it against flooding, with a total length of more than 10 000 km (Pleijster et al., 2015). In general, the failure probability of a dike system increases with the total length of the dike system, due to partial correlation or independency between different individual dike stretches (Kanning, 2012; Vrijling and van Gelder, 2002). This phenomenon is known as the length effect, which is caused by both spatial dependencies in the resistance and loading conditions and differs for each failure mechanism. Inner slope instability is a failure mechanism of the Dutch canal dikes that contributes significantly to the calculated failure probability and flood risks in the polders, where the load primarily consists of high hydraulic head peaks (Lendering et al., 2018; Rikkert et al., 2022; Van Baars and Van Kempen, 2009). The variations of the canal water levels in the drainage systems of Dutch polders are small (typically up to tens of centimetres), while the observed hydraulic head fluctuations are an order of magnitude larger than water level fluctuations. Therefore, the fluctuations of the hydraulic heads are primarily driven by rainfall and evaporation. Whether two nearby canal dikes both experience extreme load conditions after a heavy rainfall event depends on the head response of the dikes. Variations in the head response can cause extreme loads to occur after different weather events and influence the system's reliability. Furthermore, these variations in response also help water authorities identify threatening situations, as weather forecasts can be translated to potential peak head levels in dikes.

To calculate the failure probability of individual canal dikes and dike systems, information about peak head responses is essential. Currently, there is limited understanding regarding the spatial variability in head responses and head statistics in canal dikes, which is partly due to the lack of measurements and the extensive computation time required for groundwater models. Multiple studies have modelled the effects of rainfall and evaporation on the phreatic surface in dikes using different approaches (Rikkert, 2022; Jamalinia et al., 2020; van Esch, 2012). Multi-year measurements of hydraulic head levels are often lacking in dikes, which makes modelling exercises difficult to validate. Furthermore, the validation of models is hindered by the heterogeneity of dikes and the unknown field hydraulic conductivities, potentially influenced by burrowing animals, plant roots or cracks, and resulting flow paths. This means that there is also little known about the effects of climate change on head statistics and failure probabilities of canal dikes. Future climate projections indicate increasing temperatures with summers becoming hotter and drier and winters becoming warmer and wetter. This is expected to affect the stability of slopes.

Although the impact has been studied for both natural slopes (e.g. Moore et al., 2010) and earthworks (e.g. Huang et al., 2024; Rouainia et al., 2020) for canal dikes with different boundary conditions, subsurface materials and resulting head dynamics, the effects are studied to a limited extent. This study aims to assess the dynamics of peak hydraulic heads in canal dikes on a national level, caused by heavy rainfall events, by analysing the variation in head responses and head statistics. It also seeks to understand why differences in head dynamics occur by relating these variations to the physical properties of the dikes. Furthermore, the potential impact of climate change on the head statistics is quantified, indicating how flood risks in Dutch polders are expected to change in the future.

3.2 Study area and data

3.2.1 Dutch canal dike system

In the Netherlands, the threat of flooding is controlled by a system of flood defences, where distinction can be made between primary and regional flood defences. The primary flood defences are located along major bodies of water, such as the sea, the major rivers and large lakes, often referred to as outside waters, while regional defences are found along inland waters, including drainage canals, artificial lakes and smaller rivers. In general, a breach in regional defences will have a smaller impact than a breach in the primary defences, though it can still have considerable consequences, as shown in Figure 3.2. This study focuses on a subset of the regional flood defences, namely the canal dikes. The canal dikes are primarily located in the western and northern parts of the Netherlands, where the polders are located (see Figure 3.2). These cultivated lowland areas serve as agricultural land as well as for human settlement. Many cities, villages and small communities are situated throughout the polders. The water inside the polder is separated from the outside water by the primary flood defences, and the polder drainage systems manage the water inside the primary flood defences. The water is managed by discharging or pumping the polder water into canals (also called the boezem in Dutch), after which the canals release the water into the outside water, either naturally or using pumps (Steenbergen et al., 2009). Water levels in the canals are higher than the polder levels, resulting in a flood hazard for the polders that are protected by the canal dikes. The subsurface of canal dikes is characterized by low-permeability soils that mainly consist of clay and peat, and in the past many canal dikes breached (Van Baars and van Kempen, 2009).

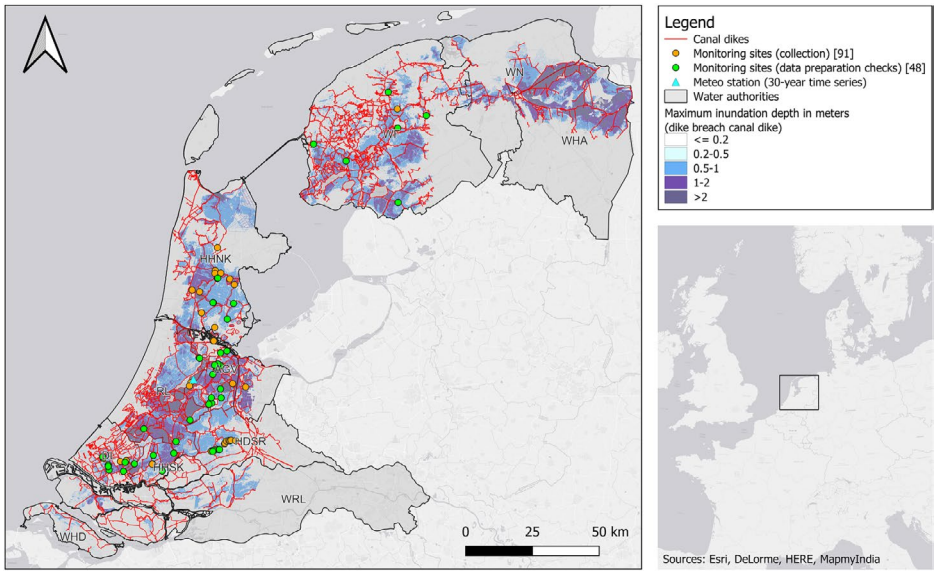


Figure 3.2 Overview of the study area with the canal dike system (subset within the regional flood defence system for the provinces of South Holland, North Holland, Friesland, Groningen and Utrecht). Monitoring sites are indicated by circle markers, with orange and green denoting collected data and the data checked and utilized for further analysis, respectively. Maximum inundation depths are depicted to illustrate potential flood impact in the polders.

3.2.2 Data collection

Head observations and preparation checks

To set up and calibrate groundwater models, head observations in canal dikes were collected and received from seven Dutch regional water authorities, namely Hoogheemraadschap Schieland & de Krimpenerwaard (HHSK), Hoogheemraadschap Delfland (DL), Hoogheemraadschap Rijnland (RL), Hoogheemraadschap Hollands Noorderkwartier (HHNK), Weterksip Fryslân (WF), Waterschap Amstel, Gooi and Vecht (AGV) and Hoogheemraadschap De Stichtse Rijnlanden (HDSR). The crest levels of the dikes with head observations vary from around NAP+1.5 m to NAP–2.5 m (all elevations are relative to the Dutch reference level called Normaal Amsterdams Peil (NAP), which is approximately mean sea level) and polder water levels ranging from NAP–2 m to NAP–6.5 m, highlighting that many canal dikes lie below sea level.

In total, 258 head observations at 91 monitoring sites were collected in this study (see Figure 3.2). Multiple head observations can be located at one monitoring site, where piezometers are aligned within a dike cross-section, e.g. measurements in the crest, inner slope and toe of the dike. At each site, up to five piezometers were installed with filter depths reaching up to approximately 5 m below the surface to measure phreatic head levels. Monitoring sites can be located near each other, e.g. 20 m, while still

exhibiting varying responses due to differences in subsurface materials, highlighting the large spatial heterogeneity. Consequently, the distance between monitoring sites was not used to exclude any observations. The heads were measured with automatic pressure loggers, with hourly measurement intervals in the period between 2006 and 2023 and were resampled to daily-mean values for the analysis. For further analysis, only time series that are longer than 2.5 years were selected. Additionally, any time series exhibiting visual anomalies attributed to failing measurement devices or odd behaviour such as pronounced drift, absence of fluctuations or inexplicable jumps were removed from the dataset. The monitoring sites were also checked for whether the head observations are located in a dike, since sometimes the dike is more like a quay without a slope. In these situations, the dike was not considered for further analyses. The head dynamics in a dike are complex and also location-dependent within the dike profile, since the head in the outer crest can respond differently from head levels in the inner slope. Only head observations that are located in the talud zone or mid-slope of the dike were used (see Figure 3.3). This is the area between the inner crest (the top of the dike at the polder side) and the toe of the dike, where the most variations in groundwater levels are expected. This is because it is farthest from the regulated water levels in the canal and polder, which are maintained at the target levels.

The resulting dataset consists of 108 head time series at 48 monitoring sites, consisting of phreatic head levels measured in the dike body. The length of the head time series varies between 3 and 9 years, with an average length of 5 years.

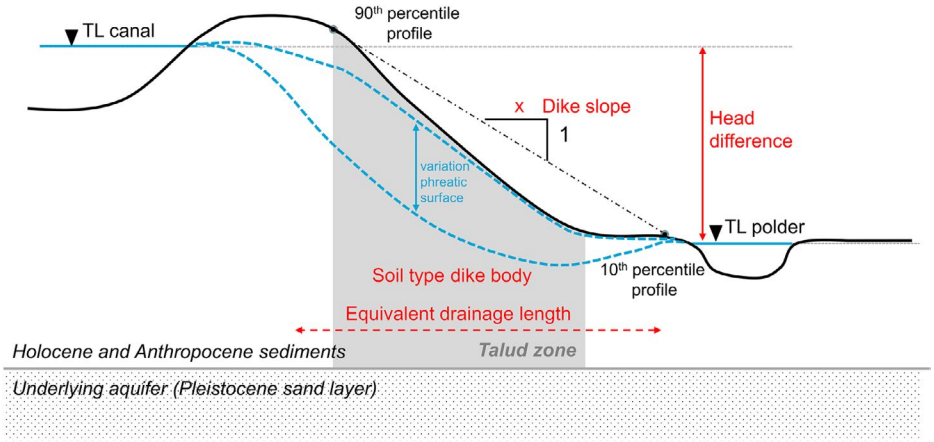


Figure 3.3 A simplified cross-sectional profile of a canal dike with regulated water levels (TL = target level) on both sides (canal and polder) and the variation in the phreatic surface (blue dashed lines). In this study, the groundwater levels in the talud zone of the dike are analysed. The conceptualization of several dike characteristics is highlighted in red, as discussed in Sect. 3.3. The 90th and 10th percentile profile points were derived from the elevation profile as an approximation of the dike slope.

Precipitation and evaporation

The Royal Netherlands Meteorological Institute (KNMI) provides several data products about weather and climate. In this study, two data products for rainfall and evaporation (on daily basis) were used, serving the purpose of (1) getting the best estimate of the historic local weather conditions at the head observation sites (for the calibration of models) and (2) getting long term time series of the weather representing the current and future climate situation in the Netherlands (for extending the head time series). First, the local historic weather conditions were extracted from rainfall and potential evaporation maps of the Netherlands, namely the radar-derived precipitation amounts (Wolters et al., 2013) and inverse distance weighting (IDW) interpolated potential evaporation amounts based on the KNMI ground stations. The evaporation maps give estimates of the daily Makkink reference evapotranspiration derived from ground observations of the global radiation and the average daily temperature (De Bruin, 1987). Secondly, to derive extended head time series encompassing more extreme events, 30 years of precipitation and evaporation time series corresponding to different (current and future) climate scenarios were used (Van Dorland et al., 2023). These time series are a representation of the climate and not an estimate of the actual weather. Since this study focuses on variations in peak heads caused by different head responses, rather than from spatial variations in the nature of the load (dimensions of weather events), the 30-year time series at one location was used for all sites. The station Aalsmeer, close to Amsterdam, was used, which lies rather central in the western Netherlands (see Figure 3.2). In total, nine 30-year time series were used. One time series corresponds to the current climate and eight to future climate scenarios with combinations of two time horizons (2050 and 2100), two greenhouse gas emission pathways and two types of regional climate responses. The emission scenarios include SSP1-2.6 (a low-emission scenario that assumes sustainable development) and SSP5-8.5 (a high-emission scenario that assumes fossil-fuel-intensive development). Each emission scenario is split into a wet-trending and a dry-trending regional response, reflecting the uncertainty in how precipitation patterns may shift in the Netherlands. All future scenarios with different emission levels and regional climate responses predict an increase in winter precipitation and drier summers, accompanied by increased evaporation and reduced precipitation. Although these trends occur across all future scenarios, their intensity varies across different scenarios. These combinations capture a wide range of future scenarios, allowing us to assess the sensitivity of head statistics to climate change across time, emission pathways and regional climate responses.

3.3 Method

This study takes a novel approach by combining a unique nationwide dataset of head observations in canal dikes with time series modelling to investigate how canal dikes respond to heavy rainfall. The approach developed for assessing the dynamics of hydraulic heads is shown in Figure 3.4. After collecting head observations in canal dikes, time series models were set up to simulate hydraulic heads in canal dikes, using precipitation and potential evaporation as the explanatory time series. Several model structures were evaluated, and the one with the overall best performance was selected. Only models that meet the specified reliability criteria (minimum goodness of fit and sufficient time series length compared to model parameters) were selected, resulting in a set of models that explain the fluctuations of the observed head levels in different dikes with a variety of head responses (step 1). These models were forced with 30 years of precipitation and evaporation time series, corresponding to different climate scenarios (current and future conditions). This was done to obtain extended head time series that encompass more extreme events (step 2), facilitating the analysis of both the dynamics and statistics associated with extreme occurrences. The variation in head responses was quantified by analysing the coincidence of the head peaks across canal dikes, selected using the peaks-over-threshold method and classifying different clusters of dikes with similar head responses. The variation and similarity of head responses were related to several physical dike characteristics, like subsurface material and dike profile, to search for explanations of the differences found (step 3). Finally, a generalized Pareto distribution (GPD) is fitted to the head peaks that describe the probability of occurrence of peak values. The variations in and properties of the head statistics are analysed, as well as the impact of climate change (step 4).

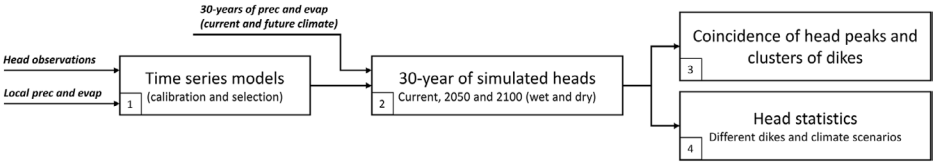


Figure 3.4 An overview of the approach and its accompanying steps.

3.3.1 Groundwater modelling in dikes

Several modelling approaches can be used to model the hydraulic head in canal dikes. Commonly used approaches comprise numerical groundwater models, like Hydrus-2D, PlaxFlow and MODFLOW, that are solutions to (systems of) differential equations that describe the flow of groundwater (Šimůnek et al., 1999; McDonald and Harbaugh, 2003). These approaches need detailed information on material behaviour for both unsaturated and saturated soils. In the case of canal dikes, Van Esch (2012) showed that it remains difficult to reproduce observed hydraulic heads in dikes because of

theuncertain conceptualization of the subsurface, spatial heterogeneity and applied boundary conditions. Time series modelling is a simplified and abstract representation of head fluctuations at one point resulting from the complex 3D movement of water in the dike (Bakker and Schaars, 2019). It is a data-driven approach that can estimate the contribution of independent drivers (rainfall, evaporation, water levels, etc.) on the observed head levels derived exclusively from observed data. This approach is used in this study.

Time series models

The basic principles of time series analysis come from the statistical sciences (Box and Jenkins, 1970). Transfer function noise (TFN) modelling is a subfield within time series analysis that aims to convert one or more input series into an output series using a statistical model. Von Asmuth et al. (2002) presented a novel form of TFN models that relies on the concepts of convolution and predefined impulse response functions and is used for many applications within groundwater science. Predefined response functions are used to estimate the effect of a unit pulse of a driver, like precipitation, on the head response. The head response is simulated through the convolution of various drivers with their response functions. The basic model structure of a TFN model to simulate heads may be written as

$$h(t) = \sum_{m=1}^{M} h_m(t) + d + r(t) \qquad [3.1]$$

where $h(t)$ is the observed heads, $h_m(t)$ is the contribution of drivers m to the head, d is the base elevation of the model and $r(t)$ represents the residuals. The number of drivers M in each model varies based on the selected model structure. The contribution of driver m to the head is computed through convolution:

$$h_m(t) = \int_{-\infty}^{t} S_m(\tau)\theta_m(t-\tau)d\tau \qquad [3.2]$$

where $S_m(\tau)$ is a time series of driver m at preceding time τ, and $\theta_m(t-\tau)$ is the associated impulse response function determining how much of that past driver still influences the head at time t. A variety of impulse response functions can be used to simulate the effects of certain drivers, where commonly used impulse response functions are the scaled gamma and the exponential response functions (Collenteur et al., 2019). The exponential response function is the simplest response function with only two parameters and may be used for drivers that have an immediate effect on the head, like the shallow head levels in the canal dikes (up to a few metres below ground level). Together with the small geohydrological dimensions (the dimensions of the dike are typically only a few tens of metres), a relatively rapid response is expected, which is

confirmed by measurements where head observations respond quickly to rainfall, despite the presence of low-permeability soils.

Various time series model structures

The head fluctuations are explained by the contribution of various hydrological drivers that are convoluted with a response function. Various model structures can be chosen, incorporating different drivers with different response functions. A driver can also be the combination of multiple drivers. The net groundwater recharge $R(t)$ is frequently used as a driver that is derived from rainfall $P(t)$ and Makkink reference evaporation $E_p(t)$ as inputs (e.g. von Asmuth et al., 2008):

$$R(T) = P(t) - fE_p(t) \quad [3.3]$$

where the parameter f is the so-called evaporation factor used to scale the reference evaporation. This model is referred to as a linear recharge model, after which S_m is substituted by the net recharge $R(t)$ and then convoluted with a response function to determine the impact of recharge on the head. In this formulation, processes such as surface runoff are not accounted for but may be relevant for canal dikes and can be incorporated using non-linear models.

Non-linear recharge models, such as those based on soil-water balance concepts, like FlexModel or Berendrecht, offer a way to incorporate additional hydrological processes, like surface runoff. They introduce more complexity and typically require more model parameters (Collenteur et al., 2021; Berendrecht et al., 2006). Furthermore, these models can account for the non-linear response of the head to precipitation and evaporation by using connecting reservoirs, such as interception and root zone reservoirs, which can include short-term water retention in the soil. Another non-linear model structure is the threshold autoregressive self-exciting open-loop (TARSO) model (Knotters and Gooijer, 1999). This structure consists of two regimes (upper and lower), which are separated by a threshold. Each regime has its own exponential response function with corresponding drainage levels, but only when the head reaches the upper drainage level will the upper response function become active. Therefore, this model can be useful when the head response is different above a certain head level. This threshold value is not fixed but is estimated during the calibration process, just like other parameters.

The head time series were modelled using time series models as implemented in the Python open-source package Pastas (version 1.6.0) (Collenteur et al., 2019). Time series models were set up, assuming that the head dynamics in canal dikes are primarily influenced by rainfall and evaporation, and only these two drivers are included in the model. This assumption is supported by the observation that canal water levels

fluctuated minimally (on the order of tens of centimetres) during the measurement period, while the observed average head range was more than 1 m. Additionally, the models demonstrated an overall good fit. In total, four different model structures were employed (see Table 3.1). The model structure with the highest averaged goodness of fit across all models was used for further analysis. Selecting a single model structure ensures consistent comparison across different locations and simplifies the interpretation of results.

Table 3.1 Model structures employed in this study and their characteristics, like type of recharge model, impulse response function and number of fitted parameters.

	Recharge	Impulse response function	Number of fitted parameters
Linear – Exp	Linear	Exponential	4
Linear – Gamma	Linear	Gamma	5
Flex-model	Non-linear	Gamma	7
TARSO-model	Non-linear	Exponential	7

Model calibration and selection

The time series models were used to characterize and simulate the heads of canal dikes with a single deterministic parameter set. For every head time series, time series models were set up, where the full series were used for calibration to maximize data utilization. This was done because the length of the available time series was limited and to avoid the missing information of the head response when splitting up the data for model calibration and model validation. Although validation tests of the models can indicate that the models are performing well and are adequate to achieve good quality model predictions in post-validation model application, previous studies showed that the most robust models are achieved when all data are used for calibration (Shen et al., 2022; Arsenault et al., 2018), which is in line with the goal of this study. Overfitting is mitigated by employing time series models with up to seven parameters and using head calibration data with more than 1000 data points. The model parameters were estimated using the least squares method, employing a warmup period of 10 years and without incorporating a noise model to represent the residuals. After choosing the best model structure, the calibrated models of that structure were evaluated using two criteria. These two criteria were used to determine whether a model is reliable for further analysis:

- Goodness of fit: the model's goodness of fit, measured by the *R* squared (r^2), must be equal to or greater than 0.7 in the calibration period, indicating a minimum acceptable level of fit. This is also known as the coefficient of determination, which is a measure of how well observed outcomes are explained by the model. The sensitivity to threshold selection was tested using threshold values of 0.6 and 0.8. Although the number of reliable models changed, both thresholds resulted in

models with similar peak head responses within comparable limits. Therefore, our findings are robust to the exact choice of threshold.

- Response time: the 95 % response time, the time it takes for 95 % of the influence of an impulse (groundwater recharge) to dissipate, must not exceed the length of the measurement series. Time series should be long enough to cover the head response in order to estimate parameters accurately (Knotters and van Walsum, 1997). This criterion eliminates models for which the time series data are not long enough considering the estimated model parameters.

When there are multiple reliable models of head time series available at one monitoring location, the model that provided the best fit was selected as the representative model for that location.

3.3.2 Peak selection and extreme value analysis

Peaks in hydraulic heads often occur in groups over time: an extremely high hydraulic head is likely to be followed by another since the groundwater system within dikes contains autocorrelation or memory. For extreme value analyses, we are interested in independent peaks to avoid biases and underestimation of the variability of extremes. Therefore, peaks were filtered out of the time series such that the peaks were mutually independent from each other in time by using the peaks-over-threshold (POT) method in combination with a time window. The POT method was applied with a threshold set at the 90th percentile of the analysed head series, and a time window of 30 d was chosen to guarantee the selection of independent peaks. Afterwards, only the n highest peaks were selected, where n corresponds to the length of the time series in years, ensuring that only the most extreme values were included.

Next, a generalized Pareto distribution (GPD) was fitted to the peaks that describes the annual probability of occurrence of peak values. The GPD has three main forms, Type I, Type II and Type III distributions, which differ in the number of parameters and the flexibility of the tail behaviour. The cumulative distribution functions of the GPD are defined by

$$f(x) = \begin{cases} 1 - \left(1 + \left(\xi \frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right), & \xi = 0 \end{cases} \qquad [3.4]$$

where x is the hydraulic head, and the three parameters of the GPD are called the scale (σ), shape (ξ) and location (μ) parameters. When $\xi = 0$, the GPD is equivalent to the exponential distribution. All generalized Pareto distributions were explored in a sensitivity analysis. In the case of peak hydraulic heads, the exponential distribution

was preferred due to its relative simplicity and its suitability for the process, as suggested by visual inspections of the tail of the distribution and the peak values. One key characteristic of the extreme value distribution is the decimate height. It is defined as the increase in head level that occurs when the return period increases by a factor of 10 and is a relevant characteristic in load and dike safety analyses (Wojciechowska, 2015; Schweckendiek, 2014).

The effects of climate change were quantified by estimating the extreme value distribution, or head statistics, of simulated head time series for different time horizons and climate scenarios (including emission scenario and regional response). To compare future scenarios with the current climate, the increases or decreases in head levels at a certain return period were compared, and the so-called probability factor was used. This metric expresses how the frequency of a head level with a 100-year return period is expected to change under climate scenarios. A probability factor of five implies that a head level that occurs, on average, once every 100 years in the current climate occurs once every 20 years under future climate conditions.

3.3.3 Clustering head responses and relating to physical dike characteristics

The relationships between the head responses and several physical dike characteristics were examined. These characteristics include the subsurface material of the dike body, the head difference (water level difference between the target levels of the canal and polder), the dike slope and the equivalent drainage length (see Figure 3.3). Dike slopes can be seen as a measure of the hydraulic gradient within the dike body, which influences the horizontal groundwater flow and the head response. In addition, steep dike slopes can reduce recharge, as they can increase surface runoff that limits the possibility for water to infiltrate into the dike. The slope of a canal dike was not always obvious, since the dikes do not always have a typical geometry, where the profile is irregular and the crest, the berm and the toe of the dike are not clearly recognizable. The dike slope was obtained from an elevation map of the Netherlands with a horizontal resolution of 0.5 m (Actueel Hoogtebestand Nederland, 2022) and was calculated by referencing two percentile points on the elevation profile (10th and 90th percentile), as an approximation of the slope between the crest and the toe of the dike. The target levels of the canal and polder were obtained from the local water authorities, which were used to calculate the head difference. The equivalent drainage length was determined by dividing the head difference by the estimated dike slope, providing an estimate of the horizontal distance over which water is effectively drained from the dike. The subsurface of the dike was based on borehole descriptions or cone penetration tests (CPTs) at the monitoring sites. In the absence of soil investigation, a detailed three-dimensional model (GeoTOP) of the upper 30 m of the subsurface of the Netherlands was used (Stafleu et al., 2012).

Clustering head responses

Differences in peak head behaviours were examined by analysing the coincidence of selected head peaks across all dikes in the 30-year simulated head time series. These simulations were based on 30 years of rainfall and evaporation at the same KNMI station to isolate the effect of differences in peak heads from the spatial variability of weather events. If peak heads at two different dikes occurred on the same day, they were assumed to coincide. By calculating the percentage of coinciding peak levels for each dike pair, a coincidence matrix was formed. This matrix provided a quantitative measure of how often peak heads align across different dikes, indicating their response to similar weather events. Based on this matrix, dike clusters were identified. These clusters consist of dikes where peak heads were driven by similar weather events, while dikes in different clusters experienced peak heads caused by distinct weather events. The clusters are estimated using the *k*-means clustering algorithm (Hartigan and Wong, 1979), where the number of clusters (*k*) has to be given beforehand and is based on the mean Silhouette score of all samples (Rousseeuw, 1987) and the "elbow method", as implemented by Yellowbrick (Bengfort and Bilbro, 2019).

Statistical tests

The relationships between these physical dike characteristics and characteristics of the impulse response functions, as well as clusters of dikes, were examined. Various statistical tests were employed to assess these relationships by calculating the *p* value for different types of variables, both categorical and continuous: the Wald test for comparing two continuous variables, the Wald chi-squared test used for two categorical variables, and the Kruskal–Wallis test for one continuous and one categorical variable. The analyses of relationships with subsurface materials were limited to clayey and peaty dikes, as the dataset includes only one sand dike, and Dutch canal dikes are generally composed of these materials.

3.4 Results

3.4.1 Modelled head responses in canal dikes

Selecting reliable time series models

For each of the 108 head time series across 48 monitoring sites, time series models with various model structures were created, calibrated and evaluated. To illustrate the performance of different model structures, Figure 3.5 gives an example for the monitoring site at Molenlaan (site DL4). The linear models (exponential and gamma response functions) are not able to model the head response for the full range of head levels, as can be seen in the scatterplot. For this location, it appeared that the head response was non-linear. Both the TARSO and Flex models provide a better fit across the entire range of head levels. Notably, the TARSO model shows improved performance in capturing head levels at the most extreme ends of the range.

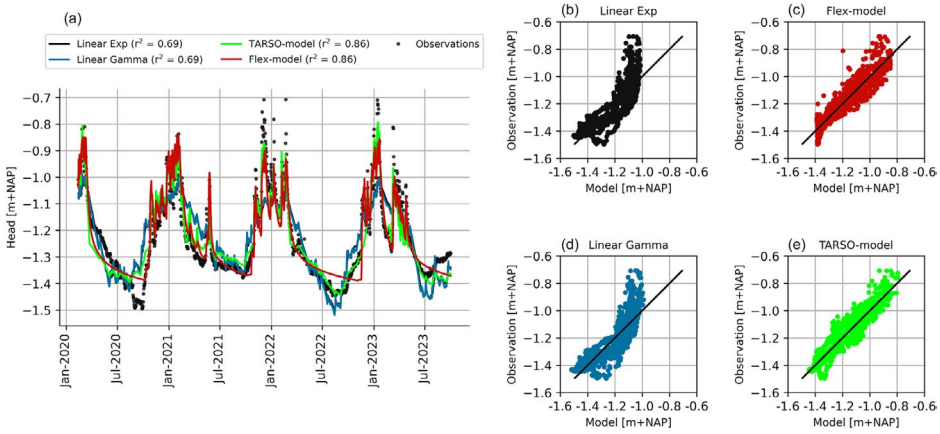


Figure 3.5 The performance of different model structures at Molenlaan (DL4). Panel **(a)** shows the simulated and observed head levels, and **(b–e)** show four scatterplots of the model structures showing the observed and simulated daily head levels (the black line indicates the 1:1 line).

The TARSO model demonstrated the best performance among all the calibrated models. It has the highest average goodness of fit (an average r^2 of 0.74) and performed as the best structure for 81% of the models. The second-best model was the linear recharge model with the gamma response function (an average r^2 of 0.68), while the Flex model structure performed, on average, the worst with an average r^2 of 0.63. However, the Flex model still performed best for 12% of the models, indicating that in some cases a more detailed non-linear representation of recharge processes can be beneficial. Overall, the head dynamics in Dutch canal dikes can be best modelled with a non-linear structure that incorporates two regimes, which are separated by a threshold. This non-linear behaviour can be the result of various soil layers in the dike body, each with distinct hydraulic properties, and changes in infiltration rates or non-constant storage capacities of the unsaturated zone during the dry season (Knotters and de Gooijer, 1999). Next, the calibrated TARSO models were evaluated using the reliability criteria (goodness of fit and response time), and for every monitoring site, the model that met the reliability criteria and had the highest goodness-of-fit score was selected. At 38 out of 48 monitoring sites (79%), models were developed with r^2 of 0.7 or higher. The other reliability criterion, that the 95% response time should not exceed the length of the measurement series, reduced the number of sites with reliable models to 35 (73%). In Figure B1 in Appendix B.1, detailed information is provided on the best-performing TARSO models at each monitoring site, including the r^2 values and corresponding response times. In addition, plots of the measured and simulated heads during the measuring period for all selected monitoring sites are shown in Figure B2 in Appendix B.1. This set of models was used for further analysis.

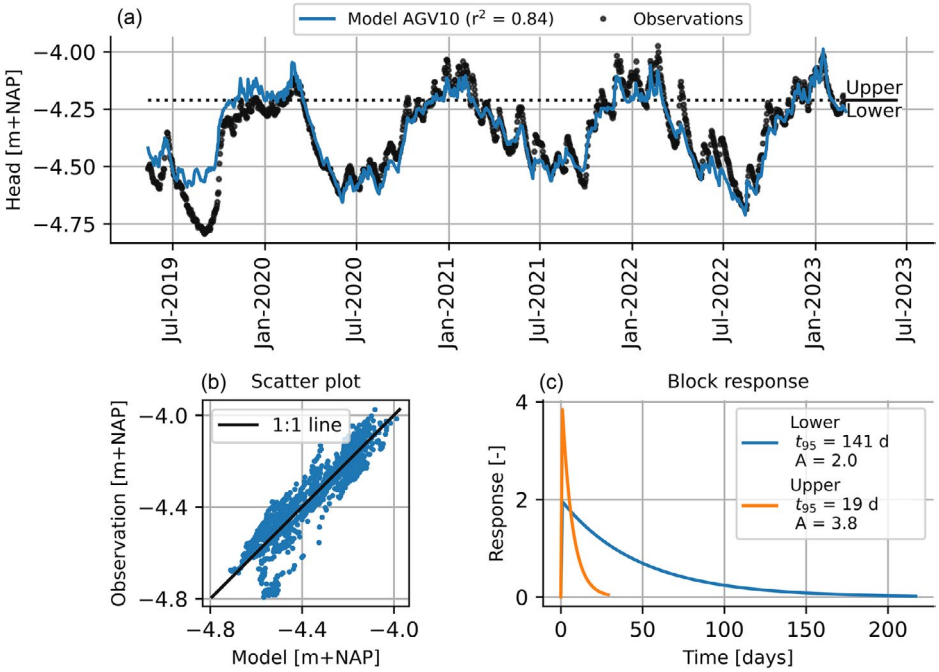


Figure 3.6 Illustration of the simulated head levels with the time series model in comparison with the observations for dike AGV10. Panel **(a)** shows the observations and model series where also the upper and lower regime are indicated with the black dotted line, and panel **(b)** shows the scatterplot of observed and simulated heads. Panel **(c)** shows the impulse response functions of the upper and lower regimes, including the parameters of the peak of the block response (A) and the response time (t_{95}) in the legend.

The model results of one of the selected time series models are shown in Figure 3.6. The simulated heads closely match the observations with an r^2 of 0.84, indicating a close one-to-one relationship between simulations and observations. The model overestimates the heads in the summer of 2019, which was particularly dry in terms of head levels. These differences may be due to inaccurate precipitation data used in the model (as summer precipitation can be highly localized) or disturbances during installation, affecting the head levels in 2019. The other dry summer in 2022 was captured more accurately. The calibrated block response functions for both the upper and lower regimes are shown in the lower right graph in Figure 3.6. These functions are dimensionless and represent the unit head response to a 1 d recharge event of 1 mm. The actual head change is obtained by scaling this response with the recharge input. Two key characteristics of these functions are (1) the peak of the block response and (2) the 95 % response time. The peak of the block response represents the maximum increase in head level that would occur. The 95 % response time, further on referred to as the response time, is a measure of the memory of the groundwater system and, in this

case, represents the time it takes for 95 % of the influence of an impulse (groundwater recharge) to dissipate. In the TARSO models, the upper and lower regimes are separated by a fitted threshold parameter, with head responses differing in each regime. For this model, the head response in the upper regime reacts more strongly to recharge events (higher peak block response) and dissipates more quickly (lower response time) than the lower regime. Whether this behaviour is consistent across all dikes is examined in the next subsection.

Characteristics of impulse response functions

The head dynamics of various dikes were quantified by examining the two key characteristics of the impulse response functions, namely the peaks of the block responses and the response times. Figure 3.7 shows these characteristics for both the upper and lower regime of the TARSO model, where the colours indicate in which water authority region the dikes are located. The colours initially appear random, indicating no clear spatial pattern in the variation of head responses across regions, likely due to the heterogeneous subsoil conditions and other dike characteristics within the canal dike system. The response times of the upper regimes of the dikes are generally short, mostly ranging from about 2 to 50 d, with two exceptions where response times exceed 250 d. These two exceptions have a very small peak of the block response compared to other dikes, which have a large variation in the peak response of the upper regime, with values reaching up to nearly 15. This variation may result from different soil storage capacities and the redistribution of infiltrated water within the dike, which can accumulate in the talud zone causing large head responses. The response times of the lower regime show more variation than those of the upper regime, with most dikes ranging between 100 and 600 d. Meanwhile, the peak block response for the majority is below 5.

Two key patterns of non-linearity appear in nearly all locations when analysing the response functions. First, the response time is longer for the lower regime than for the upper regime (34 out of 35 sites). Second, the peak of the block response is higher in the upper regime than in the lower regime (32 out of 35 sites). These observed differences may be explained by underlying physical processes. For example, longer response times in the lower regime can be caused by head gradients (as a driver of groundwater flow) that depend on the head level itself, the presence of various soil layers with different permeabilities, or the fact that head levels closer to the surface increase the degree of water saturation which affects the hydraulic conductivity and response times in a non-linear way. The lower peak of the block response for the lower regime can be caused by non-linear processes in the unsaturated zone, where lower head levels generally allow more water storage, and root water uptake further increases storage capacity (Berendrecht et al., 2006). In contrast, when head levels are higher, capillary action

draws water upward from the saturated zone, increasing moisture in the unsaturated zone and reducing the amount of air-filled pore space. This limits the potential for additional water storage.

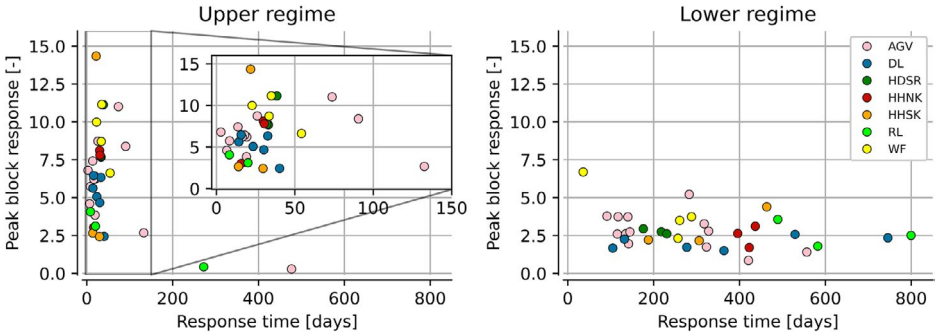


Figure 3.7 The characteristics of the impulse response functions at 33 canal dikes (peak of the block response and 95 % response time) for the upper and lower regimes, where the colours indicate the various water authority regions where the dikes are located. The inset for the upper regime shows a zoomed-in view.

3.4.2 Variation in peak head responses

Coincidence of head peaks and dike clusters

The coincidence matrix, shown in the left graph in Figure 3.8, describes how often peak heads occurred on the same day across different dikes in the 30-year simulated head time series. This illustrates whether peak heads tend to result from the same or different weather events. Based on this matrix, the elbow method identified an optimal number of clusters at $k = 4$, while the Silhouette score indicates that either $k = 2$ or $k = 4$ could be optimal, with scores of 0.442 and 0.439, respectively. Therefore, the number of clusters that are identified using the k-means clustering algorithm was set at four. The resulting clusters of dikes are called clusters A, B, C and D and are indicated within the coincidence matrix; see Figure 3.8. While the three clusters A, B and C are less distinct from each other with still moderately high percentages of coincident peaks, sometimes exceeding 50 %, cluster D has a very distinct peak behaviour. This cluster consists of only two dikes of which the time series models have deviant impulse response functions, longer response times and smaller peak block responses in the upper regime compared to the other clusters (upper right graph in Figure 3.8). In general, the distinctive peak behaviour between clusters is strongly influenced by the response time of the upper regime, with average response times of 16, 32 and 77 d for clusters A, B and C, respectively. This twofold increase in response times for each cluster results in distinct rainfall events leading to peak heads. These longer response times cause peak heads to be driven by more prolonged rainfall events, resulting in peaks that typically occur

later in the winter. More details on the differences in seasonality of peak heads can be found in Appendix B.2, including analyses on the average timing and distribution of peak heads throughout the year.

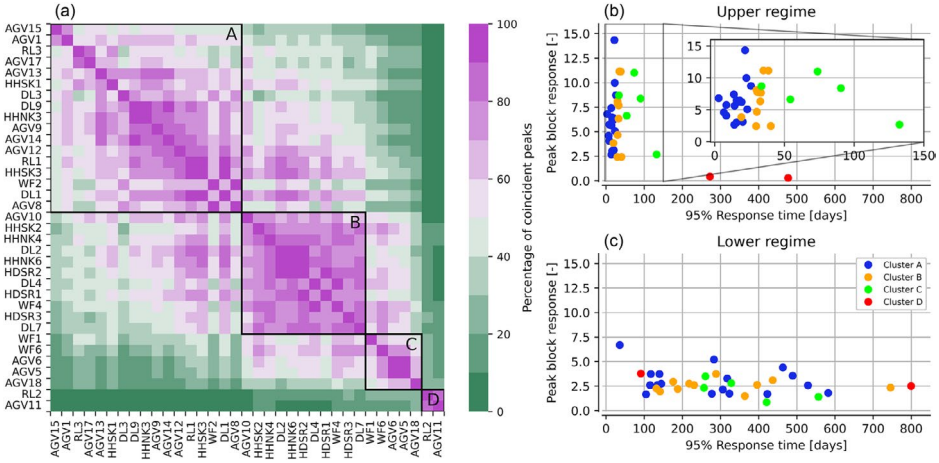


Figure 3.8 The coincidence matrix of the head peaks at the canal dikes, including four identified clusters **(a)**. Within every cluster, the locations are ranked based on the 95% response time. The characteristics of the impulse response functions of the dikes within every cluster are shown in panels **(b)** and **(c)**.

The dike clusters do not exhibit a clear spatial pattern, as shown in Figure B4 in Appendix B.3. In some regions, such as within the water authorities AGV and HDSR, dikes appear to be mainly in clusters A and B, respectively, while in other regions no clear spatial pattern is observed. Although canal dikes within the same polder may have similar dike characteristics, there can still be large spatial variabilities of those characteristics across a region (see Figures B6 to B9 in Appendix B.4). Moreover, even dikes that initially appear to have similar characteristics can exhibit different head responses, which are examined in more detail in the following paragraph.

Clusters in relation to physical dike characteristics

Can differences in peak head responses be explained by physical dike characteristics? Table 3.2 shows the *p* values for the relationships between physical dike characteristics and both the clusters and the characteristics of the impulse response functions (upper and lower regimes), indicating that most relationships were not statistically significant (*p* values > 0.05). However, the subsurface material of the dike appears to play an important role as a distinguishing characteristic in the dike clusters, with a *p* value of 0.03. Statistical tests used to calculate the *p* values can be inappropriate due to violations of test assumptions, small sample sizes, multiple comparisons and data

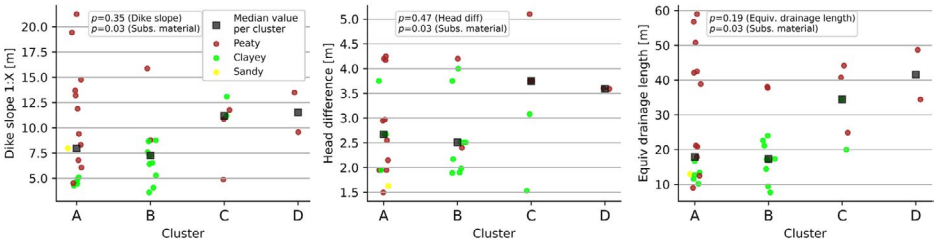


Figure 3.9 The occurrence of dike characteristics across clusters (A, B, C and D), where the colours indicate the subsurface material. The median values per cluster are marked (■).

Table 3.2 The p values of the relationships between the considered physical dike characteristics (rows) and the clusters of dikes, as well as for the 95% response time and peak block response for both the upper and lower regimes (columns).

		Upper regime		Lower regime	
	Clusters	95% response time	Peak block response	95% response time	Peak block response
Subsurface material	0.03	0.32	0.32	0.32	0.32
Equiv. Drainage length	0.19	0.19	0.31	0.22	0.39
Dike slope	0.35	0.64	0.90	0.96	1.00
Head difference	0.47	0.11	0.08	0.08	0.06

dependency, all of which can lead to misleading statistical conclusions (Greenland et al., 2016). Therefore, the physical dike characteristics within every cluster are also visually analysed, and, as expected from the statistical tests, there is a large variation of dike characteristics within every cluster; see Figure 3.9. The clayey and peaty dikes appear in all dike clusters, with most clayey dikes in cluster B (see also Figure B10 in Appendix B.4). Furthermore, small dikes, with drainage lengths less than 20 m and often associated with steep slopes, are found only in clusters A and B, which are clusters with the smallest response times. Furthermore, the median equivalent drainage length of the clusters increases from clusters A to D. This indicates the importance of dike geometry, where shorter distances from a drain (the canal or ditch) lead to faster dike drainage and smaller response times. This can be explained by the fact that the hydraulic gradient, which drives water towards the drain, increases with shorter distances. Yet, determining the cluster to which dikes with certain dike characteristics belong is not straightforward, since all clusters include dikes with various characteristics.

3.4.3 Statistics of head peaks

Head statistics and decimate height

The selected hydraulic head peaks in the 30-year simulated head time series are used to estimate the extreme value distribution of peak head levels by fitting an exponential distribution. The statistics of hydraulic head levels at a dike along the Beemster Polder (HHNK3) are used as an illustration; see Figure 3.10. The decimate height at HHNK3 is approximately 7 cm, while across various dikes the values range from around 5 to 50 cm with a median decimate height of 15 cm (as seen in the right graph in Figure 3.10). Lower decimate heights are found at dikes with smaller peak block responses in the upper regime, in combination with shorter response times. Since these characteristics of impulse response functions do not exhibit a clear relationship with dike characteristics (see Table 3.2) or a distinct spatial pattern, the decimate height also does not follow a spatial pattern, as shown in Figure B11 in Appendix B.5.

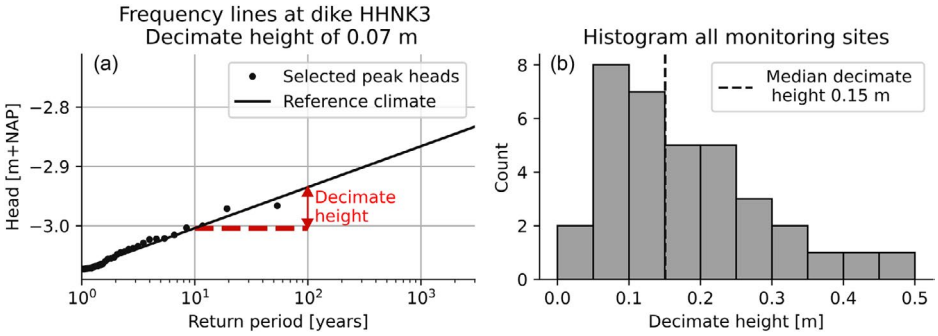


Figure 3.10 (a) Frequency lines of hydraulic head levels at a dike along the Beemster Polder (HHNK3) based on precipitation and evaporation corresponding to the current climate. **(b)** A histogram of the decimate heights of all dikes considered of which the median decimate height is 15 cm.

Impact of climate change

For one location, the resulting frequency lines of various climate scenarios in the year 2100 are shown in Figure 3.11 (left graph). For this location, all climate scenarios result in more extreme head levels or head levels to occur more frequently. In the Hw scenario (high emissions and wet regional climate response), which has the largest increase in winter precipitation according to Van Dorland et al. (2023), the head level at a return period of 100 years increases by 10 cm. Due to the small decimate height, the original head level, occurring once every 1000 years, is expected to happen once every 15 years, indicating a sixfold increase in frequency or a probability factor of six. The probability factors across all dikes, shown in the right graph in Figure 3.11, range from about three times less frequent to seven times more frequent across all climate scenarios in 2100.

This variation could not be directly linked to the clusters of dikes; however, it was found that dikes with longer response times in the lower regime appear to be less impacted by climate change. This can be explained by the fact that these dikes dry out more during drier summers. As a result, a more dried-out dike allows for greater water storage when rainfall returns, causing head peaks to occur less frequently. However, the impact of drier summers on the occurrence of extreme head levels is counterbalanced by wetter winters, which both occur in all climate scenarios. Therefore, the characteristics of both the head response to changing winter precipitation and summer evaporation determine the overall impact of climate change on extreme heads. Under the low-emission scenarios, changes in the frequency of extreme head levels remain small in both 2050 and 2100. This is due to relatively moderate increases in winter precipitation and summer evaporation, which appear to balance each other. Under high-emission scenarios, the impact of climate change in 2050 is, on average, moderate. However, at some dikes, under the dry regional response, the frequency of extreme heads reduced by more than a factor of three. By 2100, however, high-emission scenarios indicate an increase in the frequency of extreme head levels, caused by wetter winters, with the most significant impact observed in the wetting regional response.

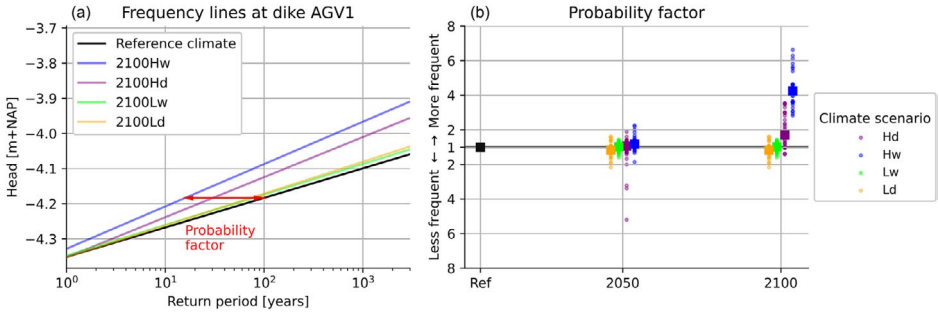


Figure 3.11 (a) The frequency lines of head levels for various climate scenarios in 2100 (High/Low emissions; dry/wet regional climate response) at the dike in a polder south of Amsterdam. **(b)** The probability factor of all dikes (·) and the median value (▪) for different time horizons and climate scenarios.

3.5 Discussion

3.5.1 Limitations and recommendations

Given the extensive canal dike system with thousands of kilometres of canal dikes and the heterogeneity of dike bodies, the number of available observations, both in terms of locations and measurement duration, is limited. As a result, it remains uncertain to what extent this set of 35 reliable models used in this study is representative of the head responses in Dutch canal dikes. In addition to the limited number of observations, two factors further affect representativity. First, the chosen TARSO model structure

influences which head responses are included in the dataset. The reliability criteria applied in this study filtered out observations and associated head responses that could not be modelled adequately, meaning that certain head responses that did not fit the selected models were also excluded. As a consequence, the selected models may not fully capture all relevant processes influencing the head response of canal dikes. For instance, changes in hydraulic conductivity and additional non-linear processes, such as surface runoff from excessive rainfall, are not explicitly accounted for. This limitation also affects the reliability of hydraulic head estimates beyond the measured range of head variations, making them more uncertain. The model uncertainties associated with the time series models were not considered in this study. Furthermore, hypotheses and possible explanations of why the TARSO model fits the head response of canal dikes best require further research. Field measurements and 2D numerical modelling could help improve the understanding of water flow and validate the suitability of TARSO models. Second, it was assumed that each dike has a single head response, as only the best-fitting model within the talud zone was selected to represent that location. However, variations in head responses can exist within the talud zone of a single dike. At 14 monitoring sites, multiple reliable models were available within the talud zone, allowing for an assessment of this variability. At 10 of these sites, the response times of the upper regime differed by only 10 d, indicating similar peak behaviour. Despite the uncertainties surrounding the representativity of this model set, this study demonstrates the value of observations and makes the data publicly available. We hope this will encourage further long-term measurement campaigns to extend the available data and improve the understanding of head responses in canal dikes.

In this study, the *k*-means clustering algorithm was used to cluster dikes based on the coincidence of peak heads. However, two key factors influence clustering outcomes. First, the choice of clustering method affects the results. Various clustering methods exist, and different methods may lead to different clustering results (Everitt et al., 2011). Second, the input data used for clustering determines the outcome. In this study, clustering was based on the percentages of coinciding peak heads across dikes, as this parameter best aligns with our objective of evaluating peak head variability in canal dikes. However, for different purposes, alternative parameters or datasets can be chosen to cluster dikes. While clustering provides a practical way to manage variability, in reality there is no clear distinction or a clear cut between different dikes in terms of head responses and peak behaviour. Instead, it is a gradual shift across a spectrum of head responses.

Next, the variation in head responses and the found clusters were linked to dike characteristics. These dike characteristics themselves were difficult to determine clearly, leading to uncertainty. For instance, the profile of the canal dike can be very irregular and in many cases there is no singular, defined slope. Additionally, canal dikes

are typically made up of multiple soil types, making it difficult to classify a peaty or clayey dike. With all these uncertainties, it is challenging to define clear characteristics for dikes and identify patterns, if not impossible. The uncertainty of the subsurface material of dikes is especially large for those based on the subsurface model GeoTOP rather than borehole descriptions or CPTs. To evaluate the impact of GeoTOP data on the findings, a sensitivity analysis was conducted by testing the effect of excluding this data. While the *p* values changed slightly, the overall results remained consistent, with the *p* value for the relationship between subsurface material and clusters increasing to 0.05, highlighting the importance of subsurface material in the analysis. Furthermore, we examined the relationship between the head response and various dike characteristics, looking at each characteristic separately. It is possible that considering the combination of various characteristics might reveal a clearer pattern. However, it is important to keep in mind that even if there is a relationship between head responses in canal dikes and dike characteristics, it is challenging to determine the local dike characteristics of each individual dike stretch because of the heterogeneous nature of the dike system. This heterogeneity exists in both longitudinal and cross-sectional directions. As a result, it can be expected that the head responses of the canal dikes have large spatial variations and can even differ for dikes that are close to each other.

Lastly, there are limitations to how the impacts of climate change on peak heads were modelled. This study considered the impact of climate change effects only in terms of changes in precipitation and potential evaporation, assuming that the head response itself remains unchanged. However, head responses can change over time due to shifts in hydraulic conductivity and water retention capacity of soils. These changes may result from cyclic wetting and drying of soils, leading to swelling, shrinkage, soil consolidation and alterations in dike structures over time (Stirling et al., 2021; Azizi et al., 2020). As climate change is expected to intensify dry–wet cycles, these processes may become more pronounced, potentially affecting the stability of dikes. Additionally, dike resistance can be further influenced by factors such as shear-strength reduction, soil compaction and peat decomposition. Quantifying all the effects of future climate scenarios is challenging, as both the hydraulic and mechanical behaviours of soils are intertwined and impacted. Translating these effects into slope instability is even more complex, which is beyond the scope of this study. To better understand and quantify changes in head responses within dikes, continuous long-term monitoring, preferably exceeding 10 years, is essential.

3.5.2 Implications for dike safety

This study quantified variations in peak head responses of canal dikes, which are relevant for estimating regional or national flood risk levels in polders. Consider an imaginary canal dike ring along a small polder where each individual dike section is assumed to have a failure probability of 1/100 per year, with fully spatially correlated

load and strength characteristics. The weather conditions across this polder are uniform, with equal precipitation and evaporation everywhere. Under these conditions, the probability of flooding in the polder equals the highest failure probability of the individual dike sections. In contrast, suppose there were four different types of dikes, each with its own peak head responses, resulting in four statistically independent load conditions. If their strength characteristics remain fully spatially correlated, the flood probability in the polder increases to approximately 1/25 per year, which is a factor of four higher than with fully spatial correlated load characteristics. In general, this can be calculated by $P_{f,sys} = 1-(1-P_{f,sect})^n$, where $P_{f,sys}$ is the failure probability of the dike ring (or flood probability of the polder), $P_{f,sect}$ is the failure probability of the individual dike sections and n is the number of statistically independent dike sections. In reality, strength uncertainty is significant and correlation lengths are small, meaning dike sections are not fully correlated in strength, which increases the flood probability of the polder further. The spatial correlations in loads and strength are crucial for accurately estimating flood risk levels in regions, and this study provides valuable insights into the variation of loadings in dikes by considering the different head responses that affect these estimates. Besides variations in head responses, regional risk levels are also affected by natural spatial variability in weather events. This spatial variability can increase the number of statistically independent load events, depending on the scale considered. This topic is discussed in more detail in Chapter 4.

Gariano and Guzzetti (2016) reviewed the literature about the impact of climate change on landslides, both in natural and engineered slopes, and concluded that the risk of shallow landslides can increase (triggered by short and intense rainfall events), while the risk of deep-seated landslides may decrease or show no significant change (related to long rainfall periods). This is primarily due to changing meteorological conditions that lead to higher head levels, reducing the shear strength, soil suction and cohesion and increasing the weight (wet density) of slope materials, all of which contribute to increasing slope instability. Deep-seated landslides appear to decrease or show no significant change, because these types of landslides depend on monthly and/or seasonal rainfall amounts. These more prolonged rainfall events are expected to decrease in regions like the Alps (Rianna et al., 2014; Gariano and Guzzetti, 2016). Canal dikes show similar behaviour to shallow landslides, with, on average, an increased risk of instability in the future. However, the impact of climate change on canal dikes can still vary considerably, with extreme head levels projected to occur either more or less frequently depending on the specific head response. While this study could not identify definitive explanations for these differences, the results suggest that both the head response to changing winter precipitation and summer evaporation play a role in the overall impact of climate change. These insights are relevant for assessing dike safety over time and enhancing adequate dike design.

Another interesting finding for dike safety assessments is the small decimate height in the canal dike head statistics. For dikes with a decimate height of only 5 cm, yearly occurring head levels (T = 1 year) are only 15 cm lower than extreme head levels that occur on average once every 1000 years. Measurements in these dikes are particularly valuable for improving safety assessments, as observed heads are closer to extreme levels, allowing for more accurate extrapolation with fewer uncertainties in modelling extreme levels (Wojciechowska, 2015). Furthermore, short-term measurements can already capture relatively high observed loads, providing valuable insights into the actual dike strength through reliability-updating techniques (Schweckendiek, 2014). Nevertheless, for dikes that are marginally stable, this small increase in head level can still be a trigger for dike failure.

3.6 Conclusions

This study aimed to assess the dynamics of peak heads in Dutch canal dikes at a national level by analysing variations in head responses, head statistics and the impact of climate change. This was done using time series models calibrated on a unique dataset of head observations in dike systems, consisting of 108 head time series across 48 monitoring sites. Various model structures were evaluated, and it was found that the non-linear TARSO model outperformed the other (non-)linear models. This model consists of two regimes (upper and lower), separated by a threshold, each with its own exponential response function and drainage levels. This threshold non-linear behaviour in canal dikes can be attributed to several factors. Groundwater flow is driven by head gradients and hydraulic conductivities, which both can vary vertically and depend on the head level itself, while rising head levels near the surface non-linearly affect hydraulic conductivity and water storage in the unsaturated zone. The TARSO model was selected to calibrate time series models for all head time series, which were then evaluated using two reliability criteria (goodness of fit and response time), resulting in a set of 35 reliable models.

Differences in peak head behaviour between various canal dikes were examined by analysing the coincidence of head peaks across all dikes in 30-year simulated head time series. Four clusters of dikes were identified, consisting of dikes where peak heads were driven by similar weather events, while dikes in different clusters experienced peak heads caused by distinct weather events. The differentiating factor was the response times of the upper regime of these dikes, where longer response times caused peak heads to be driven by more prolonged rainfall events. The identified dike clusters do not exhibit a clear spatial pattern. The reasons are the large spatial variability of dike characteristics and the fact that even dikes with similar characteristics can exhibit different head responses. While the subsurface material and dike width appeared to be important factors influencing the variations in head responses, their presence in

multiple clusters indicates that these characteristics alone do not definitively determine the head response.

Next, peak head statistics were derived across the canal dikes, revealing that the median decimate height is only 15 cm, ranging between 5 and 50 cm. This indicates that yearly occurring head levels are, on average, relatively close to extreme events. Dikes with lower decimate heights were associated with smaller peak block responses and shorter response times in the upper regime. Since these characteristics of impulse response functions do not exhibit a clear relationship with dike characteristics or a distinct spatial pattern, decimate heights also do not follow a spatial pattern. With climate change driving higher winter precipitation and summer evaporation, head statistics are changing. Results showed that head levels with a return period of 100 years are expected to occur about three times less frequently to seven times more frequently by 2100, depending on the climate scenario and the type of canal dike. Dryer summers can reduce the frequency of extreme peak heads by lowering head levels during summer, which increases dike water storage capacity when rainfall returns. However, most climate scenarios project a higher frequency of extreme head levels by 2100, caused by a wetter winter trend. The varying impact of climate change on dikes is largely attributed to the response times of the lower regime. Dikes with longer response times seem to be less affected by climate change, as they experience greater drying during drier summers. However, this increased drying during summer can have other negative consequences, as climate change is expected to intensify dry–wet cycles, potentially leading to soil degradation (Stirling et al., 2021; Azizi et al., 2020).

Code and data availability

The measurement data used to establish the model set, consisting of 108 head time series across 48 monitoring sites and the local historic rainfall and potential evaporation, are available from 4TU.ResearchData at https://doi.org/10.4121/4004f445-b71b-4996-bd7d-10b1fafbc86b (Strijker, 2024). This dataset contains the following:

- An overview of the monitoring sites and piezometers, along with their geographic locations (CSV[1] file and shapefile).
- Hydraulic head time series from the piezometers (CSV file)
- Time series of local precipitation and potential evaporation (CSV file).

Furthermore, the script, output data (models and figures) and other relevant data (e.g. dike characteristics and meteorological time-series data representing future climate scenarios) are shared to enhance the reproducibility of this study.

A readme file is added that describes the files in the dataset. The data source has a CC0 licence, which entails the waiver of all copyright and related rights, enabling unrestricted use of the data for any purpose.

4

Estimating spatial dependencies of peak heads in Dutch canal dikes: a national-scale hindcasting analysis

This chapter examined the length-effect factor of peak hydraulic heads across various spatial scales, like polder scales and the scale of the entire dike system. This factor indicates how much more frequently peak heads occur at larger spatial scales than at the section level. This was estimated by analysing 60 years of hindcast head levels under homogeneous and heterogeneous head-response configurations. For the polder scale, three polders were analysed, with dike lengths of about 15, 40, and 60 km protecting areas of 5, 70, and 180 km^2. At this scale, the length-effect of peak heads primarily arises from variations in head responses, while the spatial variability of weather events contributes moderately, though it increases with larger polders. Spatial weather variability can double the occurrence of extreme peak heads at the polder scale, whereas variation in head responses can raise this by a factor of four to five. At the scale of the entire canal dike system, head response variation has similar effects, but the length-effect of peak heads becomes dominated by the spatial weather variability. It increases with event extremity: from about 6 for yearly occurring heads to around 40 for heads with an exceedance frequency of 1/100 per year. The length-effect of peak heads also varies across head responses: for dikes with shorter response times, shorter precipitation durations become more relevant, accompanied by smaller spatial extents, resulting in larger length-effect factors. The estimated spatial dependencies imply that peak heads with a 1/100 per year exceedance frequency at the section level are expected to occur somewhere in the entire dike system each year, with a probability between 33% and 90%, depending on the variation in head responses.

This chapter is based on a manuscript submitted for publication in Journal of Coastal and Riverine Flood Risk: Strijker, B., Jonkman, S.N. & Kok, M. Estimating spatial dependencies of peak heads in Dutch canal dikes: a national-scale hindcasting analysis.

4.1 Introduction

Canal dikes are located along internal drainage canals and are crucial for draining excess water from the polder and preventing pluvial floods. This type of secondary flood defence is common in polders worldwide (Martín-Antón et al., 2016; Morton & Olson, 2018; Lendering et al., 2018; Triet et al., 2018; Warner et al., 2018). The Netherlands has an extensive system of canal dikes spanning more than 10,000 km dikes, as shown in Figure 4.1. The water levels in the canals can reach several meters above the surrounding polder, making polder areas vulnerable to flooding in the event of canal dike breaches. To estimate the flooding probability of the polder, the canal dike along it is often regarded as a series system of individual elements, also known as dike sections. The spatial dependencies among these sections influence the polder's flood probability (Jongejan et al., 2020; Roscoe et al., 2015). For instance, if multiple dike sections are subjected to loads that act independently, the overall flooding probability in the polder rises because each independent load introduces its own probability of causing failure somewhere. Generally, due to spatial independencies, the longer the dike system, the higher the probability of encountering either an extreme load or a weak spot (i.e. low resistance), and the higher the flood probability of the polder (Vrijling et al., 2011). In dike safety, this phenomenon is also known as the length effect (Kanning, 2012; Roscoe et al., 2020). Accurately assessing this effect requires accounting for spatial dependencies between resistance and loading conditions. This influences the estimation of flood probabilities in polders, as well as on larger spatial scales, such as the national level.

When assessing flood probabilities at the national level, meaning a dike fails somewhere within the entire dike system, the number of dike sections is much larger than at the polder scale, and their spatial dependencies become more important. The flood probability at the national level can be relevant for financial institutions, which are increasingly integrating flood risk into their reporting and regulatory frameworks, including climate stress-testing protocols for bank portfolios (ECB Banking Supervision, 2020; ECB, 2022). In the insurance and reinsurance sectors, risk assessments often cover national and international spatial scales. In the Netherlands, insurers already cover property damage from floods caused by failures of non-primary dikes, such as canal dikes (European Insurance and Occupational Pensions Authority, 2025). To test their financial resilience, they estimate the losses that could occur during a flood with a 200-year return period within specific regions, expressed as a percentage of the total insured value. To assess this risk, insurers estimate how strongly flood losses in different regions are linked, for which spatial dependencies are key for setting up so-called aggregation or correlation matrices.

Furthermore, while failure observations at the section level may be rare, at the system level, more failures are expected, due to the presence of more independent dike sections. By understanding spatial dependencies, failure observations across the entire dike system, or their absence, can be used to verify failure probability estimates at section levels.

The failure probability of the Dutch canal dikes often depends primarily on the failure mechanisms of erosion of the inner-slope by overflow and inner-slope instability (Lendering et al., 2018; Rikkert et al., 2022). The instability of the inner slope can occur when rising hydraulic head levels reduce the effective stress and shear strength of soils, while saturated soils can also increase the driving moment (Van Baars and Van Kempen, 2009; Ridley et al., 2004). As a result, high hydraulic head peaks within the dike body, along with the traffic load, form a primary component of the load affecting the stability of canal dikes (Lendering et al., 2018). The fluctuations in hydraulic heads are primarily driven by precipitation and evaporation, rather than high water levels, and the head response can vary across different canal dikes (Strijker et al., 2024; Strijker & Kok, 2025). As a result, some canal dikes experience extreme hydraulic heads during a heavy one-day precipitation event, while another dike requires more prolonged periods of heavy precipitation. This study focuses on the spatial dependencies of peak hydraulic head levels, influenced by both variation in head responses across canal dikes and spatial variability in weather events. These dependencies increase the probability of a high hydraulic head somewhere in the dike system and make the probability of an extremely high head somewhere along the polder larger than that for any single dike section. The question is: how much larger?

Spatial dependencies can be expressed in several ways, for instance, through spatial correlation lengths, scales of fluctuations, or semi-variograms (Fenton & Griffiths, 2008; Kanning, 2012). Each measure quantifies pairwise relationships. System-scale patterns across a national dike system can be inferred from these measures, but only through further analysis that aggregates pair-wise dependencies. The length-effect factor can be a useful measure that provides more direct insight into how spatial dependencies influence overall risk at different spatial scales. This factor is defined as the ratio between the frequency of events occurring locally, at dike sections, and at larger spatial scales, such as polders or the entire dike system, as shown in Figure 4.2.

Accordingly, we address the following research question: *What is the length-effect factor of peak hydraulic heads at various spatial scales for Dutch canal dikes?*

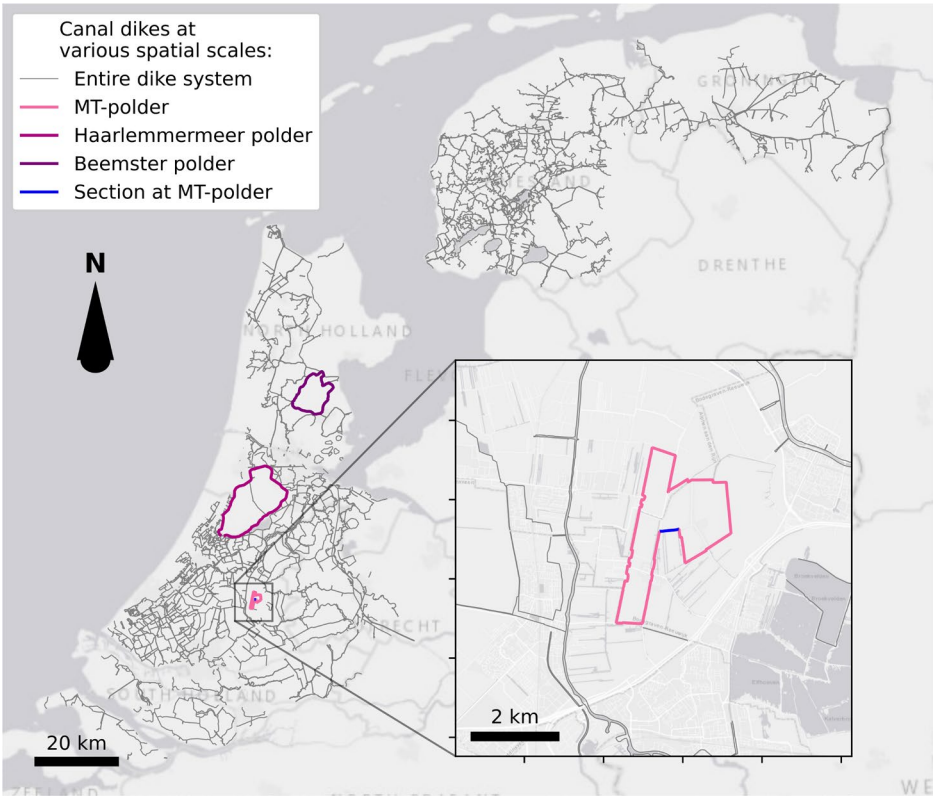


Figure 4.1 Map of the Netherlands showing the canal dike system at different spatial scales: the scale of the entire dike system, the polder scale is illustrated with three polders, namely Beemster polder, Haarlemmermeer polder and MT-polder, and the section level is illustrated with a dike section at the MT-polder.

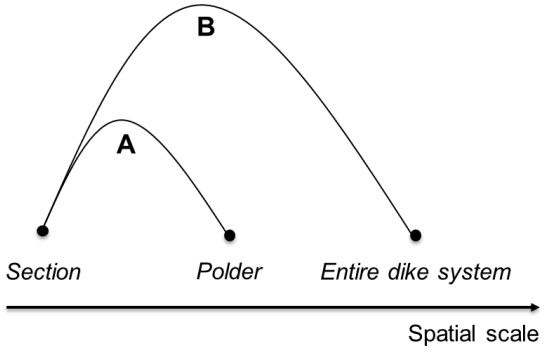


Figure 4.2 Illustration of the length-effect factors examined in this study, linking extreme hydraulic head peaks at the section level to those at the polder scale (A) and scale of the entire dike system(B).

Throughout this chapter, the terms length-effect and length-effect factor refer specifically to those of peak hydraulic heads, unless stated otherwise. Note that in classical dike safety practice, the length-effect typically refers to the increase in failure probability along a dike, rather than the increase in load occurrences as used here.

4.2 Context on system reliability

4.2.1 Accounting for spatial dependencies in dike systems

Estimating the failure probability of a dike section requires accounting for the various possible failure mechanisms, often expressed with a limit state function Z. In this function, the resistance R is typically compared to the load S. Failure occurs when the load exceeds the resistance, meaning $Z=R-S<0$. In probabilistic safety assessments, engineers then consider the probability that the limit state function is exceeded $P(Z<0)$, which is referred to as the probability of failure. To assess the reliability of a dike system, the failure probabilities of its individual elements or dike sections must be combined using system reliability theory. The method for combining these failure probabilities accounts for the correlations between loads and resistances across dike sections. In principle, the failure probability of a dike system, $P(F_{sys})$, will lie between the maximum of the failure probabilities of individual dike sections and approximately their sum, $max(P_i) \leq P(F_{sys}) \leq \Sigma P_i$. These upper and lower bounds already show that if one dike section stands out as the weakest link, its high failure probability dominates the system's failure probability, and the spatial dependencies between sections matter very little. Spatial dependencies become important for the system's failure probability when sections have similar failure probabilities.

Spatial correlations between dike sections exist because both the resistance and load variables can be correlated. For load variables like the water level and wave height, the correlation lengths are expected to be several kilometres (Vrijling & van Gelder, 1998). However, for hydraulic head levels in dikes, little is known about the correlation lengths, especially for canal dikes where extreme heads are driven by heavy precipitation. While load variables between sections are expected to be largely correlated, in contrast, the resistances between sections can be weakly correlated. Especially for the failure mechanism of inner-slope instability, where the resistance depends on soil properties, the correlation lengths in the horizontal direction are typically only tens to hundreds of metres (Phoon & Kulhawy, 1999; de Gast, 2020). Therefore, for Dutch primary flood defences, the length of independent equivalent stretches is often set at 50 metres (Jongejan et al., 2020). This length is based on fluctuation patterns of soil properties using random field models (Vrouwenvelder & Calle, 2003), together with historically observed failure lengths. When the resistances of dike sections are treated as independent, the question arises: do load dependencies between dike sections matter?

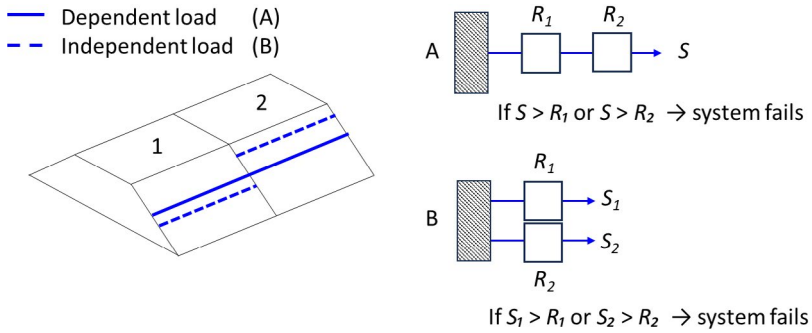


Figure 4.3 Simplified example of a dike system with two sections, each having independent resistances, subjected to a dependent load (A) and an independent load (B). Figure adapted from Kanning (2012).

To answer this question, a simple case is considered of a dike system with two dike sections, sections A and B, see Figure 4.3. Both sections are assumed to have identical resistance distributions and identical load distributions, although the resistance and load distributions can differ from each other. This setup clearly demonstrates how dependencies between loads affect the overall system reliability. Both resistances and loads are modelled with Gumbel distributions: resistance with the right-tailed form and load with the left-tailed form. In this distribution, the location parameter defines the centre of the distribution, while the scale parameter determines the spread; higher values correspond to larger variability. The resistances are assumed to be independent, while the load dependencies are varied and modelled using a Gaussian copula. The analysis examines how different degrees of load variability affect system reliability while keeping the resistance constant. The location parameter of the load distributions is set so that each section has a failure probability of 1/100 per year. The ratio between system failure probability and section failure probability (P_{sys} / $P_{section}$) is used to quantify the effect of load correlation. A system ratio close to two indicates independent sections, while a ratio closer to one indicates dependent sections.

When the scale parameter of the load is small, corresponding to small load variability, a high correlation between loads has little influence on system reliability, and the system ratios are close to 2, as can be seen in Figure 4.4. In these cases, the uncertainty in resistance dominates the system failure probability, and failures occur mainly when resistance is low, while variations in the load have little influence (see top-right graph in Figure 4.4). As load variability increases, its correlation begins to influence the system reliability. When both loads and resistances are independent (ρ_S and ρ_R are 0), the system again behaves logically as two independent sections. However, as the load variability increases, the load begins to dominate the failure behaviour: failures occur mainly under high load levels, while variations in the resistance have little effect

(see bottom-right graph in Figure 4.4). Consequently, load correlation becomes an important factor influencing system reliability, with the system ratio approaching 1 when the loads are fully correlated. This simple example demonstrates that, for dike systems in which dike-section resistances are independent, load dependencies between sections become relevant when load variability is large relative to the dike sections' strengths and dominate the failure probability.

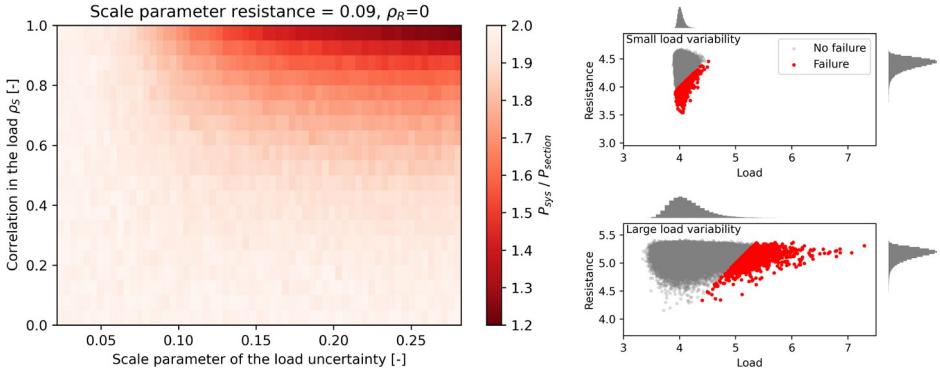


Figure 4.4 System reliability expressed as the ratio $P_{sys}/P_{section}$ for two identical dike sections. The left panel shows how this ratio varies with load correlation (vertical axis) and load variability (horizontal axis, defined by the scale parameter). The right panels show Monte Carlo samples of resistance and loads for both sections, where the top graph corresponds to small load variability, and the bottom to large load variability.

4.2.2 Expected failure observations in dike systems

Longer dike systems are more likely to fail somewhere along their length, due to the independency between sections. If each dike section has the same annual failure probability $p_{section}$, and behaves independently, the system probability of at least one failure is

$$P_{sys} = 1 - (1 - p_{section})^n \quad [4.1]$$

Where n is the number of equivalent independent reaches, which is the total length of the system divided by the length of independent equivalent dike sections. Interestingly, calculated system probabilities appear to exceed the observed rates of dike failure, for example, in New Orleans or Sacramento (Christian & Baecher, 2011). This same pattern appears for the Dutch Canal dikes.

In the Rijnland water authority area, about 300 kilometres of canal dikes are assessed as not meeting the safety standard with a maximum allowable annual failure probability of about 1/500 per year. The primary failure mechanism is the instability of the inner slope caused by heavy precipitation. Suppose we take $p_{section}$ =1/500 and use equivalent

independent lengths of 2,000 m and 200 m (both conservative for soil-property correlation), the yearly system-failure probabilities should be about 26% ($n \approx 150$) to 95% ($n \approx 1{,}500$). Yet no inner-slope instability, due to heavy precipitation, has been observed in the past 60 years. This gap likely reflects 1) section-level probabilities that are too high and/or 2) different spatial dependencies and equivalent independent dike sections length that reduces the effective number of independent reaches. One factor contributing to this gap is how uncertainties in soil properties are addressed.

For inner-slope instability, the soil properties of the dikes are often the dominant source of the resistance uncertainty in the calculated failure probability (Schweckendiek et al., 2017; Zhang et al., 2011). These properties are practically time-invariant and spatially variable uncertainties. Spatial variability is commonly attributed to geological processes occurring in the deposition of soil layers (Phoon & Kulhawy, 1999; Vanmarcke, 1983). At the section level, local site-specific data are often insufficient for meaningful statistical analyses and reliable estimates of geotechnical parameters used in slope stability calculations. Therefore, the parameters and their uncertainties are typically derived from soil data coming from a larger region. While each dike section has a specific deterministic resistance, its value is unknown due to insufficient information or knowledge.

Because of these knowledge uncertainties, the calculated system probability may not match the observed failure rates. This occurs because weaker soil properties dominate the calculated failure probability at the section levels, while stronger properties have no influence, even though both exist in reality. Accounting for this knowledge uncertainty affects not only the section failure probabilities but also the spatial dependencies between sections. In the hypothetical case where the resistance of each dike section is known exactly, the section failure probabilities would be attributable solely to load variability. The load varies over time and is inherently uncertain, and can practically never be reduced to zero. Consequently, the load variability will dominate the failure probability, and the dependencies between loads would become the key factor for the system reliability, as also shown in the example in Section 4.2.1. Although this is a hypothetical scenario, it highlights that spatial load dependencies play a key role in explaining observed failures.

4.2.3 Spatial variability and spatial dependency

When assessing the safety of a dike system, the probabilities of hydraulic head occurrences within dike sections need to be quantified. This probability can involve both inherent and epistemic uncertainties. For example, when piezometer measurements are available for a particular section, the T10 and T100 head levels can be estimated by setting up a groundwater model and simulating extreme conditions. This estimate

may include model uncertainties, but in Dutch dike safety practice, the larger source of uncertainty comes from the scarcity of piezometer data. Large spatial variability in soil properties makes estimating head levels in dikes without measurements highly uncertain. However, although hydraulic head levels can continuously fluctuate in space due to natural variability in the subsurface and discontinuous fluctuations due to local anomalies, they can still be spatially dependent. This is because hydraulic heads may respond to the same external forcing, like water level or precipitation, meaning that both heads will tend to be high at the same time. On the other hand, they can also be spatially independent, when the head responses of dikes differ, and extreme heads occur separately from each other. For example, when sections respond to precipitation events of different durations, extreme heads peak at different times. This is illustrated in Figure 4.5, which shows the T100 head level along a polder dike for a hypothetical situation. The spatial variability of the subsurface results in fluctuations in absolute head level along the dike, making it challenging to estimate extreme head levels at location B based on measurements at location A. However, the head levels at these two locations may still respond to the same driver, making them spatially dependent. This study does not focus on spatial variability, but on the spatial dependency between peak heads. Therefore, the interest lies in differences in head responses and the spatial dimensions of the underlying drivers, and not the absolute level of the hydraulic heads.

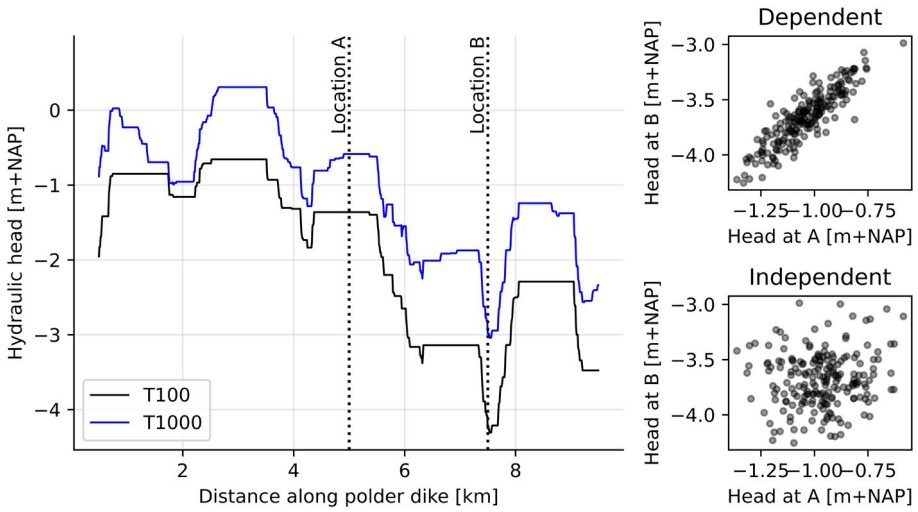


Figure 4.5 An illustrative and hypothetical example of hydraulic heads along a polder dike for return periods of 100 years (T100) and 1,000 years (T1000). The spatial variation in head levels reflects differences in local subsurface conditions, while the underlying drivers can be similar or different, making head levels either spatially dependent (upper right) or independent (bottom right).

4.3 Methodology and data

4.3.1 General approach

Step-by-step approach

To estimate the length-effect factor of peak hydraulic heads at different spatial scales, historical hydraulic head levels were hindcasted across the entire Dutch canal dike system, and the occurrence of peak heads was analysed. By analysing when and where these peaks occur, as well as their extremity, the spatial dependencies within the dike system can be quantified. The main steps are outlined below and discussed in more detail in the following sections:

1. **Hindcasting head levels**: The Dutch canal dike system was gridded to a 1km by 1km raster, assuming uniform meteorological conditions within each cell. For each cell, hindcasts of heads were generated from 1962 to 2022 (60 years) using time-series models driven by historical meteorological data.
2. **Selecting temporal independent peaks of head levels and estimating exceedance frequencies**: For each cell, independent peaks were selected using the Peaks-Over-Threshold method. The exceedance frequencies of these peaks were then estimated using head statistics derived from 100-year head simulations, representing the current climate and various rainfall regime.
3. **Estimating the length-effect factor:** The length-effect factor was determined by counting the number of peak head occurrences at various spatial scales. For example, how many peaks with an exceedance frequency of 1/10 per year occur within a single cell along a polder dike, and how many occur across the cells covering the polder dike? The ratio between these counts defines the length-effect factor, which was estimated for several exceedance frequencies.

Assumptions on head responses

Since head responses are only measured at a limited number of monitoring sites, the head levels across the canal dike system were modelled using a set of possible head responses. This model set was based on a study that examined variations in peak head responses in canal dikes using time-series models calibrated on a unique dataset of head observations (Strijker & Kok, 2025). For each cell, a whole range of possible head responses were modelled, after which, in estimating the length-effect factor, it was assumed that within each raster cell, the head responses are homogeneous or heterogeneous. These are the two head response configurations used and can be seen as upper and lower limits of the actual length-effect factor. For the homogeneous response, all dikes within each raster cell have the same head response. It is still unknown which possible head response occurs, therefore, the average and the variation of resulting from various head responses are shown in the results. In contrast, in the heterogeneous case, dikes within the same cell can exhibit different responses, reflecting the heterogeneous character of the dikes, as illustrated in Figure 4.6. In

reality, the true behaviour of the canal dike system is expected to fall between these two conceptual extremes

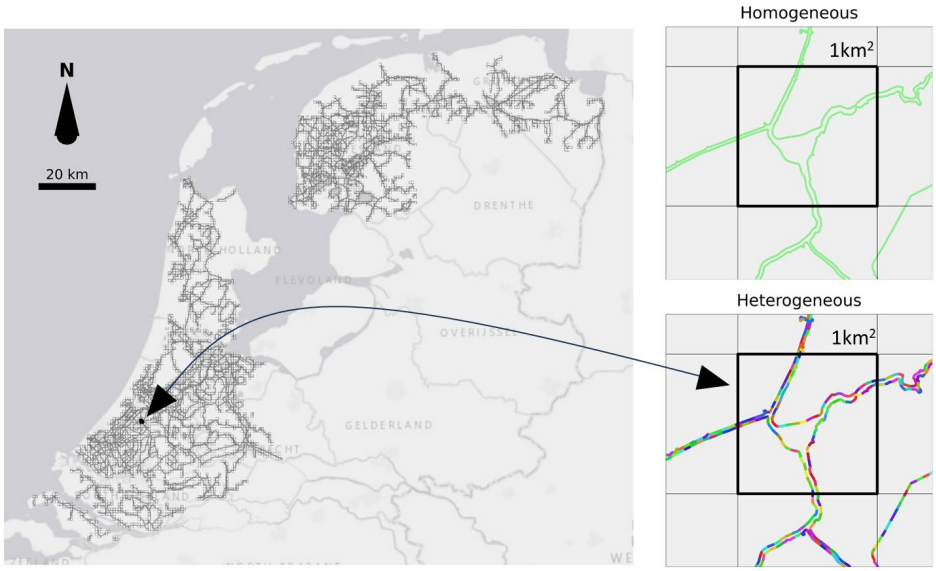


Figure 4.6 The dike system is divided into 1 km × 1 km cells, where the head response within each cell may be homogeneous or heterogeneous, as illustrated on the right.

The following sections describe in more detail how head levels were hindcasted using time-series models and historical meteorological data, followed by the methods used to select peak heads, estimate exceedance frequencies, and determine the length-effect at different spatial scales.

4.3.2 Hindcasting head levels

Time series modelling

Time series models can describe head fluctuations at a single point as a result of complex 3D groundwater movements (Bakker & Schaars, 2019). This data-driven approach quantifies the contributions of independent drivers, such as rainfall and evaporation, to observed heads by considering head observations and the relevant drivers. Von Asmuth et al. (2002) introduced a novel form of Transfer Function Noise (TFN) models that rely on the concepts of convolution and predefined impulse response functions (IRF). This approach has become widely applied in groundwater studies and is implemented in the open-source Python package *Pastas* (Collenteur et al., 2019).

Strijker & Kok (2025) examined the dynamics of peak heads in canal dikes across the Netherlands using time-series models calibrated on a unique national dataset of

head observations. Head observations were collected across the Netherlands, and those measured in the talud zone or mid-slope of the dike were analysed to represent that dike's characteristic head response. Several model structures were tested, and the Threshold AutoRegressive Self-exciting Open-loop (TARSO) model (Knotters and Gooijer, 1999) was found to perform the best. The TARSO model has a nonlinear structure composed of two regimes divided by a threshold. Within each regime, the head responds in its own characteristic manner. In total, 35 time-series models were developed, which represent the diverse head responses observed across the canal dike system. This set was used to hindcast hydraulic head levels across the entire dike system based on historical observed meteorological conditions.

Meteorological data and hindcasting

For hindcasting the hydraulic head, historical meteorological data on precipitation and evaporation were used. These datasets include gridded data on daily precipitation sums and daily Makkink evaporation for the Netherlands covering the period 1910-2022 (Pezij & Lugt, 2023). The precipitation data were based on observations from stations and gridded using ordinary kriging. The number of observation stations varied over time, namely 102 sites for the time period 1910-1951 and 260 sites for the period 1951-2012. The gridded Makkink evaporation data are based on fewer observation stations, namely 5 sites for the period 1910-1981 and 35 sites for the period 1981-2015, although less spatial variability is expected for the Makkink evaporation. The interpolation method used for the evaporation is Thin Plate Spline Interpolation. Only meteorological data from 1950 to 2022 were used to simulate head levels, with a 12-year warm-up period. Consequently, hindcasted head levels between 1962 and 2022 were analysed, covering a 60-year period.

Considering only the gridded meteorological data already provides an indication of spatial dependency. For example, by counting the number of days on which daily precipitation exceeded 44 mm along a selected line, see Figure 4.7. It was found that this threshold was locally exceeded on average nine times, but 71 times along the entire line. Therefore, this precipitation amount occurs about a factor of eight more frequently at the system level (along the line) than at the section level (at one point along the line). Furthermore, the spatial patterns along the line where this precipitation amount is exceeded vary widely, from about half of the line to only small local areas. The heavy precipitation events that cause these exceedances often cover large regions, but the highest rainfall amounts may occur only locally.

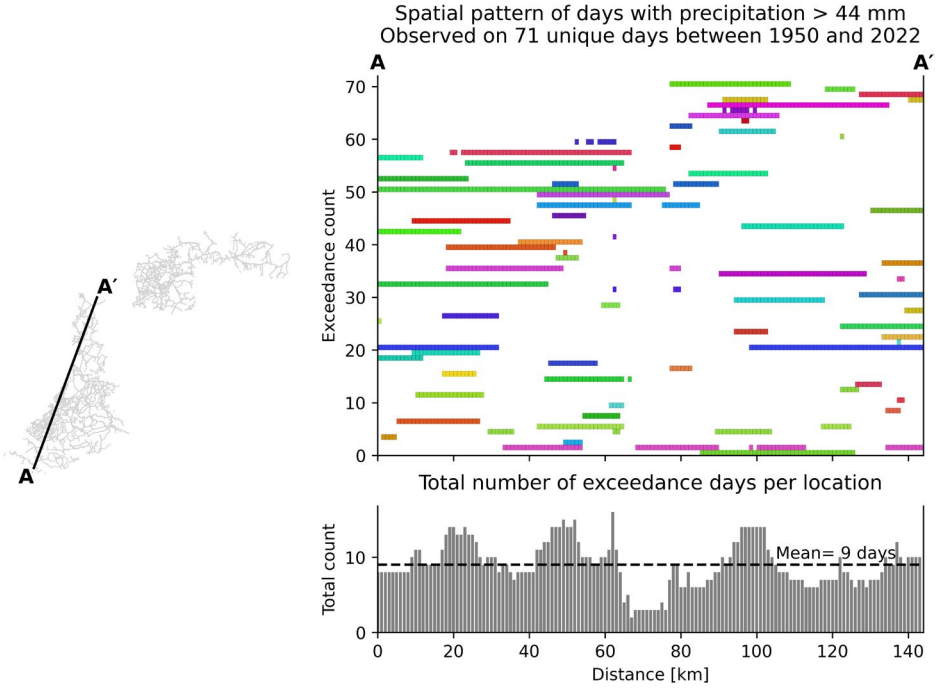


Figure 4.7 Overview of the analysed line AA', with in the background the Dutch canal dike system (left), and the number of exceedances of heavy precipitation along this line (right). The bar heights indicate how often the daily precipitation exceeded 44 mm during the period 1950–2022, where the random colors indicate the unique dates at which this occurred.

4.3.3 Estimating spatial dependences

Selecting independent peaks and estimating exceedance frequencies

To compute the length-effect factor, peak heads were first selected from each hindcasted head time series across all raster cells. The selection of independent peaks was done using the Peaks Over Threshold (POT) method, applying a minimum 45-day separation between peaks to ensure independence. The threshold was determined in two steps. First, an initial threshold was set at the 80th percentile of the head time series. Then, a second threshold was applied to retain only peaks with an exceedance frequency of once per year. For example, in a 100-year time series, only the 100 highest peaks were retained.

The exceedance frequencies of peak head for different head responses were derived from 100 years of simulated precipitation and evaporation, representing the current climate and various rainfall regimes (STOWA, 2024; STOWA, 2019). These rainfall regimes reflect regional differences in the frequency of extreme precipitation across the Netherlands. For example, the western part of the Netherlands experiences more extreme precipitation

than the Northern part, as illustrated in Appendix C.1, which shows a map of the different rainfall regimes. Next, a Generalised Pareto Distribution (GPD) was fitted to the selected peaks. An example is shown in Figure 4.8. The lower panel displays the 100-year simulated head time series for the reference rainfall regime, with peaks identified using POT method. The upper panel shows the exceedance frequencies of these peaks based on plotting positions, together with the fitted GPD. The fitted distributions for all head responses and various rainfall regimes are provided in Appendix C.2.

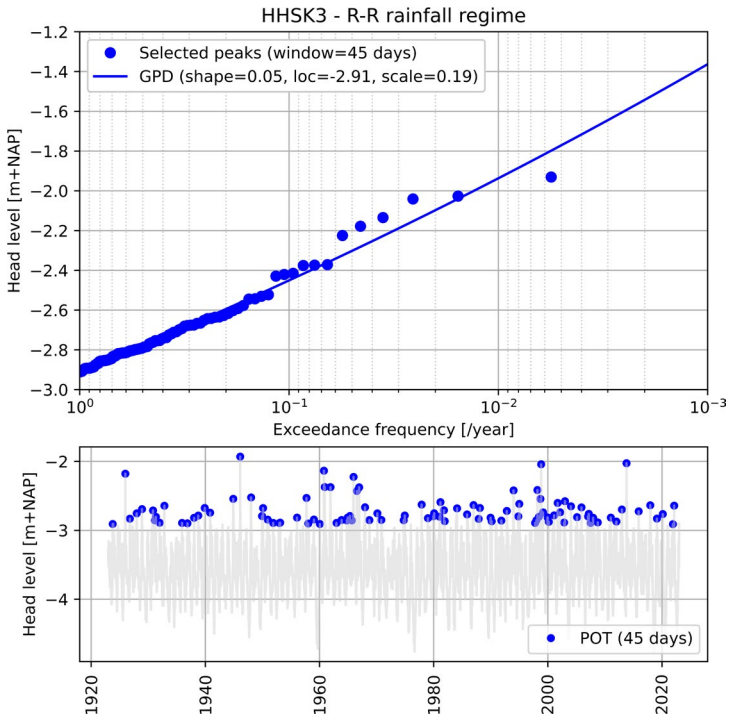


Figure 4.8 Example of selected peaks and fitted GPD.

At various spatial scales, selecting independent events involves both temporal and spatial components, as peaks at different locations can occur on different days but still result from the same weather event that propagates across the country. Therefore, a time delay of ±3 days was applied to head peaks across the dike system to treat them as one independent peak event.

Calculating the length-effect factor

The length effect is defined as the increase of the probability of failure with the increasing length of a dike, due to imperfect correlations and/or independence

between cross-sections and/or sections (Schweckendiek et al., 2013). To express this independence, the length effect factor (*LE*) can be used, which is the ratio between the probability of system failure and the probability of failure for a cross-section (Vrijling et al., 2011). It is often used to determine safety targets for dike cross sections derived from the acceptable probability of failure of the dike (sub)system. In addition to failure probabilities, the length-effect factor can also be applied to loads. For loads, it provides insight into the number of independent load occurrences within the system and indicates how much more frequently peak heads occur at the system level compared to the section level. It is defined as the ratio between the exceedance frequencies of a particular load level at two spatial scales, such as the section level and system level:

$$LE_m = \frac{F_{m,sys}}{F_{m,section}} \quad [4.2]$$

where m denotes the exceedance frequency of a particular load level at the section level. $F_{m,sys}$ and $F_{m,section}$ are the corresponding exceedance frequencies of that particular load level at the system and section level, respectively. These frequencies can be estimated empirically by counting how often the load level $x(F_m)$ is exceeded in time series at both spatial levels:

$$LE_m = \frac{\#\{X_i > x(F_m)\}_{sys} \;/\; T_{per}}{\#\{X_i > x(F_m)\}_{section} \;/\; T_{per}} = \frac{\#\{X_i > x(F_m)\}_{sys}}{\#\{X_i > x(F_m)\}_{section}} \quad [4.3]$$

Where X_i are the observed load values in the time series, #{▪} denotes the number of exceedances and T_{per} is the length of the observation period.

By counting the exceedances of hindcasted head levels at the section and system levels for various frequencies, a graph such as Figure 4.9 can be constructed. The blue and red lines show the number of exceedances at the section and system levels, respectively, as a function of the exceedance frequency of the head levels. The length effect factor is the ratio between these two lines, as indicated by the black arrow. At the section level, the number of exceedances is expected to be close to the theoretical expectation, according to probability theory, $E(X)=F_m T_{per}$, where F_m is the exceedance frequency and T_{per} the time period. For example, a load level with an exceedance frequency of 1/10 per year is expected to occur six times in a 60-year period. This assumes a Poisson process, where events happen randomly and independently over time at a constant average rate. For two independent sections, the expected number of exceedances doubles, since there are twice as many independent trials. In reality, deviations can occur due to statistical uncertainty and non-homogeneous time series. For instance, precipitation amounts that occur on average once every ten years today were more extreme in the past, due to the colder climate.

The length effect can vary with exceedance frequencies, as illustrated in Figure 4.9. For example, $LE_{1/50}$ can be larger than $LE_{1/2}$. Higher-intensity precipitation events are typically accompanied by larger precipitation areas (Lochbihler et al., 2017). However, the spatial extent of the highest intensities can actually decrease. For example, for both extreme river flows and precipitation in Great Britain, Keef et al. (2009) showed that events with increasing return periods become more spatially localised. This decrease in spatial dependence was attributed to the more localised character of heavy events.

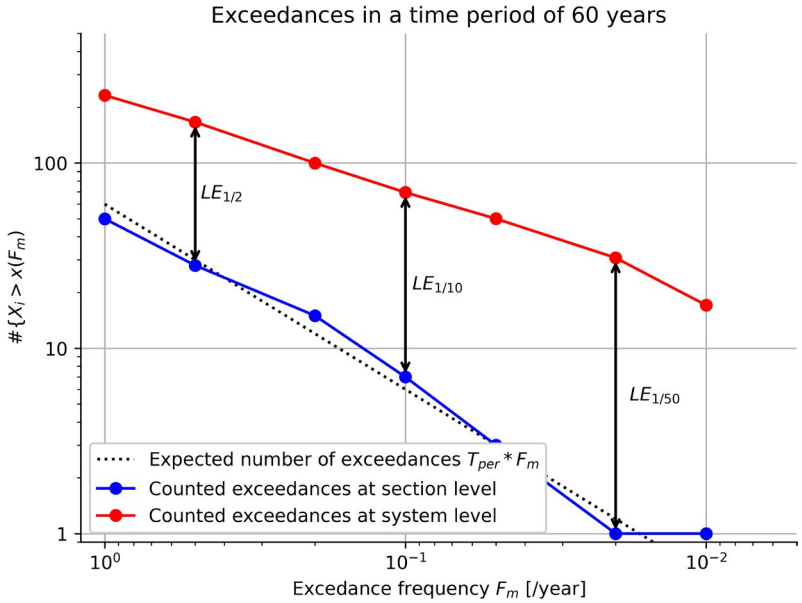


Figure 4.9 An example of empirical exceedance counts for 60 years of simulated annual maxima. The black dotted line shows the expected number of exceedances N F_m. The blue and red symbols give exceedance counts at the section level and the system level, respectively. The length-effect factor can be derived at different exceedance frequencies; examples for 0.1 and 0.5 are highlighted with arrows and formulas.

4.4 Results

This section presents the results of this study. First, the spatial dependencies are explored in Section 4.4.1 through analyses of the underlying data and the time series models used. For example, the influence of variation in head response characteristics on the length-effect of peak heads is quantified, and the spatial patterns of extreme peak heads and corresponding weather events are analysed. Next, the length-effect factors at various spatial scales, namely at the polder scale and the scale of the entire dike system, are estimated.

4.4.1 *Exploring spatial dependencies of peak heads*

Influence of head response characteristics on the length-effect

The length-effect of peak heads arises from the natural spatial variability of weather events and from variations in head responses. The length effect arising from variations in head responses across the 35 time-series models is quantified by analysing peak exceedances within the 100-year simulated head levels used to derive head statistics. Using head statistics, the number of exceedances is counted for various head levels, corresponding to exceedance frequencies from once per year to 1/100 per year. It can be seen that the number of exceedances lies close to the expected number based on probability theory, as indicated by the black solid line with markers in Figure 4.10. This was expected, as the head statistics were based on the same 100-year simulated head levels. However, there is some spreading across the head responses, which reflects the statistical uncertainty in the tail of the distribution. The average number of exceedances across the various head responses is referred to in this study as the homogeneous response configuration.

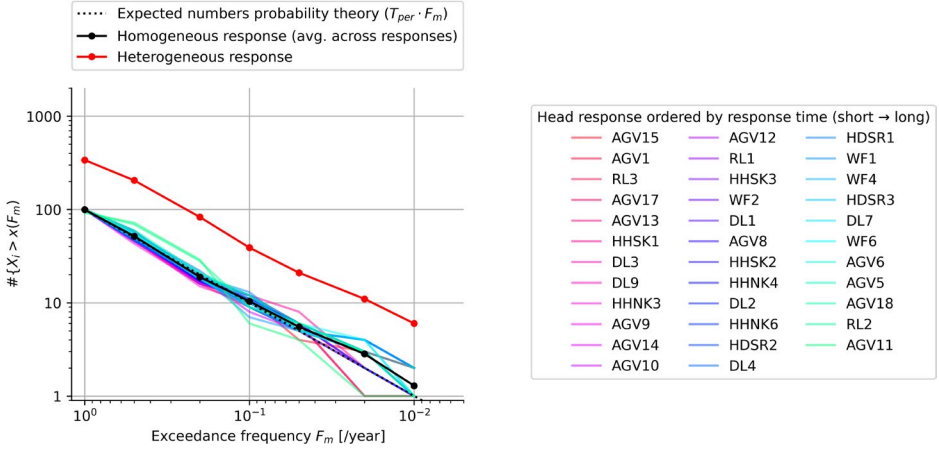


Figure 4.10 Exceedance counts from 100-year head simulations compared with the expected number under a Poisson process, for heterogeneous and homogeneous response.

When the exceedances across all 35 head simulations are counted, treating peaks occurring within three days of each other as a single exceedance, the number of exceedances does not increase by a factor of 35, because many head responses are correlated. In fact, the number of exceedances increases by roughly a factor of four to five. For example, head levels with an exceedance frequency of 1/10 per year occur on average 10 times in a single series but occur at 40 distinct moments when all series are combined. The total exceedance count across all head responses is denoted here as the heterogeneous response configuration. This factor four to five represents approximately

the length-effect arising from variations in head responses, and aligns with the four dike clusters identified by Strijker & Kok (2025).

One key characteristic influencing the extent to which peak heads occur simultaneously across dikes is the response time (Strijker & Kok, 2025). In the case of peak heads, it is the time required for 95% of the effect of a precipitation impulse to dissipate or drain away. Dikes with short response times show rapid drops in head levels after precipitation events. In contrast, dikes with long response times dissipate more slowly, leading to high head levels that persist for several days. Figure 4.11 illustrates these contrasting behaviours by zooming in on the highest head peak within 100-year simulated heads from five time-series models. From the lowest line to the highest, rapid-draining dikes with short response times show sharp peaks that follow heavy daily precipitation, while slow-draining dikes with longer response times exhibit broader and more persistent head elevations driven by a prolonged period of precipitation.

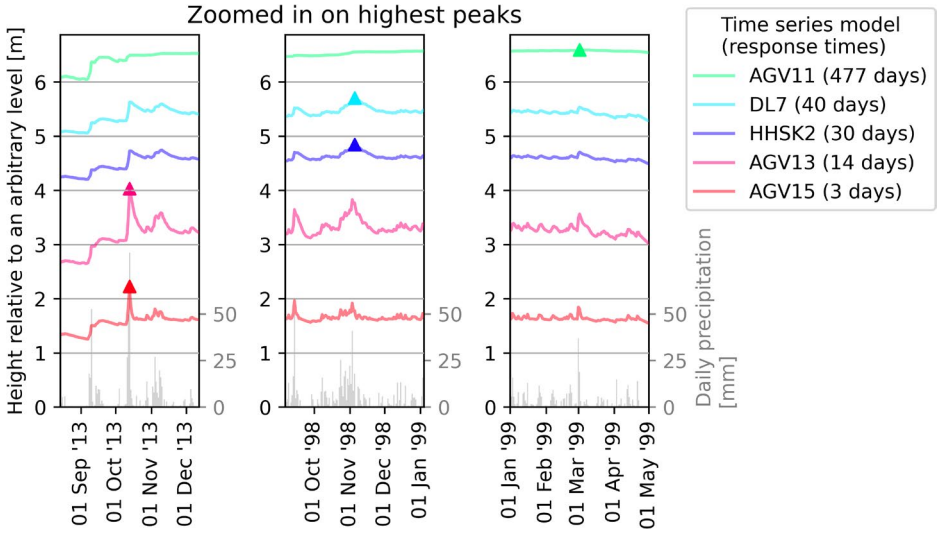


Figure 4.11 Periods in which the highest peaks occur for five different time series models with various response times. The highest peak in every head series is indicated by a triangle (▲). Heights are shown relative to an arbitrary reference level, with the series stacked above one another.

Spatial patterns of peak heads and associated weather events

A first impression of spatial dependencies arising from the natural spatial variability of weather events can be obtained by examining the spatial pattern of dates on which maximum peak heads occurred, assuming a uniform head response across the entire dike system. Figure 4.12 shows the dates of locally maximum hindcasted heads for two head responses, namely AGV15 and DL7. First of all, the dates at which the maximum

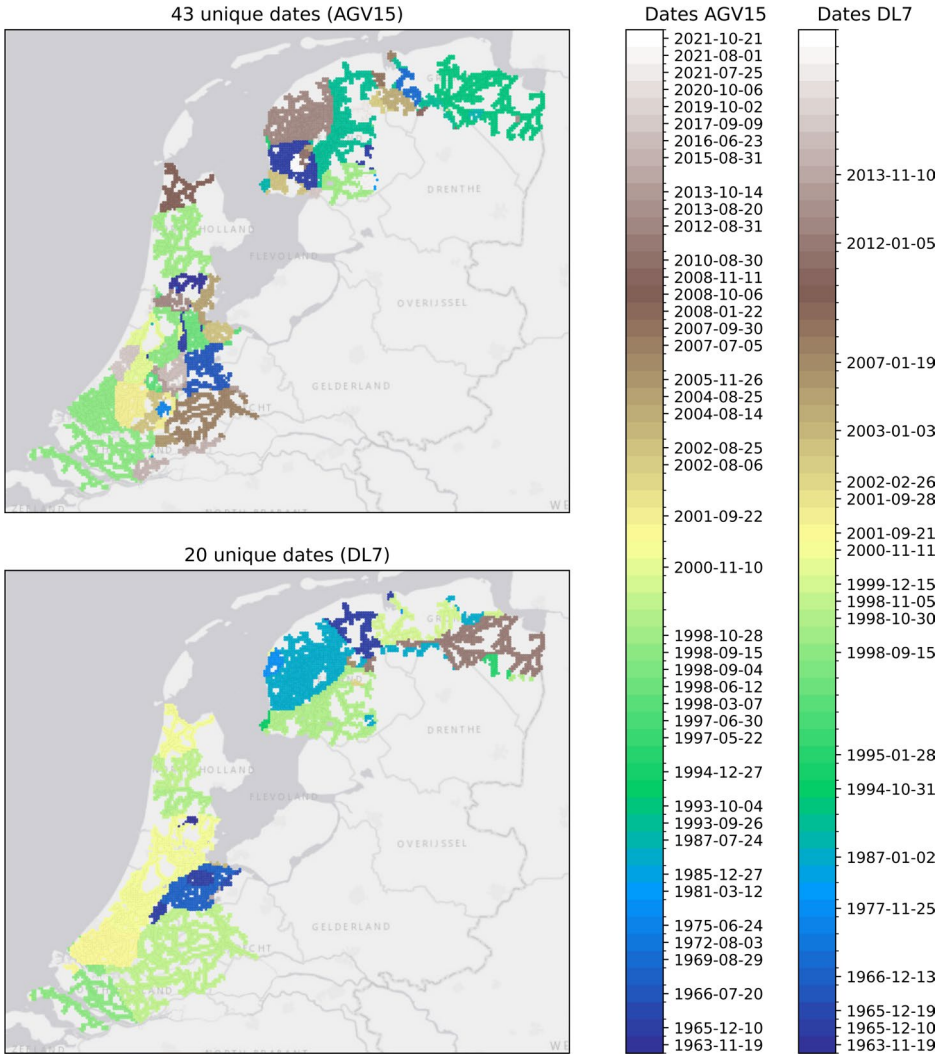


Figure 4.12 Dates of maximum hindcasted heads assuming a particular head response across the entire canal dike system, namely AGV 15 (top) and DL7 (bottom).

local heads occur clearly differ across the two head responses, indicating that different weather events cause peak heads to occur. Secondly, the spatial patterns differ: for AGV15, the number of unique dates of maximum local heads is larger and more localised than for DL7. The reason is that weather events with different typical precipitation durations also exhibit distinct spatial patterns. For AGV15, the response time of the hydraulic head is shorter than for DL7, and shorter precipitation durations are causing peak heads to occur. Extreme precipitation amounts in short periods are more localised than those in prolonged periods, resulting in more unique dates of locally maximum heads.

To illustrate how peak heads form, two examples of the underlying weather events are shown. For the time series model AGV15, extreme heads occurred in July 2021, when exceptionally heavy daily precipitation was recorded during an already wet winter in the province of Friesland, in the northern part of the canal dike system. On that day, more than 100 mm of rainfall fell within 24 hours, with an estimated return period of about 50 years under thecurrent climate. In contrast, for DL7, extreme heads are associated with multi-day rainfall events, such as the one in November 2000. The spatial patterns of precipitation and corresponding head exceedance frequencies for these two major rainfall events in the Netherlands are shown in Figure 4.13. Both events affected a large spatial area, but the most extreme rainfall amounts and peak heads were more localized. The short and intense event of 25 July 2021 produced sharp, confined zones with high head exceedance frequencies. In contrast, the long-lasting and widespread event of 10–13 November 2000 led to extreme head levels across broader regions, with a spatial footprint extending across multiple provinces. This illustrates how localised extreme precipitation within larger rainfall systems, together with varying head responses, shapes the spatial dependencies between peak heads in dike systems.

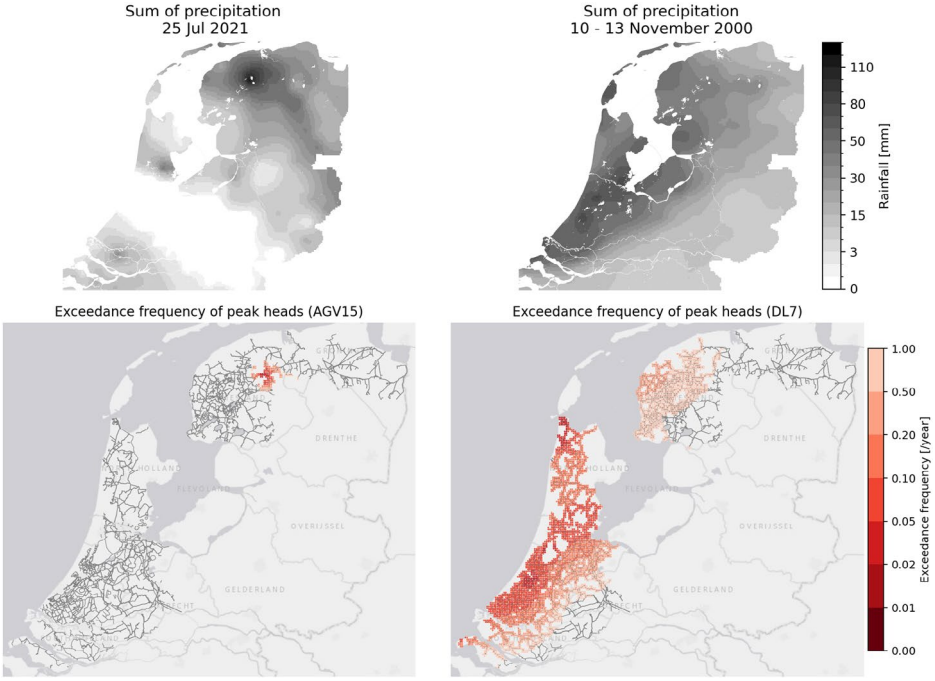


Figure 4.13 Illustrative example of spatial patterns of precipitation and corresponding head exceedance frequencies for two major rainfall events in the Netherlands. The top panels show total precipitation on 25 July 2021 and 10–13 November 2000, filtered to display only rainfall over land. The bottom panels show resulting peak head levels as exceedance frequencies (<1 /year), illustrating localised (left) and more widespread (right) peak heads occurring.

4.4.2 Length-effect factors at various spatial scales

In this section, the length-effect factor is analysed across different spatial scales and under two head response configurations: homogeneous and heterogeneous. At the polder scale, three polders with varying spatial dimensions were examined: the MT-polder, the Haarlemmermeer polder, and the Beemster polder. The dikes surrounding these polders are approximately 15, 40, and 60 km long, protecting areas of about 5, 70, and 180 km², respectively. Their geographical locations are shown in Figure 4.1.

Polder scale

Figure 4.14 shows the exceedance counts from 60 years of simulated heads for three polders, analysed under different spatial scales and head response configurations. The upper row shows the average exceedance counts at the section scale for all possible head responses across all three polders. Because meteorological conditions vary along the polder dike, this average indicates what typically occurs at the section scale. Overall, the exceedance counts are lower than the expected numbers according to probability theory (black dotted line). This discrepancy arises because the derived head statistics are based on current climatic conditions, in which the effects of already-occurring climate change have been corrected, while the hindcasted head levels were not adjusted. Consequently, head levels expected to occur once every ten years would theoretically appear about six times in a 60-year period, but in practice, they occurred on average only once or twice at the section scale.

At the polder scale, the number of exceedances increases compared to the section scale, as can be seen in the lower row of Figure 4.14. Under the homogeneous configuration, where all sections within a polder respond in the same manner, this increase relative to the section scale reflects the role of natural spatial variability of weather events that cause peaks to occur more frequently somewhere along the polder dike. However, this effect remains modest: the ratio of exceedances at the section and polder scales reaches values of about 2, see Table 4.1. This factor increases slightly for larger polders but remains relatively small, indicating that natural spatial variability contributes only marginally to independence between sections, and that peak heads across the polders remain strongly correlated. When heterogeneous head responses are assumed, meaning all possible head responses can occur within each square kilometre, the number of exceedances rises more strongly, by approximately a factor of four to five (see the red lines in Figure 14). The resulting length-effect factors derived from comparing exceedance counts at the polder scale under homogeneous and heterogeneous response configurations with the section-scale are shown in Table 4.1. For peak heads with an exceedance frequency of 1/5 per year, the length-effect factor at the polder scale ranges from about 1 to 14, depending on the response configuration and polder size. In general, the length-effect factor increases with lower exceedance frequencies. This

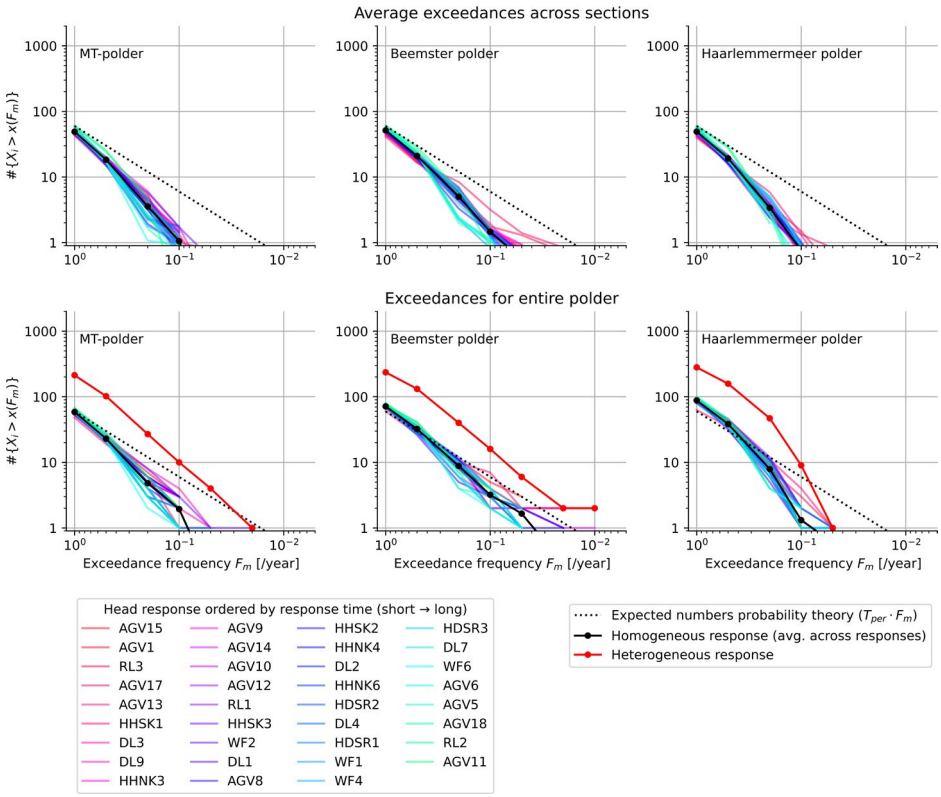


Figure 4.14 Exceedance counts of simulated peak heads for three polders with different spatial dimensions. The upper panels show the average exceedances across sections, and the lower panels show exceedances for the entire polder. The black and red lines represent homogeneous and heterogeneous responses, respectively, with colored lines showing individual head responses ordered by response time.

increase can have a physical cause, since more extreme weather events tend to have smaller spatial footprints, as will be discussed in the next section. However, it may also be partly due to statistical uncertainty. As can be seen, the length-effect factors remain small under the homogeneous head response configuration but increase under the heterogeneous configuration. This indicates that the length effect at the polder scale is driven mainly by variation in head responses, with spatial variability in the weather playing a secondary role.

Scale of the entire dike system

The same analysis was done at the scale of the entire dike system. Starting at average exceedances at the section scale, it can be seen that extreme peak heads for various head responses show some spreading, where they occur less frequently for longer response times than for shorter ones, see the upper graph in Figure 4.15. This can be

Table 4.1 Length-effect factors for different head response configurations, defined as the ratio of the average number of exceedances across sections and head responses to the number of exceedances at the polder scale for homogeneous and heterogeneous configurations.

Exceedance frequency F_m	MT-polder (A=5km², L=15km)		Beemster polder (A=70km², L=40km)		Haarlemermeer polder (A=180km², L=60km)	
	$LE_{Fm,hom}$	$LE_{Fm,het}$	$LE_{Fm,hom}$	$LE_{Fm,het}$	$LE_{Fm,hom}$	$LE_{Fm,het}$
1	1.2	4.3	1.4	4.6	1.8	5.7
1/2	1.3	5.6	1.5	6.3	2.0	8.2
1/5	1.3	7.6	1.8	8.0	2.3	13.9
1/10	1.8	9.5	2.2	11.0	1.7	11.8

explained by the fit of the extreme-value distributions: for longer response times, the shape parameter of the fitted distributions becomes more negative, resulting in a heavily bounded tail (see Appendix C.2). This makes the estimated exceedance frequency rather sensitive to small changes in peak heads. Therefore, the peak heads can be underestimated for longer response times based on the derived head statistics, making the hindcast peak head less extreme. As this affects the number of exceedances at various spatial scales in a similar way, its influence on the estimated length-effect factors is expected to be limited.

Considering the exceedance counts at the scale of the entire dike system, the spread in exceedance counts across the head responses increases further: for shorter response times, the number of exceedances at the system scale is even higher. These extreme peak heads are caused by weather events with heavy precipitation amounts over shorter durations, whereas longer response times are linked to longer-duration events. These short-duration heavy rainfall events typically cover smaller areas (Overeem et al., 2010; STOWA, 2019), making it less likely to affect a large part of the dike system simultaneously. This reduces spatial dependencies and results in more exceedances of peak heads at the scale of the entire dike system.

While the number of exceedances tends to increase with shorter response times, they also increase with more extreme peak heads. This is reflected in Figure 4.15, where the average exceedances across sections (top graph) and the curves at the system scale (bottom graph) diverge in opposite directions from the theoretical expectation: for more extreme events, the number of exceedances at the section scale falls below the theoretical line, while at the system scale it rises above it. The decreasing spatial extent with more extreme precipitation events likely contributes to this reduction in spatial dependency (STOWA, 2019).

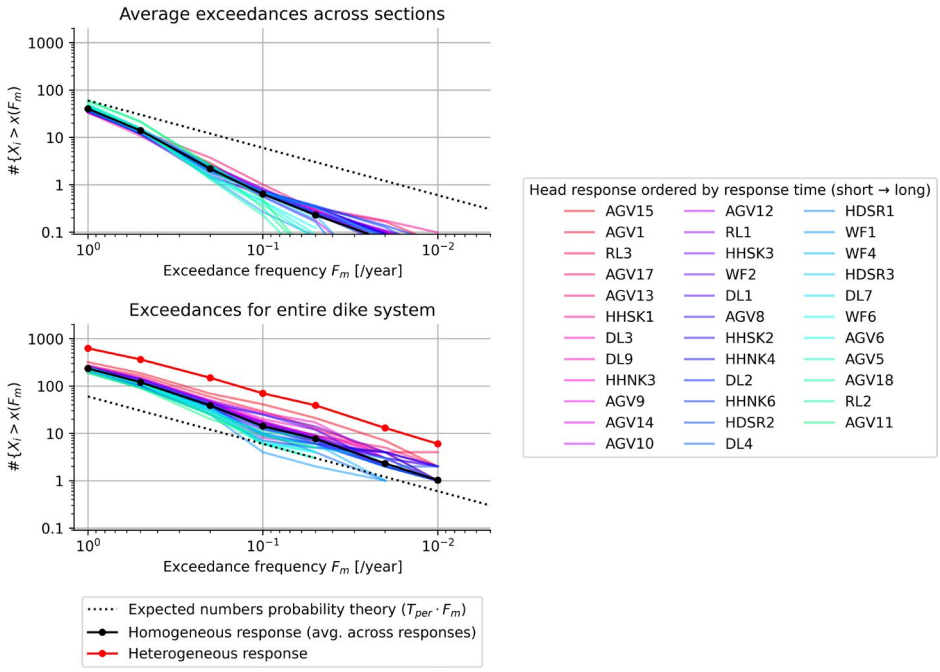


Figure 4.15 Exceedance counts of simulated peak heads over 60 years. The upper panel shows averages across sections, and the lower panel shows totals for the entire dike system. The black and red lines represent the homogeneous and heterogeneous responses, respectively, with colored lines showing individual head responses ordered by response time.

For the homogeneous configuration, where only one head response is assumed for the entire dike system, the length-effect factor is, on average, about 22 for head levels with an exceedance frequency of 1/10 per year, as can be seen in Table 4.2. This factor increases with smaller exceedance frequencies, and the length-effect reaches a factor of about 34 and 40 for head levels with exceedance frequencies of 1/50 and 1/100 per year, respectively. The heterogeneous head responses, at the system level, again result in an increase in the number of exceedances by about a factor of 4 compared to the homogeneous configuration. This factor is roughly constant across exceedance frequencies.

To conclude, while the length-effect at the polder scale is dominated by variation in head responses, the entire dike system is dominated by the spatial variability of weather events. The spatial dependencies identified suggest that peak heads with an exceedance frequency of 1/100 per year at the section level are expected to occur somewhere throughout the entire dike system each year, with a probability ranging from 33% to 90%, depending on the variation of head responses.

Table 4.2 Length-effect factors for different head response configurations, defined as the ratio of the average number of exceedances across sections and head responses to the number of exceedances at the entire dike system level for homogeneous and heterogeneous configurations.

Exceedance frequency F_m	$LE_{Fm,hom}$	$LE_{Fm,het}$
1	5.8	15.6
1/2	8.6	26.2
1/5	17.8	68.4
1/10	22.4	110.2
1/20	33.6	170.4
1/50	33.8	189.9
1/100	40.0	233.3

4.5 Discussion

4.5.1 Limitations

Spatial dependencies play an important role in dike safety and flood risk management, particularly when analyses are conducted over larger areas, such as at the national scale. However, available data and understanding of spatial dependencies can be limited (Kanning, 2012). As a result, estimates of spatial dependencies are subject to uncertainty. The approach adopted in this study also involves uncertainties, both in the underlying data and in the modelling framework, which are discussed in this section.

Hindcasting head levels

Head levels were hindcast using time series models. As with any model, outcomes outside the range of the calibration period are uncertain, unless it is known that the system behaviour remains unchanged. Deviations in head responses may occur, for instance, due to surface runoff caused by extreme precipitation amounts. These limitations are not expected to have a substantial impact on the findings, since they affect absolute levels rather than the occurrence and exceedance frequencies of extreme events. Furthermore, it was further assumed that the heads are primarily driven by precipitation and evaporation. This is generally reasonable, as water level fluctuations are typically small and rapid, only a couple of decimetres lasting a few days (Strijker et al., 2024). However, this assumption may not hold everywhere; for some dikes, local water-level variations may play a larger role, affecting their head levels and their spatial dependencies.

This study used historical precipitation data from observation stations that were spatially interpolated using ordinary kriging. As a result, the gridded data represent an approximation of the real-world weather patterns, which may be more irregular in nature. The use of radar-based precipitation data can improve the representation of these spatial patterns. The impact on the findings of this study is expected to be

limited, as spatial dependencies were quantified through the length-effect factor. This factor implicitly accounts for spatial patterns but primarily reflects the probability of any exceedance occurring along the dike system during a given weather event, rather than the actual spatial extent of that event.

Estimating spatial dependencies

In the analyses, a GPD was fitted to the selected peaks to estimate the exceedance frequencies of peak heads. As shown in Appendix C.1, this distribution generally provides a good fit, however, for certain locations, the tails of the distribution deviate from the data. Whether this comes from statistical uncertainty or has a physical explanation is difficult to tell. The choice of the fitted distributions influences the estimated exceedance frequency of peak heads and can affect the calculated length-effect factors. Next to the GPD, an exponential distribution was used to assess the sensitivity of various distributions. Overall, this distribution provided a worse fit to the peak heads, while the length-effect factors increased by roughly a factor of 1.5 for exceedance frequencies of 1/50 and 1/100 per year. The impact on the length-effect for higher frequencies was very moderate.

The derivation of head statistics was based on 100 years of weather data, corrected for already-occurred climate change, and defined for various rainfall regimes in the Netherlands. This was considered the most consistent approach. Nevertheless, alternative methods exist, for instance, deriving local head statistics for each raster cell directly from hindcasted heads. While such an approach would capture local variability more explicitly, it would also make the exceedance frequencies more sensitive to random spatial fluctuations in precipitation and to local model noise, potentially obscuring large-scale patterns and reducing the comparability of results across regions.

This study used hindcasted head levels from the past 60 years. Although extreme events occurred locally through the entire dike systems, events with exceedance frequencies around 1/100 per year remain relatively rare; tens of exceedances within the dike systems. In practice, however, even more extreme conditions are relevant for dike safety. Consequently, the spatial dependencies found for such low exceedance frequencies are subject to considerable statistical uncertainty. To obtain spatial dependencies representative of normative conditions for dike safety, results for peak heads with exceedance frequencies of about 1/10 – 1/50 per year can be used and extrapolated to less frequent, more extreme conditions.

4.5.2 Implications

The findings of this study have several implications for understanding how spatial dependencies in hydraulic heads influence the reliability of canal dike systems and for

improving approaches to dike safety management at both regional and national scales. The following subsections discuss how the results relate to the occurrence of failures in canal dike systems and to the potential use of hindcasts within an Observational Method for dike safety.

Management of dike safety and flood risk at larger spatial scales

Using a risk-based approach for dike safety, target reliabilities for canal dike cross-sections are often derived from acceptable probabilities of flooding in areas where spatial dependencies are essential. By translating the acceptable flood probability of a polder to target reliabilities at cross-sections, the length-effect factor of the failure probability is often used. For inner-slope instability, the target reliabilities are often derived using an equivalent independent length of 50m, which arises from correlation lengths of soil properties and is consistent with historically observed failure lengths. So the length-effect factor of the failure probability is based on the spatial dependencies of time-invariant knowledge uncertainties, since these are often in practice the variables that dominate the failure probability (Schweckendiek et al., 2017). While peak heads are time-variant variables that trigger instabilities and can increase the dependency between sections, this is unlikely to change the length-effect factor of the failure probability used in dike stability assessments since these variables do not dominate failure probabilities. Nevertheless, it is recommended to keep in mind that each dike section has a specific deterministic resistance based on the present subsurface conditions, however, its value is unknown due to insufficient information or knowledge. As discussed in Section 4.2.2, in the hypothetical case where the resistance of each dike section is known exactly, the section failure probabilities would be attributable solely to load variability. Although this is a hypothetical scenario, it highlights that spatial load dependencies play a key role in explaining observed failures, as will be discussed in the next subsection.

The occurrence of failures in dike systems

For the inner-slope instability of dikes, hydraulic heads are important and can trigger failures, but they do not have to be the only load variable. For example, on many canal dikes, roads are located, and traffic can be an important factor in stability. Therefore, the length-effect factor for peak heads across the entire dike system provides a limited perspective on the load dependencies between dikes. Therefore, one may need a model that expresses the probability of traffic intensity conditional on peak heads, or vice versa, as illustrated by Lendering et al. (2018).

Assuming that peak heads are the dominant load variable and that epistemic uncertainties in resistance are absent, resulting in no spatial independencies from the resistance, the length-effect factors of peak heads can be used to estimate system-

failure probabilities for the Rijnland water authority area, as discussed in Section 4.2.2. Therefore, we estimated length-effect factors of peak heads specifically for the Rijnland water authority area and found values of 5 and 28 for the 1/10 per year exceedance frequency under homogeneous and heterogeneous head response configurations, respectively. Since extreme events are relatively rare at this spatial scale, the length-effect factor for the 1/100 per year exceedance frequency was estimated based on a rational increase in the length-effect at the system level, resulting in a length-effect factor of peak heads between 9 and 50. Given that 300 kilometres of canal dikes are assessed as not meeting the safety standard with a maximum allowable annual failure probability of about 1/500 per year, and assuming they are equally distributed throughout the area, the yearly system-failure probabilities should be about 2% ($n \approx 9$) to 10% ($n \approx 50$). This aligns more closely with the observed number of failures compared to system-failure probabilities determined using equivalent independent lengths of 2,000 m and 200 m, based on conservative correlation lengths for soil properties. It should be noted, however, that the calculated failure probabilities contain a large degree of knowledge uncertainty related to time-invariant geotechnical parameters, rather than natural variability. As such, they are not strictly frequencies and this comparison with observed failure rates should be treated as a rough indication rather than a direct validation. The comparison is still relevant, since differences between expected and true dike behaviour affect decisions on, for example, dike reinforcement and evacuation planning. Reducing epistemic uncertainties is therefore an important step toward better aligning estimated failure probabilities with observed failures and load conditions.

Usefulness of hindcasts in dike safety using the Observational Method

Analysing the slope stability of dikes can require a vast amount of effort and labour, with only rough estimates of physical constants that appear in the equations (Terzaghi, 1961). Alternatively, stability problems can be considered using the 'learn-as-you-go' method, also known as the Observational Method. The Observational Method emerged because, in many geotechnical applications, it is not feasible to assume highly conservative values for loads and material properties and design solely for those conditions (Peck, 1969; Baecher & Christian, 2005). Instead, reasonable estimates of the parameters and their uncertainties were made, and observations of its performance were made so that appropriate measures could be taken. A structured and quantified approach to incorporate observations and detect differences that exist between the assumed and the real subsoil conditions is reliability updating (Schweckendiek et al., 2017). In this way, hindcasts can also help obtain estimates of survived extreme head levels, at least their extremity, rather than absolute levels. This, combined with monitoring parameters that serve as proxies for dike strength, could provide an effective and practical approach to managing dike safety, as the key question often concerns the

dike's strength and its evolution over time. As long as the strength does not deteriorate, this method can be highly effective. Moreover, since load uncertainties are relatively small and high load levels occur annually, both for peak heads (Strijker & Kok, 2025) and for traffic loads, this approach may be particularly suitable for improving the safety management of Dutch canal dikes.

Discussion on the origin of spatial dependencies

The spatial dependencies found in peak heads are larger than typically seen for soil properties. This suggests that hydraulic loading across canal dikes can be more spatially coherent than the underlying geotechnical variability alone would predict. One explanation is that spatial variability influences local soil properties, but the response to extreme weather events is largely controlled by structured, permeable shallow soil layers, as peak head levels are typically close to the surface. Because such structuring is widespread in topsoils, variations in head responses between locations remain limited, thereby strengthening spatial dependencies in hydraulic loads and increasing the likelihood of simultaneous high heads along the dike. Whether these structured, permeable layers indeed explain the observed coherence remains an open question. Further observations and evidence are needed to determine how much of the spatial dependence in peak heads can truly be attributed to these shallow soil properties.

4.6 Conclusions

This study aimed to answer the following research question: What is the length-effect factor of peak hydraulic heads at various spatial scales for Dutch canal dikes? This was done by hindcasting hydraulic heads across the entire Dutch canal dike system from 1962 to 2022, a 60-year period, using time-series models. By analysing peak head occurrences at various spatial scales, the length-effect factor was estimated for three polders and the entire canal dike system.

At polder scales, it was found that the length-effect factor primarily arises from variations in head responses, while the spatial variability of weather events across polders contributes only marginally, though it increases with larger polders. The spatial variability causes extreme peak heads at polder scales to occur roughly up to twice as often as at the section level, whereas variation in head responses can increase this occurrence by up to a factor of four to five.

At the scale of the entire dike system, the length-effect factor is dominated by the spatial variability of weather events. This variability differs among head responses: dikes with shorter response or drainage times are more sensitive to short-duration rainfall, which tends to have smaller spatial extents and thus leads to larger length-

effect factors. On average, hydraulic head levels with a 10-year exceedance frequency at the section level occurred about 20 times more often at the system level, for the homogeneous head response configuration. In contrast to polder scales, the length-effect factor at the national scale increases with smaller exceedance frequencies of peak heads, consistent with findings from other studies showing that the spatial extent of extreme precipitation decreases with increasing intensity. The length effect factor reaches 40 for head levels with an exceedance frequency of 1/100 per year under homogeneous head response. Assuming that the head response is heterogeneous, again, the length-effect factor increases by about a factor of four to five, compared to the homogeneous configuration. These length-effect factors suggest that peak heads with an exceedance frequency of 1/100 per year at the section scale are expected to occur somewhere in the entire dike system each year, with a probability ranging from 33% to 90%, depending on the variation in head responses.

5

The role of load variations in assessing credible dike failure probabilities: balancing load and strength uncertainties

This chapter investigated the role of load variations in reliability updating and assessing credible failure probabilities. It was found that the impact of reliability updating increases when load variations are small, because relatively frequent load levels contribute most to the failure probability, and since there is evidence of good performance close to these levels, there is greater potential to improve the estimates through reliability updating. Moreover, a credibility check was introduced for dikes that have been stable for decades, where load levels with return periods of up to 10 years are not expected to contribute more than 50% to the failure probability, indicating an imbalance between load variation and strength uncertainty. This imbalance occurs when the inverse gradient of the fragility curve exceeds 1.5 times the decimate height of the load. Many Dutch dikes, including canal dikes and dikes along the large lakes and delta regions, have small decimate heights. For these dikes, strength uncertainties must be sufficiently small to obtain credible failure probability estimates.

This chapter is based on Strijker, B., Jonkman, S. N., & Kok, M. (2026). The role of load variations in assessing credible dike failure probabilities: balancing load and strength uncertainties. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 1-22. https://doi.org/10.1080/17499518.2026.2654177

5.1 Introduction

To provide flood protection, transportation, energy and freshwater, society depends on various earthen embankments. The safety of these structures often requires periodic assessment (European Commission et al., 2024; CEN, 2004). In the Netherlands, over 10,000 kilometres of dikes provide flood protection, including both regional and primary dikes (see Figure 5.1). These are earthen embankments that retain water. While primary dikes protect the hinterland from floods caused by major water bodies such as the sea, lakes, and large rivers, regional dikes are located along smaller inland water bodies. In the event of a dike breach, both types of floods can lead to severe damage, while fatalities are only expected at primary dike breaches. Therefore, estimating credible dike failure probabilities is a key question for Dutch society in managing flood risk. Credible implies that the failure probability is realistic and justifiable (Phoon, 2023), derived from a proper and well-informed engineering assessment using appropriate models based on available data, and not contradicting the (historical) behaviour or performance of the flood defence. Simpson et al. (1981) suggested using so-called worst credible parameter values in calculations, which have a probability of being exceeded about 0.1%, as also discussed in Phoon (2016). These values reflect loads and material properties that the designer could still realistically believe might occur (Simpson et al., 1981). In flood safety, however, relevant load levels can be more extreme than those with a 0.1% annual exceedance probability (Kind, 2014; Jonkman et al., 2011). Consequently, the interpretation of what is credible, or worst credible, depends on the geotechnical structure and its observed and expected performance. The authors prefer using the term credible when linking probabilities to observed performance. For example, suppose a dike has an estimated failure probability of 1/100 per year but withstood a load level with an exceedance probability of 1/500 per year, then that failure probability estimate is not credible. Ultimately, rare events do happen, which are unlikely, but not necessarily incredible. The pursuit of credible failure probabilities prevents dikes from being unnecessarily and unjustifiably estimated as unsafe, and supports making investment decisions based on a well-founded safety profile.

One of the failure mechanisms contributing to the overall failure probability of dikes in the Netherlands is inner-slope instability, referred to as dike instability (Jongejan & Maaskant, 2015). This failure mechanism occurs when high water pressures inside the dike reduce the shear strength and induce loss of stability. It accounts for the largest number of primary dike trajectories failing to meet safety standards (Nationaal Georegister, 2024), while historically it accounted for only 5% of the failures (Van Baars & Van Kempen, 2009). The same holds for regional dikes, where hundreds of kilometres are considered unsafe on a national scale, which contradicts observations (Rikkert, 2022). Many of the Dutch dikes are centuries old and have often been reinforced and upgraded over the years, resulting in heterogeneous dikes with varying soil

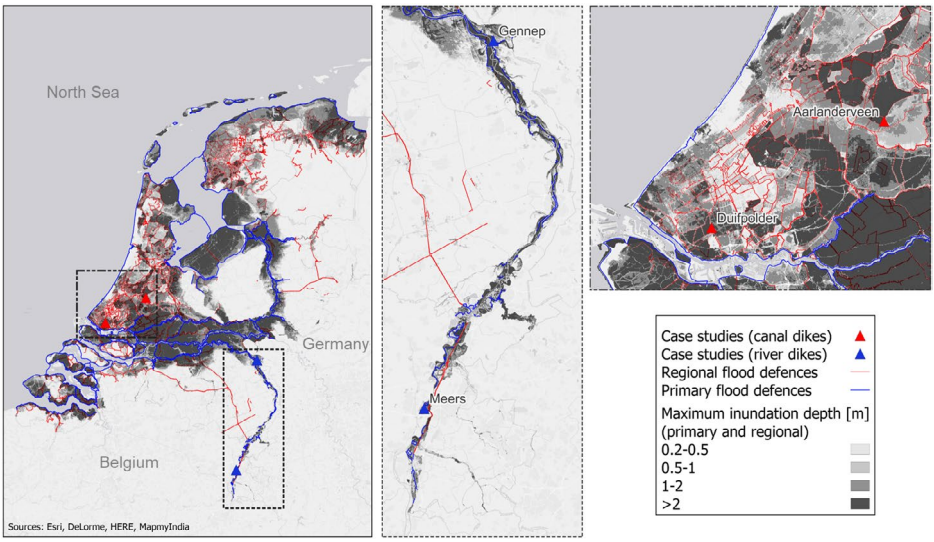


Figure 5.1 Maps of the Netherlands that show the maximum inundation depths (combined from many flood scenarios from primary and regional dike breaches) and the dike systems for both primary (blue) and regional (red) flood defences, along with the highlighted case study sites that are marked (▲) and labelled.

compositions (Van Woerkom et al., 2022). This heterogeneity, combined with limited data on the subsurface buildup and the uncertain shear strength of the soil, makes the overall strength of dikes uncertain. This uncertainty complicates the assessment of whether a dike meets safety standards and raises the question: are dikes truly unsafe, or is it challenging to demonstrate adequate safety with the available data and models?

In general, assumptions and choices in failure probability modelling tend to be conservative, or safe-sided (European Commission et al., 2024). This results in so-called hidden safety elements in safety assessments, leading to smaller observed failure rates compared to estimations made using models. This gap between observations and model outcomes can be reduced by understanding parameter sensitivity in reliability analyses (Yang et al., 2024) and the use of Bayesian inference (Huang et al., 2025; Yang et al., 2024) to reduce model parameter uncertainty and update the predicted reliability. The specific application of Bayesian inference is often called reliability updating, which uses Bayesian principles to refine reliability predictions by integrating prior knowledge with new knowledge. This can be measurements, monitoring outcomes, or observations of system performance (Straub, 2011). It is a well-established and frequently used technique in structural engineering to update the failure probabilities of structures (De Koker et al., 2020; Straub & Papaioannou, 2015; Hemel et al., 2024). While less common for dikes, it is still applied to various failure mechanisms, like overflow, piping

and slope stability (Huang et al., 2025, Schweckendiek et al. 2017; Schweckendiek et al., 2014; Schweckendiek & Vrouwenvelder, 2013; Rikkert et al., 2022; van der Krogt et al., 2022). Piping and slope stability are failure mechanisms that often involve large epistemic strength uncertainties due to uncertain model schematizations and subsurface conditions. These uncertainties are virtually time-invariant, and as shown by Schweckendiek et al. (2014), these uncertainties can be especially reduced using reliability updating. While the role of time-invariant epistemic strength uncertainties in reliability updating is well understood, the role of time-varying inherent load uncertainties, or simpler load variations, is poorly understood.

The central question addressed in this chapter is: What is the role of load variations in reliability updating and assessing credible failure probabilities?

The outline of this chapter starts with the materials and methodology in Section 5.2. This section describes the characteristics of typical Dutch dikes, the methods used to calculate the probability of dike instability, along with the applied reliability updating techniques and the schematization of the case studies. Section 5.3 discusses the results of the conceptual analysis and the case studies. Section 5.4 explores the implications of these findings for failure probability analysis and outlines the limitations. Finally, Section 5.5 concludes with final remarks.

5.2 Materials and methods

5.2.1 Some typical Dutch dikes and their key characteristics

The Netherlands has a variety of dikes that protect the low-lying land, like sea, river and lake dikes, as well as compartmentalising dikes and dikes along the canals (see Figure 5.1). All these dikes have different characteristics, like dimensions, subsoil conditions and boundary conditions. The case studies in this study considered canal and river dikes, of which the typical differences in dike characteristics are summarised in Figure 5.2 and are discussed in more detail.

The canal dikes in the Netherlands are located along drainage canals in polders in the northwestern part of the Netherlands. Water levels in these canals can be several meters above the polder levels and are maintained near target levels using pumping stations and weirs. In extreme situations, canal water levels typically rise several decimeters above the target levels. In contrast, the Dutch rivers exhibit highly dynamic water levels, often fluctuating by several meters between average annual conditions and extreme floods (Klijn et al., 2019). This difference in water level fluctuations is a key distinction between canal and river dikes and impacts the dynamics of pore-water pressures, which are crucial for dike stability (Ridley et al., 2004; Van der Meij, 2023).

Pore-water pressures in the dike body are commonly expressed by the phreatic line. This phreatic line can be influenced by both the intrusion of outside water and precipitation on top of the dike. Which driver plays the biggest role varies with the dike's specific characteristics. At canal dikes, where water level fluctuations are small, precipitation is expected to be the primary factor influencing the phreatic line, as supported by previous studies (Strijker et al., 2024; Rikkert, 2022). In contrast, river water levels can fluctuate significantly and for extended periods, often lasting weeks, making them the main driver of the phreatic line's position within river dikes.

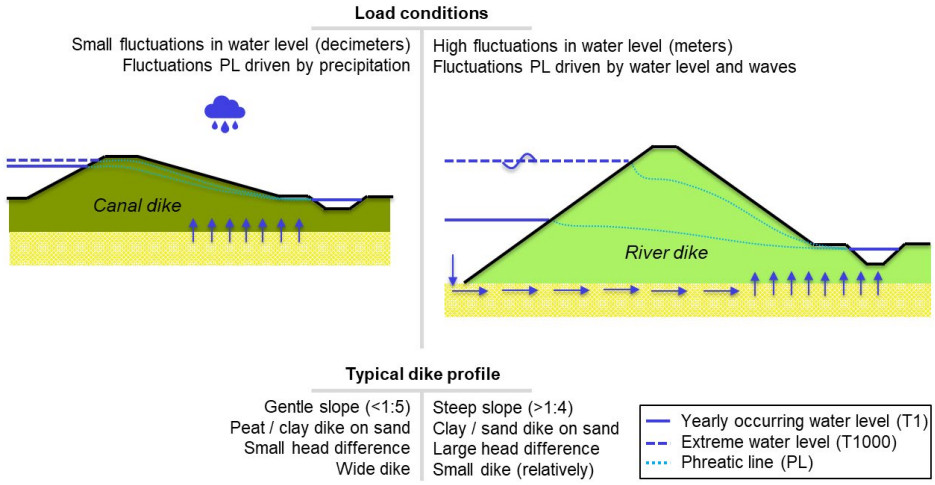


Figure 5.2 Schematic representation of a typical river dike and canal dike, illustrating characteristic loading conditions and dike profiles.

Other relevant differences include the subsurface material and geometry. The canal dikes are often heightened with locally available soils, mainly a mixture of clayey and peaty material. The dikes are, in general, rather wide with gentle slopes and small head differences. The canal dike bodies are situated on Holocene deposits ranging in thickness from a few meters to over 20 meters, overlying a Pleistocene aquifer. The river dikes have a similar basis of Holocene and Pleistocene layers, but the dike bodies are typically composed of sand and clay. Although both dikes are located on the Pleistocene layers, the pore-water pressure dynamics are different. Unlike canal dikes, the underlying aquifers of river dikes are often directly connected to the outside water level. As a result, high water levels can cause a significant increase in pore-water pressure beneath the dike, which impacts dike stability. The fluctuations in pore-water pressures in the aquifer underlying canal dikes are limited and driven by seasonal variations in precipitation and evapotranspiration (Wong et al., 2007). The response is slow and damped by the overlying Holocene clay and peat layers.

5.2.2 *Failure probability analyses and fragility curves*

Uncertainties: epistemic and inherent randomness

The probability of dike instability results from uncertainties in both load and strength. These uncertainties can arise from inherent randomness or a lack of knowledge, referred to as inherent and epistemic uncertainties, respectively. While inherent uncertainty cannot be practically reduced through additional analysis or data collection, such as fluctuating water levels and pore water pressures, epistemic uncertainties can. Examples of epistemic uncertainties are geotechnical parameters, such as the shear strength. This uncertainty arises primarily from the limited number of soil samples tested at a given dike section, and the uncertainty in estimating this geotechnical parameter reflects spatial variability and measurement error (Phoon & Kulhawy, 1999). In practice, soil samples from various locations are combined to create a regional dataset, after which factors like variance reduction from averaging spatial variability, statistical uncertainty, and measurement uncertainty are considered (Calle et al., 2021; Van der Krogt et al., 2019; Vanmarcke, 1977). Estimates of these geotechnical parameters remain uncertain and are incorporated into the reliability analyses. These uncertainties are time-invariant and, in practice, deterministic within a year. Although soil properties can naturally vary over time due to geological processes, this is not a random process, but rather a gradual and patterned evolution. When assessing the failure probability within a specific year, the temporal variation of geotechnical parameters is expected to be small relative to the present uncertainty in estimating them from available data, and the uncertainty in soil parameters is mainly epistemic. This means that the dike strength in safety assessments is often assumed to be time-invariant and epistemically uncertain, while most loads are typically classified as time-varying and inherently uncertain (Schweckendiek et al., 2017).

Fragility curves

A commonly used approach to estimate the dike failure probability is the use of fragility curves (Casciati & Faravelli, 1991; Schultz et al., 2010; Schweckendiek et al., 2014; Bachman et al., 2013). This curve describes the dike's strength by expressing the relationship between load levels and failure probabilities conditional on these loads. To calculate the failure probability of a dike, the fragility curve needs to be combined with the probability density function of the load events:

$$P(F) = \int_{S=-\infty}^{S=\infty} f_S(s)\, F_R(s)\, ds \qquad [5.1]$$

where $f_s(s)$ is the probability density function of load levels and $F_R(s)$ is the cumulative density function of the resistance R, that gives the probability of failure conditional on the load *S*, also known as the fragility curve. When analysing the role of loads, it is more practical to focus on the exceedance probability, $1\text{-}F_S(s)$, rather than the probability

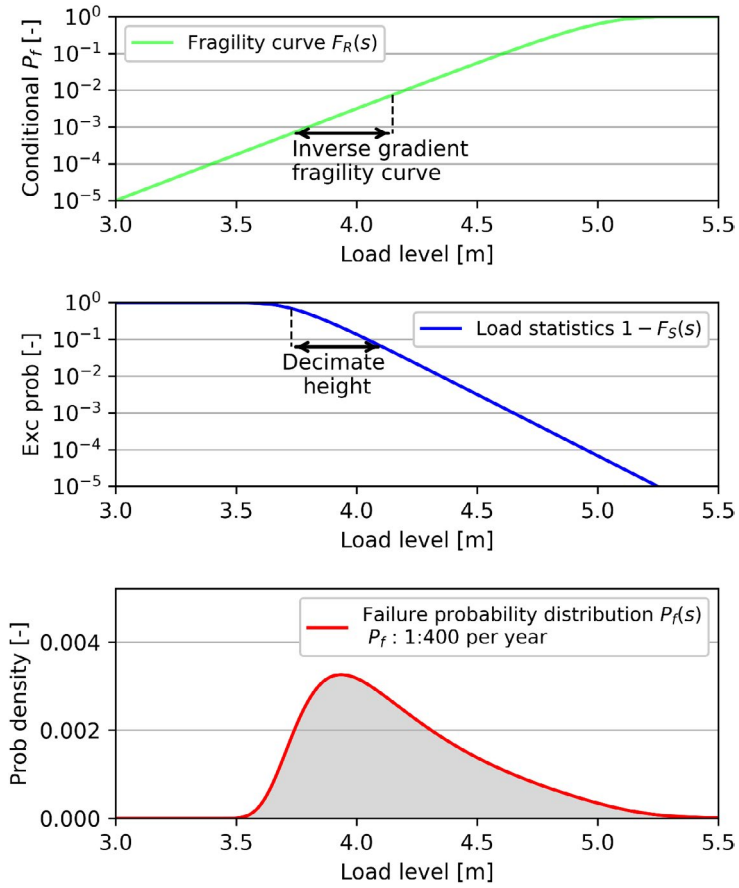


Figure 5.3 Graphical representation of the dike failure probability: the fragility curve (top), load statistics (middle), and the resulting failure probability distribution (bottom). The integrated failure probability density distribution gives a failure probability of 1:400 per year in this example.

density function. This is because extreme load events are the primary concern, and the exceedance probability effectively highlights the likelihood of these occurrences. These different components are shown graphically in Figure 5.3, where an example of a fragility curve (top panel), load statistics (middle panel) and resulting density function of the failure probability are shown (bottom panel). This density function is the product of the fragility curve and the probability density function of the load. It provides insight into the contribution of various load levels to the overall probability of failure, as well as the load event where failure is most likely to occur, which is at the peak of this distribution. The integral of this distribution gives the failure probability, which in this example is 1/400 per year.

The shape of the fragility curve reflects the amount of strength uncertainty. This depends not only on the stochastic input variables but also on the failure mechanism itself. In general, the curve approaches a step function for very well-understood, like wave-overtopping, and a more S-shaped curve corresponds to more poorly understood, like slope instability (Van der Meer et al., 2009). A measure to quantify this shape is the inverse gradient, which represents the increase in load level that results in a tenfold rise in the conditional failure probability, as illustrated in the upper panel of Figure 5.3. A higher inverse gradient corresponds to a gentler curve, indicating larger strength uncertainties. The inverse gradient can be taken at any part of the fragility curve, but in this study, the inverse gradient is considered at the load level scenario where failure is most likely to occur, so at the peak of the failure probability distribution.

A measure for the load variations is the decimate height. This is defined as the increase in load level that occurs when the exceedance probability of the load level experiences a tenfold decrease (Wojciechowsk,a 2015; Schweckendiek, 2014). Smaller decimate heights correspond to smaller load variations. This measure is illustrated in the middle panel of Figure 5.3.

Reliability updating using Bayesian analysis

Traditionally, engineers deal with uncertainties in failure probability analyses by starting with initial estimates based on available information and refining these estimates through additional field observations. This process can be made more structured by using probability theory, and specifically Bayesian probability updating, to revise prior judgments. This method can be used to update failure probability estimates based on observations, such as the survival of observed load conditions without failure indications. When survived load levels are used to update failure probabilities, Bayesian updating is applied by eliminating implausible values of uncertain strength parameters and model uncertainty, assuming the uncertainty is epistemic. In general, this can be done in two ways that are mathematically equivalent and differ only in terms of implementation: the direct and indirect reliability updating techniques (Schweckendiek, 2014; Schweckendiek, 2010). Both techniques use the Bayes' Rule (Bayes, 1763) that forms the basis for updating probabilities with new evidence:

$$P(F|\varepsilon) = \frac{P(\varepsilon|F)P(F)}{P(\varepsilon)} \qquad [5.2]$$

where F is the failure event to be predicted and ε the observed event or evidence. In this study, the evidence is survival at an observed load level s_{obs}, where survival implies that the dike strength R is larger than the observed load: $R > s_{obs}$ (inequality type of information).

$$P(F|\varepsilon) = \frac{P(F \cap \varepsilon)}{P(\varepsilon)} = \frac{P(F, R > s_{obs})}{P(R > s_{obs}))} \quad [5.3]$$

where R denotes the random variable representing dike's resistance or strength, expressed in the same units as the load variable S (e.g., phreatic head or water level).

The indirect method uses Bayes' rule to update the PDFs of the underlying random variables describing the strength, while the direct method exploits the definition of the conditional probability of failure F given the evidence ε to update the failure probability directly. In this study, the direct method using fragility curves is preferred because it is simpler to implement and aligns with the proof-of-concept focus, rather than providing detailed updates to stochastic variables. This method directly calculates the updated fragility curve and probability of failure, but does not update the underlying stochastic variables, which is beyond the scope of this study. Direct reliability updating can be applied using fragility curves and survival information, which leads to the following expression for the updated failure probability:

$$P(F|\varepsilon) = \frac{\int_{s_{obs}}^{\infty} P(s_{obs} < R < S) f_s(s) ds}{1 - F_R(s_{obs})} = \frac{\int_{s_{obs}}^{\infty} (F_R(s) - F_R(s_{obs})) f_s(s) ds}{1 - F_R(s_{obs})} \quad [5.4]$$

Since the load density function $f_S(s)$ multiplies the entire integral, the expression can be written more compactly as:

$$P(F|\varepsilon) = \frac{\int_{s_{obs}}^{\infty} (F_R(s) - F_R(s_{obs}))\, ds}{1 - F_R(s_{obs})} f_s(s) \quad [5.5]$$

From this, the posterior or updated fragility curve can be written as:

$$F_{R,upt}(s) = P(F|\varepsilon) / f_s(s) = \frac{\int_{s_{obs}}^{\infty} \big(F_R(s) - F_R(s_{obs})\big) ds}{1 - F_R(s_{obs})} \quad [5.6]$$

$$F_{R,upt}(s) = \begin{cases} \dfrac{F_R(S) - F_R(s_{obs})}{1 - F_R(s_{obs})}, & s \geq s_{obs} \\ 0, & s < s_{obs} \end{cases} \quad [5.7]$$

Figure 5.4 shows an example of the effect of reliability updating on the fragility curve and the failure probability distribution.

Note that updating only has an effect if past and current epistemic strength uncertainties are fully correlated and time-invariant. This study assumes time-invariant uncertainties, a reasonable choice for many geotechnical properties (Schweckendiek, 2014). It should be noted, however, that in practice the observed survived load level s_{obs}

may itself be uncertain due to measurement errors or spatial variability in pore-water pressures. Such uncertainties effectively reduce the information content of the survival evidence, which in turn diminishes the impact of the updating. In this study, a best-estimate s_{obs} was used. This assumption, and its resulting limitations, is important to keep in mind for practical applications.

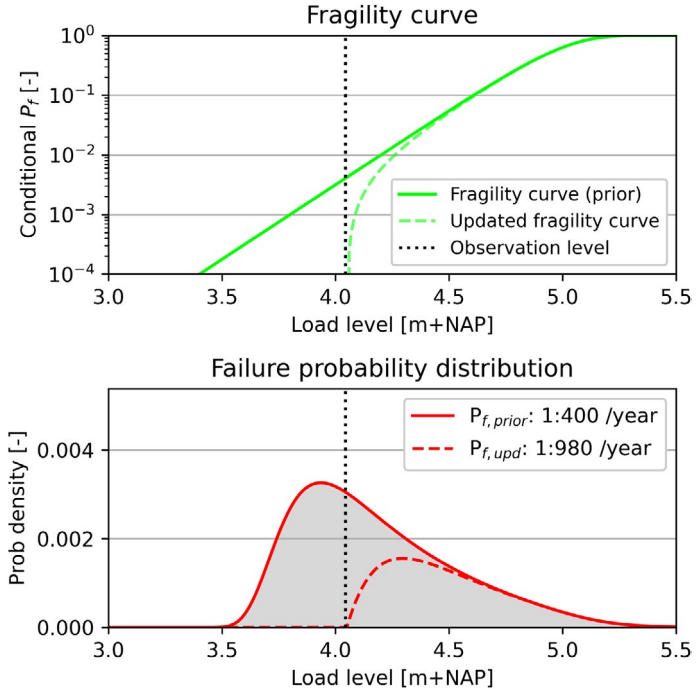


Figure 5.4 Example of prior and updated fragility curve (top), based on an observed and survived load level, and the resulting failure probability distribution (bottom) with the prior and updated failure probabilities. The load statistics are unchanged and correspond to those shown in Figure 5.3.

5.2.3 Conceptual analysis

To examine how load variations influence reliability updating, a conceptual analysis was performed on hypothetical dikes using a range of combinations of load variations and fragility curves. Both were described with two-parameter Gumbel distributions. Their spreads were varied while the prior failure probability was fixed at 1/100 per year by adjusting the relative distance between the two distributions. In practice, fragility curves and load statistics for dikes are often derived from probabilistic analyses using numerical integration and do not necessarily follow a probability density function. As a result, the decimate height and inverse gradient of the fragility curve can be estimated by taking the tangent of the curve. For this conceptual analysis, these measures follow

directly from the Gumbel distributions. The cumulative distribution function of the Gumbel distribution is:

$$F(x) = \exp\left(-\exp\left(-\frac{x-\mu}{\beta}\right)\right) \quad [5.8]$$

where μ is the location parameter and β the scale parameter. The decimate height and inverse gradient follow from the scale parameter by multiplying β with $ln(10)$.

For the conceptual analyses, the load variations were varied between decimate heights of 0.05 and 0.65 m, and the strength uncertainty between inverse gradients of the fragility curve from 0.05 to 1.55 m. For each case, reliability updating was applied using a survived load level that occurs on average once every 10 years.

5.2.4 Modelling dike stability

The dike's probability of instability is generally assessed using two-dimensional limit equilibrium methods, such as Bishop, Spencer or Uplift-Van, which calculates the factor of safety against instability considering driving forces and resisting forces acting on a slip plane (Bishop, 1955; Sharp et al., 2013). These methods, in combination with search algorithms to find the slip plane with the minimum factor of safety, are implemented in D-stability, which was used in this study (Van der Meij, 2023). Next to calculating the factor of safety, this software can also calculate failure probabilities using Adaptive Importance Sampling or First Order Reliability Method (FORM) analysis. While importance sampling includes a search algorithm, FORM analysis was carried out with a fixed slip plane, which was found using a calculation with design values (Van der Meij, 2023). Importance sampling may not converge for very small failure probabilities. In those cases, a FORM analysis was used as an alternative approach. The resulting failure probability is always calculated for a single load scenario, representing only one point on a fragility curve. The complete fragility curve can be constructed by performing multiple calculations using a range of different load scenarios with various headlines.

5.2.5 Case study description

To demonstrate the role of load variations in reliability updating and assessing credible failure probabilities, four realistic case studies are examined in this section. These case studies include two canal dikes and two river dikes. This section describes the schematization of the dikes, including geometry, subsurface materials, soil properties, load variations and the survived load levels.

Dike geometry and geotechnical properties

The two canal dikes are located in the western Netherlands and are named after the polder that they protect: Aarlanderveen and Duifpolder. The two river dikes are located

in the southern part of the Netherlands and their names were retrieved from the nearby towns they protect; Gennep and Meers. The geographic locations of the dikes are shown in Figure 5.1. The canal dikes consist more of peat and organic clay, while the river dikes are made up of clay and sand, as can be seen in the cross-sectional profiles in Figure 5.5. All dikes are located on top of Pleistocene sand layers, which are located very deep, over 15 meters below the surface, at the Duifpolder.

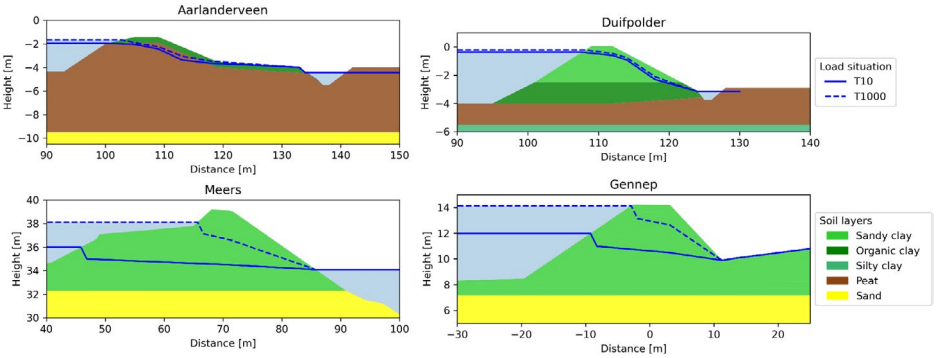


Figure 5.5 Cross-sectional profiles of the considered case studies showing the soil types and schematized phreatic lines for different probabilities of occurrences. Only two phreatic lines are shown here, whereas in the analyses, more load situations were considered, including pore-water pressures in the aquifer.

The soil parameters for all layers are listed in Table 5.1. Shear strengths were modelled using Mohr-Coulomb, based on a regional dataset of over 5,000 shear strength tests for typical Dutch geological deposits (triaxial for clay, direct simple shear for peat) and over 13,000 classification tests (STOWA, 2023). Strength parameters were treated as stochastic variables, while volumetric weights were considered deterministic using mean values. When the variance of the strength parameters in this dataset is derived, it includes the effects of spatial variability, statistical uncertainty, and measurement uncertainty (Phoon & Kulhawy, 1999; Calle et al., 2021). The strength uncertainties applied in dike-stability reliability analysis should reflect the shear strength uncertainty of a layer. Because relevant slip planes span both weaker and stronger zones, the dataset's variance is reduced to account for spatial averaging effects. However, since the dataset consists of tests within a larger region, not all variance can be reduced. Variance reduction is done by a simplification of the basic random field model, where it is assumed that 25% of the variance in the dataset is due to regional variations of the mean, while 75% of the variance is due to variations around the local mean that spatially average out (Calle et al., 2021; Calle, 1996). This results in standard deviations of shear strength parameters that are approximately half those found in the dataset.

Two additional scenarios were included to assess the impact of strength uncertainty: a moderate scenario (10% variation in cohesion, 5% in friction angle) and a high strength uncertainty scenario (20% variation in cohesion, 10% in friction angle). The high strength uncertainty was based on the NEN 9997-1, and also suggested by the JRC report on the reliability background of the Eurocodes (Vrouwenvelder et al., 2024). These coefficients of variation represent the uncertainty in spatial average ground strength parameters. In both scenarios, the deterministic soil properties remained unchanged.

Table 5.1 Deterministic and stochastic soil parameters in the reference case. The presented standard deviations are based on the full dataset, whereas these values are reduced in the stability analyses to account for averaging spatial variability.

	Vol. Weight (Deterministic) $\gamma_{sat}/\gamma_{unsat}$ [kN/m³]	Cohesion (Lognormal)		Friction Angle (Lognormal)	
		Mean [kPa]	Std. dev. [kPa] (COV)	Mean [°]	Std. dev. [°] (COV)
Silty clay	15.7 / 15.7	2.7	0.5 (19%)	29.9	0.8 (3%)
Organic clay	15.2 / 15.2	3.4	1.0 (29%)	26.5	1.1 (4%)
Peat	10.2 / 10.2	5.0	0.3 (7%)	13.7	0.3 (2%)
Sandy clay	16.2 / 16.2	3.7	0.6 (16%)	28.2	0.7 (2%)
Sand	20.0 / 18.0	-	-	35.0	1.8 (5%)

Note that this schematization can differ in various aspects from those used in actual safety assessments conducted by the water authorities. For example, a less detailed soil schematization and different soil strength parameters are used. The case studies aim to provide insights using various schematizations of dikes, rather than estimating a single best failure probability.

Load variations

The fragility curves are constructed conditional on load scenarios, which comprise a combination of 1) the phreatic line, including the water level acting against the dike and head levels in the dike body, and 2) the pore-water pressures in the underlying aquifer. Pore-water pressures are assumed hydrostatic up to the first cover layer above the Pleistocene sand, and linearly interpolated from there to the top of the aquifer. There are no roads on the dikes, and external traffic loads are not considered in the analysis. These load components have two-dimensional aspects. These are represented by a parameter for the height of the phreatic line, which serves as the load variable for the conditional failure probabilities. Therefore, the height of the phreatic line is defined as the average head in the dike body and is expressed in m+NAP, the Dutch vertical reference system. Appendix D.1 contains statistics on water levels and head at the inner crest used for the load scenarios. The underlying assumptions of these scenarios are discussed here.

For the load scenarios at canal dikes, it is assumed that high canal water levels and high head levels are strongly correlated and occur simultaneously, as both result from heavy rainfall (Lendering et al., 2018; Strijker & Kok, 2023). Water level statistics were derived using log-linear interpolation between target levels and normative water levels with a 100-year return period, received from the water authorities. The decimate heights are 15 cm at Aarlanderveen and 7.5 cm at Duifpolder. The statistics of head levels from the inner crest and further down the dike profile were based on measurements and time series models (Strijker & Kok, 2025, Bakker & Schaars, 2019) and are shown in the right-hand graphs of Figure 5.6, with decimate heights under 10 cm. Heads between the canal and the inner crest are linearly interpolated.

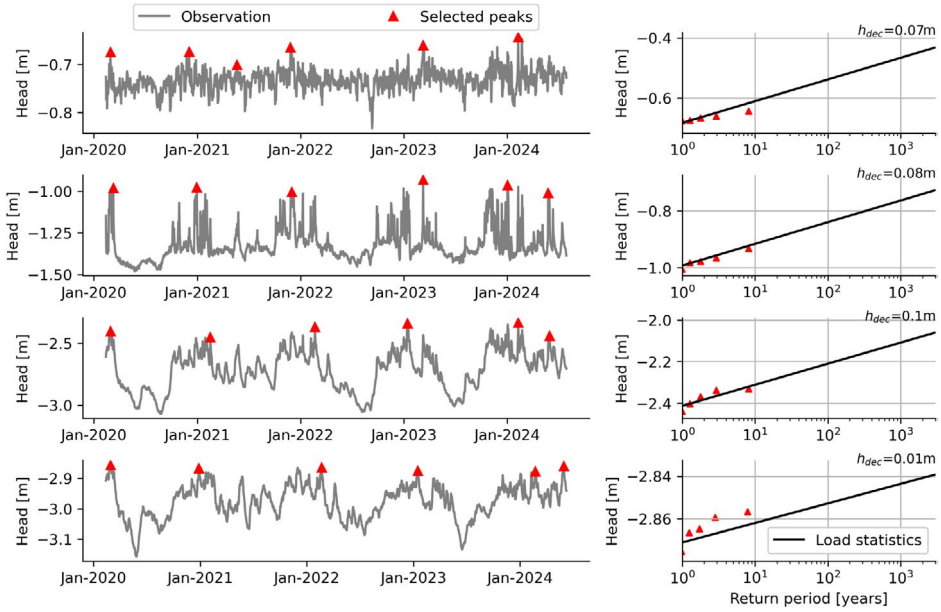


Figure 5.6 Time series of the observed head level at Duifpolder with selected peaks (▲) at the inner crest (top), talud zone (middle) and toe (bottom). Right: plot positions of the selected peaks with extrapolation lines. Red triangles in the right and left panels mark the same peaks, but axis scaling differs.

For the load scenarios at the river dikes, the phreatic line and pore-water pressures under the dike were schematized using the official flood risk assessment toolkit for primary flood defences (Slomp et al., 2016; TAW, 2004). This toolkit includes the Waternet Creator, which models water pressures based on parameters, such as the maximum water level, dike geometry, leakage lengths and subsurface conditions (Van der Mei, 2023). For the river dikes, the water level statistics were based on the probabilistic tool Hydra-NL, which provides water level statistics across the Netherlands

(Geerse, 2011). Subsurface conditions differ between the two river dikes. The Meers dike consists of a dike body of clay overlying a sand layer, while the Gennep dike has a clay body with an underground of clay and peat layers on top of a sand layer. As a result, pore-water pressure cannot build up in the sandy aquifer at Meers, but can at Gennep. The leakage length of the sandy aquifer at Meers, expressing the distance over which the head decays to about 37%, is about 180m. This results in high pore water pressures at the toe of the dike. These pore-water pressures can cause uplift (Van et al., 2005), where the hydraulic head in the sand layer is maximised by the weight of the uplifted overlying layer.

For all cases, the fragility curves were set up using at least five load scenarios, ranging from yearly occurring levels to levels with a probability of 1/10,000 per year. Adjacent scenarios differ by at most a factor of ten in probability. The fragility curves were constructed using log-linear interpolation of the conditional failure probabilities between the load scenarios. Nonlinearity checks were performed, and additional scenarios were added where needed. The overall failure probability was then obtained by combining the conditional failure probabilities with the load distribution and integrating over all load levels.

Survived load levels in reliability updating

Extreme events during monitoring offer valuable data and much deeper insight into dike reliability than annual occurring load levels. For instance, during the 2021 Meuse flood, water levels occurred with an exceedance probability smaller than 1:100 per year (Strijker et al., 2023). However, this study focuses on how failure probability changes across dike types and loading variations. To maintain consistency and eliminate random variability, two load levels were selected for each case with fixed annual probabilities of 1/2 and 1/10 per year, respectively called T2 and T10 events. These represent typical yearly occurring and more extreme events, helping to demonstrate the added value of observing both.

5.3 Results

To answer the research question, the results of two analyses are discussed. First, the extent to which failure probabilities can be reduced using reliability updating for different load and strength configurations is conceptually assessed. Secondly, reliability updating is applied to the case studies involving four dikes in the Netherlands. This section finishes with a reflection on balancing strength uncertainties to load variation for assessing credible failure probabilities. Unless stated otherwise, all failure probabilities are expressed as annual probabilities.

5.3.1 Conceptual analysis on reliability updating

To understand the role of load variations in reliability updating, a conceptual analysis was performed using hypothetical dikes with a range of combinations between load variations and strength uncertainty. The impact of reliability updating is expressed using the probability factor, defined as the ratio of prior to updated failure probability ($P_{f,prior}$ / $P_{f,upd}$). It was found that the largest reduction in failure probability occurs when the strength uncertainty is high and the load variations are low (top right corner in Figure 5.7). Even in cases of moderate strength uncertainties, reliability updating can lead to significantly smaller failure probabilities when the decimate height is small. This is because, under small load variations, the range of loads contributing to the failure probability is narrow and concentrated around relatively frequent load levels. This is illustrated using two cases with identical fragility curves and prior failure probabilities, but different load variations, indicated with red and blue dots in Figure 5.7. For the two cases shown in Figure 5.8, the prior failure probability is 1/100 per year, while the updated probabilities are 1/960 per year (red, small load variations) and 1/140 per year (blue, large load variations). This results in probability factors of 9.6 and 1.4, respectively. In both cases, reliability updating causes a sharp drop in the conditional failure probability near the survived load level (top row in Figure 5.8). When the load variations are small, this range contributes much more to the failure probability than in cases with larger load variations. As a result, reliability updating has a larger impact on the failure probability under small load variations.

To assess when reliability updating is most effective, sensitivity analyses were performed by varying survived load levels and prior failure probabilities (see Appendix D.2 and D.3). The key finding is that the impact of reliability updating is independent of the prior failure probability. Regardless of whether the prior failure probability is high or low, the impact of updating depends primarily on the shape of the failure probability distribution, which is influenced by load variations and strength uncertainties, rather than by the absolute level or relative distance between them

5.3.2 Case study results

The insights from the conceptual analysis were verified using four realistic case studies of canal and river dikes with varying strength uncertainties and load variations. Their position relative to the conceptual parameter space is illustrated in Figure 5.7 by the transparent blocks. The calculated probability factors for the case studies are compared with the so-called conceptual factors, defined as the expected ratio $P_{f,prior}$ / $P_{f,upd}$ for given combinations of decimate height and fragility-curve inverse gradient, based on Gumbel distributions. The results are discussed, beginning with the canal dikes, followed by the river dikes.

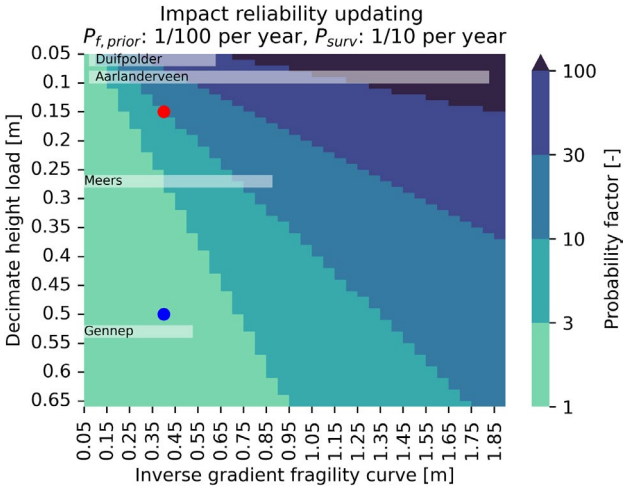


Figure 5.7 Impact of reliability updating across different load variations (decimate height) and strength uncertainties (inverse gradient fragility curve). Transparent blocks indicate the ranges of load variations and strength uncertainties of the case studies considered, and red and blue dots mark two cases of which the fragility curves, load statistics and resulting failure probability distribution are shown in Figure 5.8.

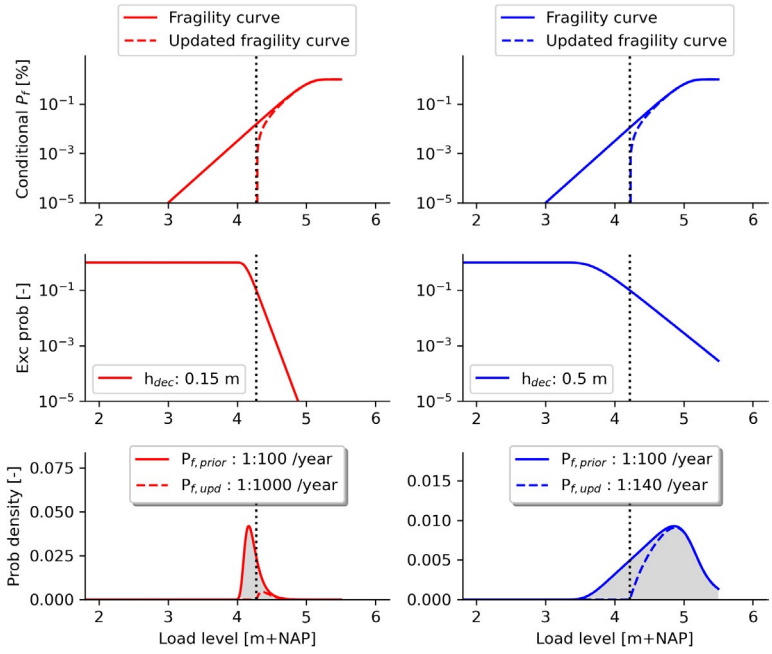


Figure 5.8 Two combinations of load statistics (red: small load variations, blue: large load variations) and the same prior fragility curves (inverse gradient of 0.4m). The vertical dotted line indicates the survived 10-year load level. Both cases have the same prior failure probability, while the updated probabilities differ.

Canal dikes

Based on the conceptual analysis, it is expected that the failure probabilities of the canal dikes can be significantly reduced through reliability updating, since the load variations are small with decimate heights of the phreatic line up to 10cm. For each dike, three different strength uncertainty scenarios were considered. The resulting prior and updated fragility curves (T2 and T10) for Aarlanderveen are shown in Figure 5.9. As strength uncertainty increases, fragility curves become flatter with higher inverse gradients ranging from 0.08m to 1.79m, as indicated in the titles. In the middle panel, the inverse gradient for the moderate-strength case is taken at the load level where failure is most likely. This point lies at a phreatic line of roughly –0.75 m and is marked by a grey solid circle. After updating, the fragility curve is capped at the survived load level and gradually converges back to the prior curve. This offset is largest under high strength uncertainty, leading to the greatest reduction in failure probability. The updated fragility curve depends only on strength uncertainty and the survived load level, while the final impact on failure probability also depends on the load statistics. Given the small load variation of the canal dikes, how much are the failure probabilities reduced?

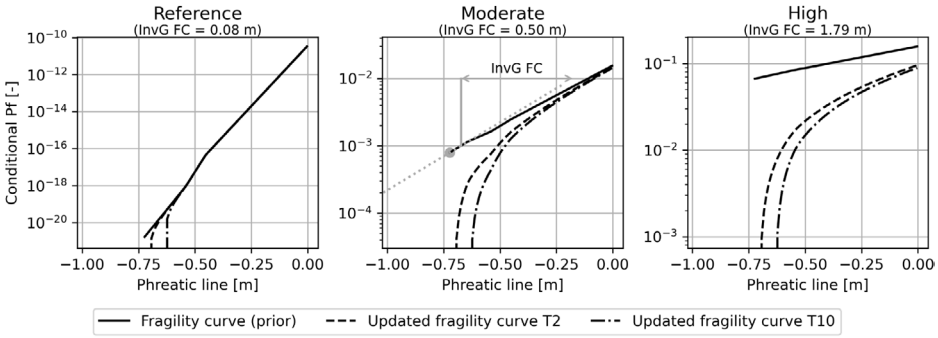


Figure 5.9 Fragility curves at Aarlanderveen for three different strength uncertainty scenarios (reference, moderate and high), including both the prior and two updated fragility curves, based on two different survived load levels (T2 and T10).

For the moderate strength uncertainty scenario and a T10 survival load, reliability updating reduces failure probability by a factor of 10.0 at Duifpolder and 42.4 at Aarlanderveen. Although Duifpolder has a smaller decimate height, its fragility curve is steeper (lower inverse gradient), which limits the impact of updating compared to Aarlanderveen. These outcomes align with the conceptual analysis, as shown in Table 5.2, where probability factors are summarised for all strength scenarios and survival loads for both the conceptual analyses and case studies. In the reference strength-uncertainty scenario, the inverse gradient of the fragility curve is small relative to the load decimate height, so reliability updating does not reduce the updated failure

probability. While the case study results closely match the conceptual factors, small differences arise. These differences can be explained by the fact that the fragility curves and load statistics of the case studies do not follow statistical distributions, as used in the conceptual analysis. In practice, both the decimate height and the inverse gradient of the fragility curve vary across the range of phreatic line levels.

Table 5.2 Prior and updated annual failure probabilities for the two canal dikes, each with three different fragility curves. Reliability updates are based on T2 and T10 survival loads, with reductions expressed as probability factors and compared to conceptual factors (in italic and shown in brackets).

		Aarlanderveen (h_{dec}=0.09m)			Duifpolder (h_{dec}=0.06m)		
Strength uncertainty scenario *Inverse gradient fragility curve [m]*		Reference *0.08*	Moderate *0.50*	High *1.79*	Reference *0.08*	Moderate *0.10*	High *0.61*
Prior failure probability		<1/1,000,000	1/1,300	1/18	1/47,000	1/120	1/8
T2	Updated failure prob	<1/1,000,000	1/12,000	1/580	1/47,000	1/370	1/130
	Probability factor *(conceptual factor)*	**1.0** *(1.0)*	**9.3** *(7.7)*	**31.8** *(30.7)*	**1.0** *(1.0)*	**3.0** *(1.9)*	**16.3** *(14.8)*
T10	Updated failure prob	<1/1,000,000	1/57,000	1/2,700	1/47,000	1/1,200	1/440
	Probability factor *(conceptual factor)*	**1.0** *(1.0)*	**42.4** *(31.8)*	**149.1** *(162.5)*	**1.0** *(1.0)*	**10.0** *(4.1)*	**54.3** *(70.7)*

River dikes

For the river dikes, the phreatic lines exhibit larger variations than in the canal dikes, with decimate heights around 30cm and 50cm. This leads to a total range of about two meters, covering load scenarios that occur on average once per year up to once every 10,000 years. As a result, the fragility curve span a much wider range of load levels, as shown in Figure 5.10, compared to about 70cm at Aarlanderveen. Despite this, the inverse gradients of the fragility curves are within a similar range as those of the canal dikes. The updated fragility curve reduces the conditional failure probability for several decimeters above the survived load level, although this effect diminishes at higher phreatic levels.

For the river dikes, the impact of reliability updating is minimal and updated failure probabilities are nearly identical to the prior values, see Table 5.3. Even for the Gennep dike, which has a relatively high prior failure probability of 1/32 per year under the high strength uncertainty scenario, a survived load level with a 10-year return period provides little new information due to the large load variation. Even more extreme events, such as the 2021 Limburg floods with a return period of 100 years (Strijker et al., 2023), add little value when updating failure probabilities. For example, the Gennep dike's failure probability under T100 conditions is only reduced by a factor of three to seven, depending on the strength uncertainty scenario. This shows that surviving extreme loads doesn't guarantee a significant reduction in failure probabilities, as the impact depends on how the load varies.

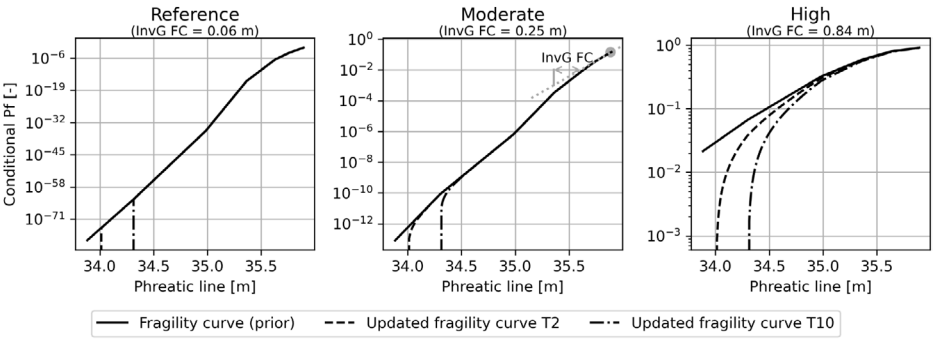


Figure 5.10 Fragility curves at Meers for three different strength uncertainty scenarios (reference, moderate and high), including both the prior and two updated fragility curves, based on two different survived load levels (T2 and T10).

Table 5.3 Prior and updated annual failure probabilities for the two river dikes, each with three different fragility curves. Reliability updates are based on T2 and T10 survival loads, with reductions expressed as probability factors and compared to conceptual factors (in italic and shown in brackets).

		Meers (h_{dec}=0.27m)			Gennep (h_{dec}=0.53m)		
Strength uncertainty scenario *Inverse gradient fragility curve [m]*		Reference *0.06*	Moderate *0.25*	High *0.84*	Reference *0.49*	Moderate *0.06*	High *0.27*
Prior failure probability		<1/1,000,000	1/240,000	1/24	1/49	1/49	1/32
T2	Upd. failure prob	<1/1,000,000	1/240,000	1/59	1/49	1/49	1/32
	Probability factor *(conceptual factor)*	**1.0** *(1.0)*	**1.0** *(1.1)*	**2.5** *(2.4)*	**1.0** *(1.2)*	**1.0** *(1.0)*	**1.0** *(1.3)*
T10	Upd. failure prob	<1/1,000,000	1/240,000	1/110	1/49	1/49	1/35
	Probability factor *(conceptual factor)*	**1.0** *(1.0)*	**1.0** *(1.6)*	**4.7** *(5.7)*	**1.0** *(1.0)*	**1.0** *(1.0)*	**1.1** *(1.8)*

In conclusion, the case studies confirm the findings from the conceptual analysis. The calculated probability factors in the case studies lie close to the ones calculated by the conceptual analysis. Therefore, the results of the conceptual analysis can be seen as a first approximation for estimating the impact of reliability updating under different load and strength characteristics.

5.3.3 Balancing strength uncertainties to load variations

To prevent incredible failure probability estimates, a credibility check is proposed. This check helps identify situations where the strength uncertainty is disproportionately high compared to the load variations. Specifically, it targets situations where the failure probability of a dike is mainly influenced by load levels that occur every 10 years or more frequently, and these frequent loads contribute more than 50% to the total failure probability. According to conceptual analysis results, such situations typically arise when the inverse gradient of the fragility curve is at least about 1.5 times greater

than the decimate height of the load, as shown in Figure 5.11. These situations may appear unlikely in dike safety, particularly when the dike has demonstrated stability over many decades. In such cases, the load variations and strength uncertainty are out of balance, leading to unrealistic strength uncertainty compared to load variations, and consequently, incredible failure probabilities. Reducing the uncertainties related to the dike's strength will lower the inverse gradient of the fragility curve, resulting in more credible failure probability estimates. This can be done by, for example, monitoring load levels in combination with reliability updating methods, or gathering more detailed data on soil properties to obtain a clearer and less uncertain assessment of the dike's strength.

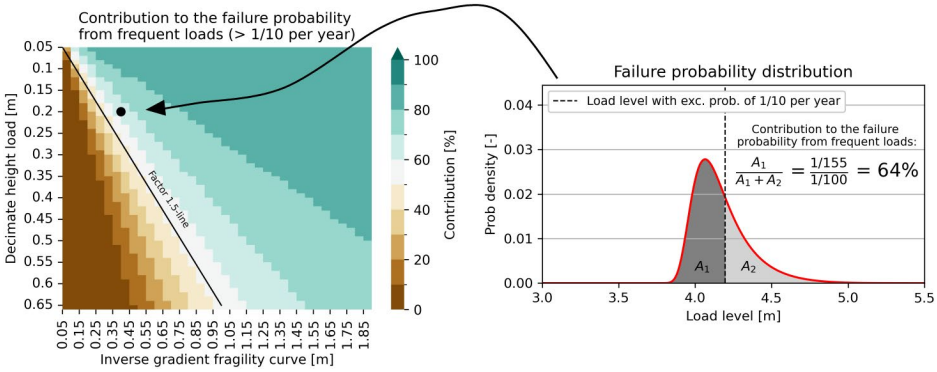


Figure 5.11 Left: heatmap showing the contribution of load levels that occur frequently (>1/10 per year) to the failure probability across different load variations and strength uncertainties. The factor 1.5-line shows where the inverse gradient is 1.5 times larger than the decimate height. Right: an example of a hypothetical dike with a decimate height of the load of 20 cm and a fragility curve with an inverse gradient of 40 cm, where frequent loads contribute 64% to the overall failure probability.

5.4 Discussion

5.4.1 Implications

This section reflects on the way the findings of this study can help estimate probabilities of dike instability in practice. It starts with identifying regions in the Netherlands where dike failure probability estimates can potentially be improved using reliability updating, based on information about the load variations. It finishes with a call to consider load variations as stochastic variables, rather than design values, even when the variations are small.

Regions in the Netherlands with small load variations

One of the key characteristics of the over 10,000km of Dutch canal dikes is small load variations in both the water level and the phreatic line (Strijker & Kok, 2025). Furthermore, there are many primary flood defences across the Netherlands with a large

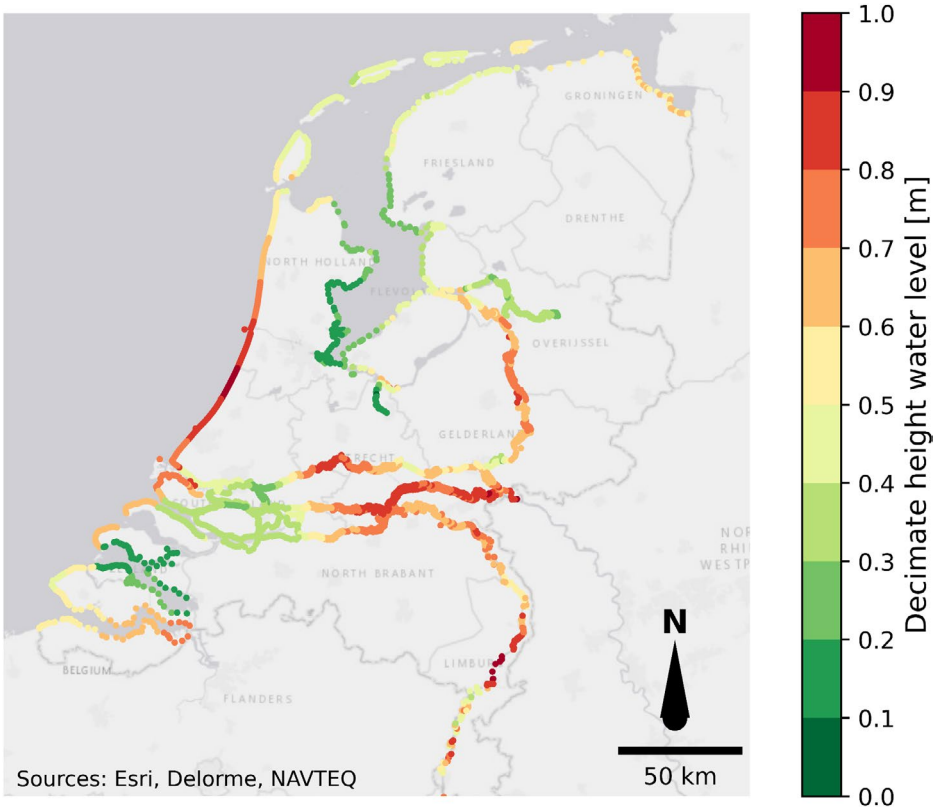


Figure 5.12 The decimate height of the water levels along the primary dikes in the Netherlands, based on data from TMR2006 (Bouw 2008).

variety in load variations. The variety in water level statistics is available throughout the Netherlands and is calculated using a probabilistic tool called Hydra-NL, which combines physical and statistical models (Geerse, 2011). As can be seen in Figure 5.12, the decimate heights of water levels along the primary dikes in the lake and delta areas are the smallest. These are areas where water levels are regulated by weirs, pumping stations, and storm surge barriers, which reduce variations in water levels. Therefore, these controlled systems with small water level variations offer advantages: primary dikes are exposed to lower absolute loads than in uncontrolled systems, and the limited variation helps to assess actual dike's strength and performance, since extreme levels are close to those occurring annually.

The value of considering the load probabilistically

In many geotechnical problems, the uncertainty in the strength dominates the reliability estimates, while the influence of loads is relatively low (European Commission

et al., 2024). Therefore, it can be practical to treat the load side of the problem semi-probabilistically using design values. This approach is also applied in the current safety assessment of Dutch canal dikes, where load variations are relatively small. However, when assessing failure probabilities, loads can provide insight into the schematization of the strength. Considering the load probabilistically helps to realistically estimate stochastic strength variables through reliability updating. It also provides insight into failure processes, such as the sensitivity to load variations that can trigger failures. Importantly, dike instabilities occur due to variables that change over time, even when their fluctuations are small. In safety assessment, time-variant and time-invariant, inherently or epistemically uncertain, are often treated similarly probabilistically. However, in reality, the time-variant variables trigger failure, and we must keep these variables in mind in safety assessments. Therefore, it is argued that time-variant loads should be considered as probabilistic variables in full probabilistic calculations, and analyse the failure probability contribution of various load scenarios.

5.4.2 Limitations and future research

Model uncertainties

The approach in this study did not include any model uncertainties. Model uncertainties were excluded to avoid introducing additional complexity and to allow for a clearer interpretation. However, there are model uncertainties associated with, for example, modelling the stability and estimating the pore-water pressures. In practice, these model uncertainties tend to be compensated using conservative and safe assumptions in safety assessments, like a safe assumption for the phreatic line in the river dikes or strength modelling (European Commission et al., 2024). Nevertheless, adding these overlooked model uncertainties is certain to increase the prior failure probability, but their influence on the updated failure probability is less straightforward, and a qualitative reasoning is given. Adding additional uncertainties in the strength, such as model uncertainty related to stability modelling, results in a fragility curve that shifts upward and becomes flatter. This, in turn, is likely to increase the impact of reliability updating, resulting in larger probability factors. Adding model uncertainty to the schematization of the load does not necessarily change the decimate height; it is more likely to affect only the absolute level. This does not affect the impact of reliability updating in terms of the probability factor, but it does increase the prior and updated failure probability. Furthermore, this study assumed a clear survived load level for reliability updating, which can be in practice rather uncertain, because of measurement errors and the translation from point measurements to spatial variables, such as the phreatic line. Accounting for these uncertainties, rather than using a best estimate, will reduce the impact of reliability updating.

Multivariate load variations

Failure probabilities were estimated using fragility curves, which express the relation between load and conditional failure probability. In the case studies, we only considered one load variable, namely the pore-water pressures in and under the dike. However, the load can also be a combination of pore-water pressures and traffic load. In these cases, a fragility plane can be used where the failure probability is conditional on two variables (Nofal et al., 2020). This can make failure probability analysis more challenging, as well as reliability updating. This traffic load is time-variant and, therefore, inherently uncertain (Lendering et al., 2018). The way this load variable is considered within failure probability analysis, deterministic or probabilistic, affects the potential of reliability updating. In future extensions, attention should be given to the dependency between loads, such as time-varying pore-water pressures and traffic load, as well as to the incorporation of multiple monitoring indicators.

Modelling shear strength and the value of historic data

When it comes to the reliability of dikes, the behaviour of dike material is never constant over the long term. Soils are subjected to weather and other forces of the environment, making soil characteristics vary over time. For example, strength parameters can evolve due to factors such as soil aging, dry-wet cycling, and human interventions like maintenance or reinforcement efforts. Additionally, load characteristics can change, because of climate change with increasing extreme precipitation and more frequent high water levels (Van Dorland et al., 2023). This dynamic nature of load and strength means that failure probabilities and dike safety change over time. Whether this has a positive or negative impact on dike stability depends on site-specific dike characteristics. Ultimately, the challenge to reliability updating lies in finding the right balance between understanding evolving strength and load parameters and leveraging historical measurements. Regular monitoring, combined with probabilistic methods that incorporate past data, helps create a clearer picture of the dike's true performance.

5.5 Conclusions

This study aims to better understand the role of load variations in reliability updating and in assessing credible failure probabilities. First, this was conceptually examined using hypothetical dikes with various load and strength configurations described by Gumbel distributions. It was found that the impact of reliability updating increases when load variations are smaller, regardless of the prior failure probabilities. The reason is that only a relatively small load range contributes to the calculated failure probability, which, in absolute terms, is close to the load levels that occur every year. By updating the fragility curve, the conditional failure probabilities closest to the survived load level are reduced the most. Therefore, reliability updating has a larger impact on the failure

probability for situations with smaller load variations. Second, these findings were confirmed in the case studies of two canal dikes and two river dikes in the Netherlands, which varied in terms of fragility curves and load variations. The two canal dikes with decimate heights smaller than 10cm benefited the most from reliability updating, while the updated failure probabilities of the river dikes changed only slightly. This is because the load variations of the two river dikes are larger with decimate heights of the phreatic line of approximately 30 and 60 cm, while the inverse gradients of the fragility curve were similar. In the moderate-strength scenario for the canal dikes, reliability updating with a load that occurs once every two years reduces the failure probabilities by a factor of three to nine, relative to the prior estimates. However, for the river dikes, even extreme load levels with a 1/100 probability per year, as occurred during the 2021 Limburg floods, contribute little when it comes to improving failure probabilities through reliability updating.

Based on the conceptual analysis, a credibility check for safety assessments was proposed, identifying situations where the failure probability is dominated by load levels that occur every 10 years or more frequently. The contribution of frequent load levels to the total failure probability exceeds 50% when the inverse gradient of the fragility curve is about 1.5 times greater than the decimate height of the load. In these situations, the load variations and strength uncertainty are out of balance, making it challenging to estimate credible failure probabilities, as this can contradict the stability of the dikes over many decades. Situations of small load variation occur widely in the Netherlands, for example, along much of the over 10,000 kilometres of canal dikes and in primary flood defences in lake and delta regions. In these situations, reducing the uncertainties related to the dike's strength will lower the inverse gradient of the fragility curve, resulting in more credible failure probability estimates. This can be achieved by, for example, monitoring load levels in combination with reliability updating methods or gathering more detailed data on soil strength.

Code availability statement

The script used to produce example cases, key figures illustrating the methodology, and the main results of the conceptual analysis are available from 4TU.ResearchData at: https://doi.org/10.4121/acd1f26c-15c4-47f6-b01a-0120d104031a

It includes the construction of fragility curves and the direct reliability updating procedure described in the manuscript.

6

Observations and insights from practice

During my research activities, I actively applied scientific insights to practical projects for water authorities, and in turn, real-world challenges inspired and refined my research. This chapter describes a selection of observations and insights from practice, which arose from projects carried out on behalf of water authorities, STOWA, and Rijkswaterstaat. These observations and insights are relevant to this dissertation because they focus on the safety of canal dikes and the impact of precipitation on hydraulic heads within dikes. Additionally, they can offer useful guidance for future research. The next sections summarise findings from the studies listed under "Selection of other projects" in the list of publications, which offer a more detailed explanation of the work. At the start of each section, a reference is made to the study or studies it builds on.

6.1 Lessons from the wet winter of 2023-2024: damages and head observations at canal dikes

This section is based on Strijker (2025a).

6.1.1 The wet winter and noticed damage

The winter of 2023–2024 was exceptionally wet in the Netherlands. The monthly and multi-month precipitation amounts were exceptional on a national scale, while the daily and multi-day extremes were more localised. The events with heavy precipitation within several days are particularly relevant for canal dikes, where shallow groundwater levels respond rapidly to heavy precipitation but also dissipate relatively quickly, as the groundwater systems of dikes tend to have a relatively short memory. For water authorities in the central and southwestern parts of the country, the estimated return periods of precipitation amounts in one to eight days reached up to 100 years, as shown in Figure 6.1.

The impact of this wet winter on canal dikes was assessed by contacting all water authorities with canal dikes and gathering information on damage incidents and crisis management specifically related to these dikes. **This inventory revealed five significant cases of damage to canal dikes during the winter within the water authorities Hoogheemraadschap van Delfland and Rivierenland. These damages include shallow inner-slope slides, subsidence, and leaks, but did not result in major dike safety issues or breaches.** These two water authorities also experienced the highest precipitation totals over one- to four-day periods. At the same time, most dikes showed no problems at all, which potentially offers valuable insight into their actual strength. Figure 6.2 illustrate the damage at one location, where cracks in the crest were formed due to deformations and the ditch of the dike was filled with sand to reinforce the dike.

Furthermore, the inventory revealed that the water authorities signal flood situations at canal dikes using canal water levels. If water levels are high, inspections of canal dikes can be conducted, and, where necessary, emergency measures can be implemented. At known problem spots, sandbags were placed at bottlenecks in the drainage canals, such as bridges that obstruct the water flow and raise the water levels upstream. However, in general, the canal water levels are often strongly regulated and were even during these heavy precipitation amounts manageable and under control. Nevertheless, damage occurred, indicating the importance of precipitation above canal water levels. **This is a clear shortcoming of the current method of signalling hazardous situations, which can be addressed by incorporating the signalling of heavy precipitation events and their effects on groundwater levels using time series models, as developed in this dissertation.**

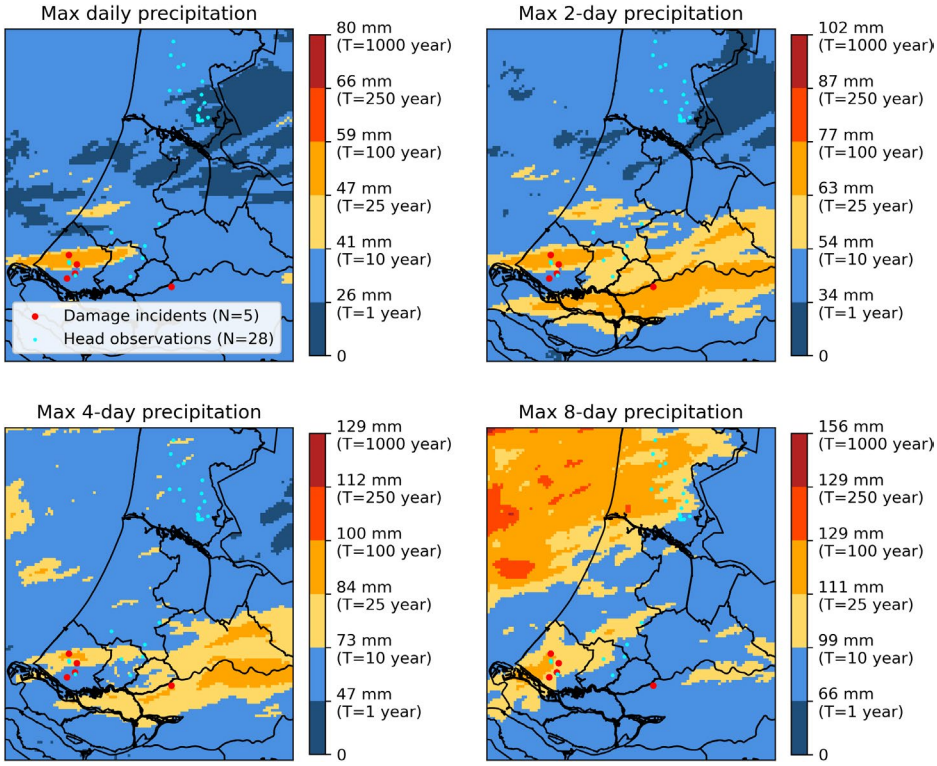


Figure 6.1 Maximum precipitation totals for various durations during the winter of 2023–2024, with the colour scale reflecting the corresponding return periods of precipitation amounts. The dots indicate the considered monitoring sites, and the five red dots mark the sites where damage occurred. The legend is only shown in the top-left graph.

Figure 6.2 Settlement and crack formation along the regional flood defences near the N468 between Schipluiden and Maasland (left). The ditch next to the dike has been filled with sand to reinforce the dike (right). Pictures from the water authority of Delfland.

6.1.2 *Typical schematizations of phreatic lines compared to head observations*

When assessing the safety of a canal dike, the phreatic line is, in practice, predominantly determined by predefined guidelines. This is done because, in most situations, no head measurements are available. These so-called schematizations describe the phreatic line based on several dike profile points. In principle, these schematizations are used as a conservative assumption. When a dike is assessed as unsafe, this assumption can be refined if it has the potential to improve the estimated stability. In practice, refinement often does not take place, and prescribed guidelines are also used for design purposes. The schematization of the phreatic line is based on interpolating head levels between various profile points. The height of the head levels at various profile points varies across water authorities, and four different schematizations of four water authorities are shown in Table 6.1. The way these different schematizations work out is illustrated using a canal dike at the polder Aarlanderveen in Figure 6.3. There is no single highest or lowest phreatic line schematization, as the lines cross each other along the dike profile. The largest differences between schematizations of different water authorities are:

1. The way the phreatic line is schematized in the crest zone, where HHNK uses an offset at the inner crest, which increases with larger crest widths, while at other water authorities, it is often assumed to be equal to the target level or a fixed offset lower than the target level.
2. At HHNK and HHSK, the shoulder berm is used a an additional profile point, while this is neglected at two other water authorities. The phreatic line is capped off at the surface, but adding this point can result in very different phreatic lines.

Table 6.1 The height of the phreatic line at several profile points of a dike. HWL= design High Water Level,

	Intersection point canal and dike	Outer crest	Inner crest	Shoulder berm	Crest berm	Inner toe	Shoulder dtich	Intersection point ditch and dike
HHNK	HWL	HWL	HWL-*off*	SL-0.1m	SL-0.1m	SL-0.1m	PWL	PWL
HHSK	HWL	TLC	TLC	SL-0.2m	-	SL-0.2m	-	PWL
Rijnland	HWL	-	TLC	-	-	SL	-	PWL
Delfland	HWL	TLC	TLC-0.2m	-	-	SL	-	PWL

*TLC=Target Level Canal, SL = Surface level, PWL=Polder Water Level, *off*=max((1:10)*width crest,0.5*(TLC-PWL))

During the wet winter, head levels in canal dikes were measured at 111 monitoring points through standpipes across 28 monitoring locations. These observations were analysed. The maximum observed head levels in the crest zone were compared to the

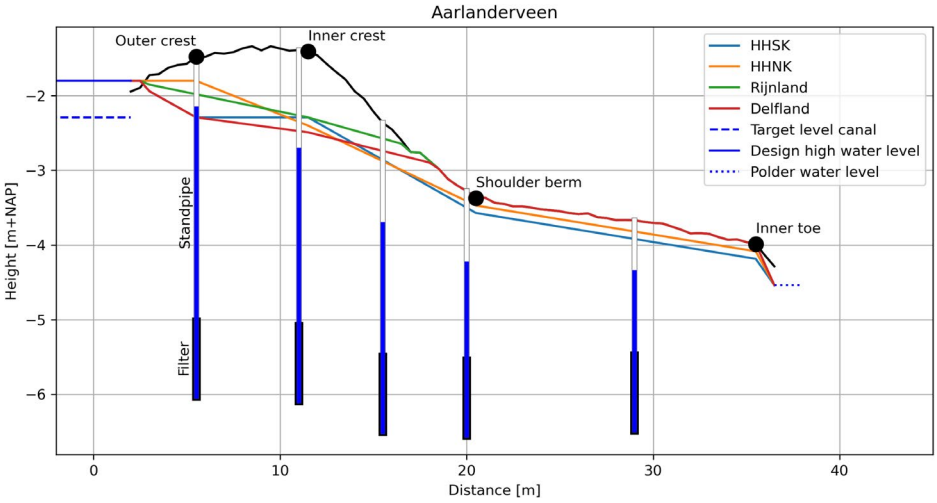


Figure 6.3 The phreatic line through a canal dike, based on guidelines from various water authorities. The blue vertical bars within the standpipes indicate the maximum measured head level during the winter 2023-2024.

target level in the canal, and head levels in the talud zone were compared to the surface level, since these levels are used to schematize the phreatic lines in canal dikes, see Figure 6.4. **In general, the maximum measured head levels within the cross-sections show a wide range of variations: from a nearly fully saturated dike to groundwater levels two meters below the surface**. At the crest, maximum groundwater levels are often higher than the outer water level. At the inner crest, groundwater levels reached up to the target canal water level, with a median maximum level 30 cm below it.

Next, all the maximum observed heads at the monitoring points were used to evaluate the schematization of the phreatic line using two guidelines: HHSK and HHNK. Although the conditions that occurred are not necessarily normative, a comparison can still provide relevant information. For each location, the phreatic line is schematized, and at points where standpipes are situated, this schematization is compared to the maximum measured head level recorded during winter 2023-2024. These differences are plotted against the normalised distance of each monitoring point in every cross-sectional profile. The results are shown in Figure 6.5. The following conclusions can be drawn:

- The median of the maximum measured head levels in the cross-section lie mainly below the schematized line, while the 95th percentiles lie mainly above it. This indicates that the current schematization guideline should not be regarded as a conservative first estimate. For a conservative estimate, it is recommended to assume a fully saturated dike.

- In the crest-zone, the head levels can rise above the target level, and the way the phreatic line is schematized, according to HHNK guidelines. This can occur when precipitation is stored in the crest before it is drained away.
- The 5–95% percentile range around the inner crest and upper slope (between 0.75 m and 1.5 m) exceeds 1 m. Such variation significantly influences the stability of the dike and demonstrates that the absolute phreatic level can vary considerably across the canal dike. Furthermore, these observations can be used to estimate the

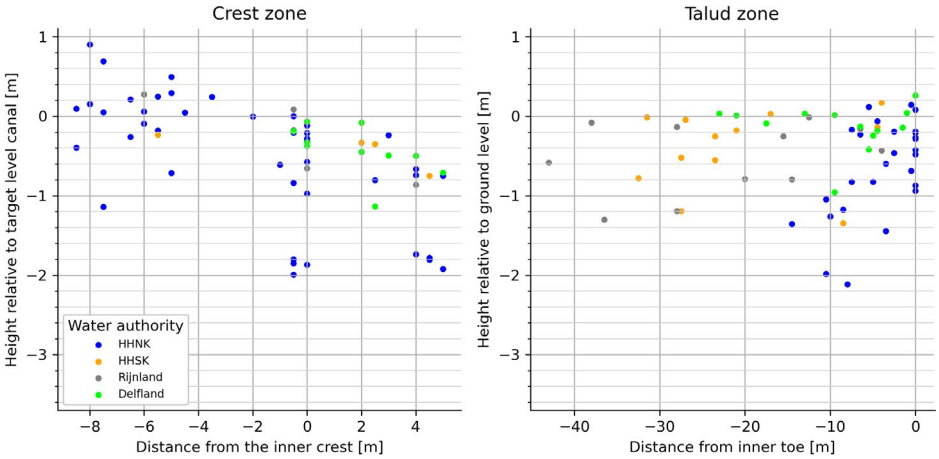


Figure 6.4 The maximum measured head level relative to the target canal water level in the crest zone, shown in relation to the distance from the inner crest where the measurement was taken (left), and maximum measured head level relative to the ground level in the talud zone, in relation to the distance from the inner toe (right).

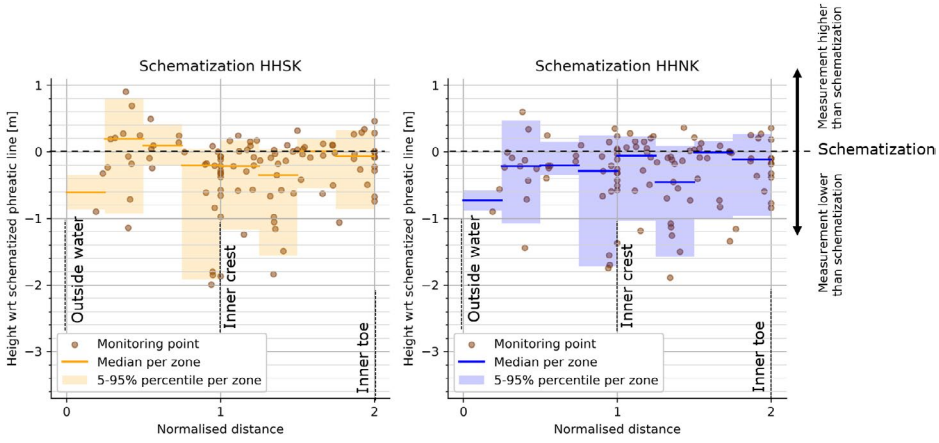


Figure 6.5 The difference between the maximum measured head levels and the schematized phreatic line (left=HHSK, right=HHNK) along the cross-section, normalised as follows: 0 = outside water, 1 = inner crest, 2 = inner toe.

uncertainty of the phreatic line rising from spatial variability, which is an important variable for dike safety, despite limited data to quantify this uncertainty, and is often considered deterministic.

6.2 Schematizing phreatic lines for safety assessments of canal dikes using time series models

This subsection 6.2.1 are results from Strijker & de Snoo (2024a) and Strijker (2023a). Subsection 6.2.2 is based on Strijker (2025a).

6.2.1 Examples of the probability estimate of the phreatic line

In Section 6.1 the head observations during the wet winter 2023-2024 were compared to schematizations of phreatic lines. Although the head observations were the highest recorded at many locations, they likely rise even higher during more severe precipitation events. Therefore, the phreatic line for normative conditions, such as 1/10 or 1/100 probability per year, can be even higher. The phreatic line under normative conditions is an important variable in safety assessments. Therefore, the water authorities of Rijnland and Schieland were interested in a probability estimate of the phreatic line at several monitoring sites. A similar approach was utilised by Strijker & Kok (2025), in which time series models were set up to extrapolate head levels to more extreme events. While Strijker & Kok (2025) quantified the variation in head responses in the talud zone of dikes, this approach was applied to various piezometers at a single monitoring site to estimate the statistics of the phreatic line, using linear interpolation between these points. In total, six monitoring sites were analysed, with five piezometers installed at each site. For each piezometer, a time-series model was set up and forced by long-term precipitation and evaporation time series to produce a long-term head time series, for which an extreme-value analysis was conducted.

For two monitoring sites, the resulting phreatic lines at various probabilities of occurrence are shown in Figure 6.6 and Figure 6.7, for both high and low phreatic lines. Consistent with the analysis of the wet winter 2023-2024's impact on canal dikes in Section 1.1, it was, again, observed that there is significant variation in the absolute level of the phreatic line: the phreatic line of two dikes with, on beforehand, more-or-less similar profiles and subsurface materials can be either near ground level or approximately one meter below it, see Figure 6.6 and Figure 6.7. Furthermore, phreatic lines under normative conditions (in these cases, an exceedance probability of 1/100 per year) are close to the observed groundwater levels that occur annually. **Across the six monitoring sites studied, the phreatic level that occurs on average once every 100 years lies up to 20 cm above the yearly-occurring levels**. Lastly, in some cases, the phreatic level can even drop below the polder water level, causing the ditch to stop

draining and instead allow water from the ditch to infiltrate the dike. This is observed, and the time series models further extrapolate it, but additional non-linear processes that may not be accounted for in the model can limit this.

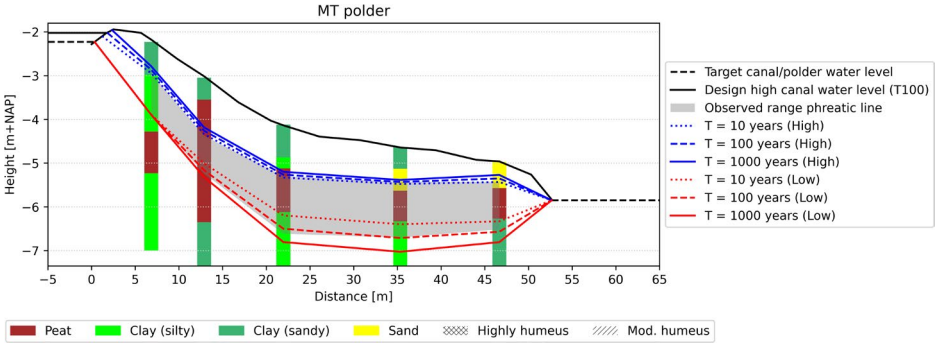


Figure 6.6 Cross-section of the MT polder showing borehole profiles and soil types, with phreatic lines for high and low conditions estimated for different return periods using time-series models

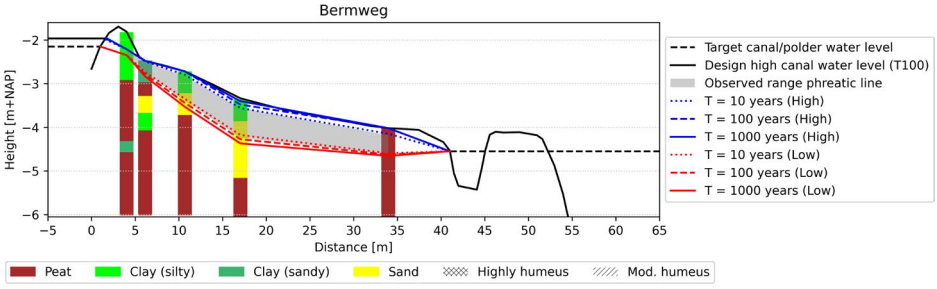


Figure 6.7 Cross-section of the Bermweg showing borehole profiles and soil types, with phreatic lines for high and low conditions estimated for different return periods using time-series models.

6.2.2 The reason for small variations in high phreatic lines.

Strijker & Kok (2025) found that the decimate height of head levels, defined as the increase in level that occurs when the return period increases by a factor of ten, across 35 canal dikes is on average 15 cm. The locations where the phreatic line was analysed, as described in Section 6.2.1, show smaller decimate heights, namely not above 10cm. **Can these decimate heights, ranging from a few centimetres to decimeters, be explained physically and statistically?**

The rise in head levels is primarily driven by infiltration of precipitation into the dike that recharges the groundwater. Since head levels are shallow, this often leads to a relatively rapid increase in head levels, while the stored groundwater drains away more

slowly. The different time-scales of infiltration and drainage can be explained by two main factors:

1. The topsoil is often more permeable and contains structural features caused by swelling and shrinking, which increases the infiltration of water, whereas drainage occurs through deeper, saturated soil layers that are typically less permeable.
2. The hydraulic gradient is steep during vertical infiltration of precipitation, driving rapid downward flow, but becomes much smaller once the water reaches the groundwater table and drains more slowly in the horizontal direction with a smaller hydraulic gradient.

While the drainage time determines which precipitation durations are relevant for head peaks to occur, as shown in Strijker & Kok (2025), the increase of head levels to precipitation is more relevant for the decimate height. The increase in head level in canal dikes to 1mm of precipitation can be quantified using simplified storage models (e.g. $\Delta H = R / ((1–S)\cdot n)$, where R is amount of water that needs to be stored in the soil column, S is the degree of saturation and n is the porosity) or time series analysis of head observations. Strijker & Kok (2024) showed that the head response to 1mm of precipitation in winter, with high hydraulic heads, is typically between 2.5 and 10 mm. Combining this with precipitation statistics for different durations gives an estimate of the increase in heads to more extreme precipitation events and the resulting decimate height of head levels, as shown in Figure 6.8. In this way, the found decimate heights can be substantiated, as they should be approximated by the head response and the increase of precipitation amounts with larger return periods.

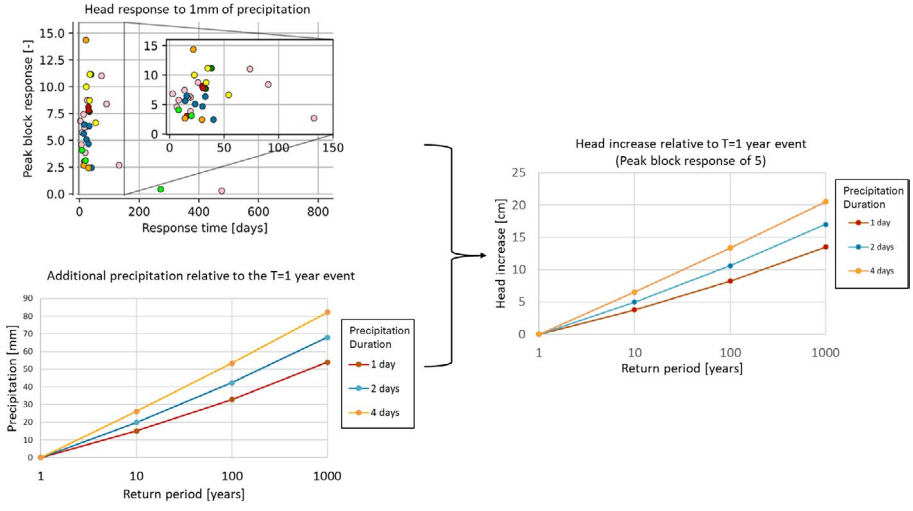


Figure 6.8 The head response, including the head increase per 1 mm of precipitation and the response times, combined with the increase in precipitation for more extreme events (left), results in an estimate of the head increase (right).

6.3 Groundwater drought-indicator to signal droughts based on the exceedance frequencies of groundwater levels in dikes

This section is based on Strijker (2023a).

6.3.1 Signalling drought in dikes through statistics

Water authorities conduct inspections to detect signs of potential failure and, if necessary, take measures. This is done under both normal conditions and hazardous conditions, like during high water levels and prolonged droughts. During hazardous conditions, water authorities have to decide when and where to start inspecting the dikes. For prolonged droughts, meteorological drought indicators are often used, like the precipitation deficit. This indicator appeared to be the most reliable meteorological drought indicator for droughts in canal dikes, for both soil moisture and groundwater levels, outperforming the standardized drought indicators, like *SPEI* and *SPI* (Strijker et al., 2024). Nevertheless, it was also shown that the precipitation deficit has some limitations as an indicator for soil moisture and groundwater levels. The precipitation deficit provides insight into the lack of precipitation and potential evaporation, but it doesn't directly indicate how dry the soil in the dike actually is. The actual level of drying varies between dikes due to geohydrological differences. Two important limitations are:

1. Dikes can start to dry out during the winter months, even when there's little sunshine and low evaporation, simply because the groundwater drains away. This effect isn't captured by the precipitation deficit, which only starts counting from April 1st.
2. After summer ends, the precipitation deficit can remain high for a long time, while measurements show that the dike has already started recovering from the drought. This is due to the incomplete link between precipitation deficit and actual drought conditions within the dike, which are influenced by geohydrological factors.

Therefore, STOWA, Rijkswaterstaat, Hoogheemraadschap Hollands Noorderkwartier (HHNK), Hoogheemraadschap Schieland en de Krimpenerwaard (HHSK) and HKV lijn in Water has developed a new drought indicator specifically for canal dikes: the groundwater drought indicator. **This new groundwater drought indicator provides an estimate of how extreme the groundwater level is in a dike at a specific location, expressed in terms of return periods**. A value of 10, for example, means that the groundwater level at that spot typically occurs on average once every 10 years. This estimate is based on head observations from 34 canal dikes. Using time series models and statistical analyses, the indicator determines how exceptional the drought is at a given location and time. In this way, it reflects the actual dryness within the dike while also accounting for variations in drying across different canal dikes.

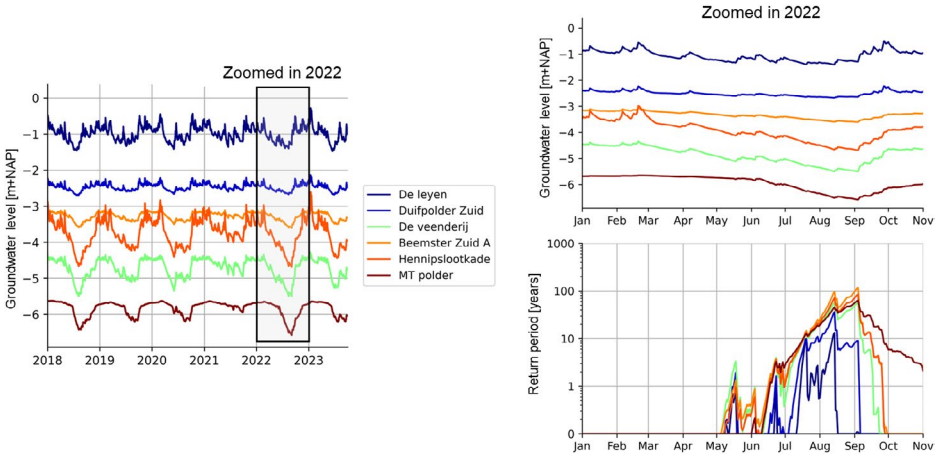


Figure 6.9 Groundwater level time series for several dike sections, showing seasonal fluctuations that also vary from each other. The right panels zoom in on 2022: the upper graph shows groundwater levels, while the lower graph displays corresponding return periods of high groundwaterlevels throughout the year.

6.3.2 Web application to support decision-making for drought inspections

After the conceptual development of the groundwater-indicator, a web application has been developed to provide an up-to-date overview of the groundwater drought indicator for every dike in the Netherlands. The goal is to make the new indicator easily accessible within the organization. This web application was used during the summer of 2025 within the drought protocols to support decisions on organizational scaling (alert phase) and inspection planning. The application also shows recent dry years and allows users to compare the development of this new drought indicator with the traditional precipitation deficit (see figure below).

The groundwater drought indicator is used alongside the precipitation deficit. By combining both indicators, it becomes possible to compare their performance, build confidence in the new approach, reduce the risk of unexpected anomalies or other teething problems, learn from their differences, and support a transparent transition. In addition, HHNK's drought protocols consider a broader set of parameters, including the dike's soil type and known damage patterns.

The development of the groundwater drought indicator is a step toward better understanding and managing the risks that drought poses to canal dikes. It offers an additional tool for obtaining a more direct and nuanced view of actual drought conditions in canal dikes, beyond what traditional meteorological indicators can provide. The indicator is expected to provide a clearer picture of spatial variations in drought across dikes and support better prioritization of when and which dikes require inspection.

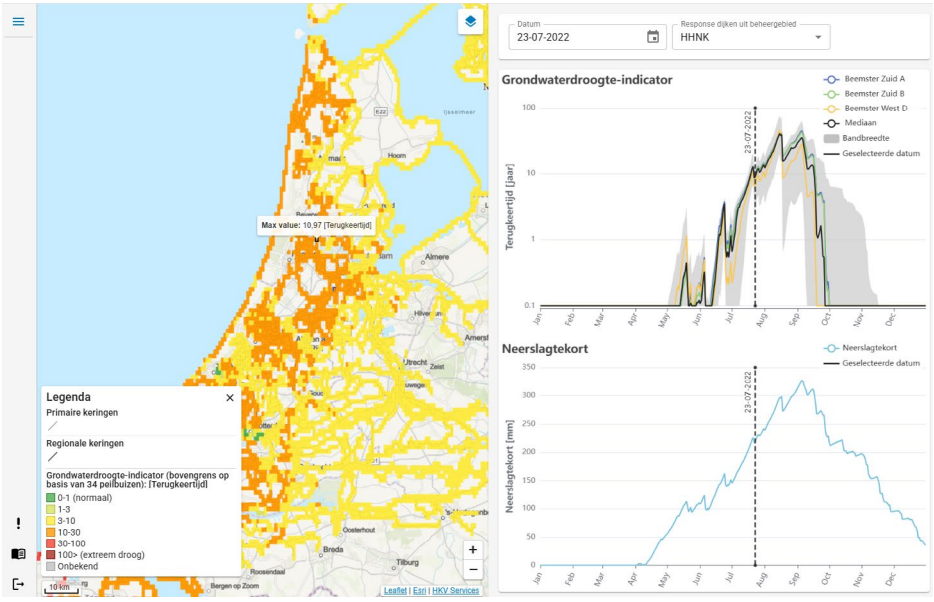


Figure 6.10 Screenshot of the web application. On the left, a spatial overview of the groundwater drought indicator is shown for a selected date, with colors indicating the return period of groundwater levels. On the right, time series of the groundwater drought indicator (top) and precipitation deficit (bottom) are shown for the year corresponding to the selected date. These graphs illustrate the progression of drought over time and help understand the differences between the two indicators.

6.4 The role of precipitation in the inner-slope stability of primary flood defences in the Netherlands

These sections are based on Strijker (2024b) and Strijker (2025b).

6.4.1 Inner-slope instability of primary flood defences

The Netherlands has over 3000 km of primary flood defences of which over 1000km of dike trajectories do not meet the safety standards and need to be reinforced. The highest reinforcement costs arise from dike sections with insufficient inner-slope stability, driven by increased pore-water pressures in and under the dike. The trajectories that do not meet the standards for inner-slope instability are shown in Figure 6.11. Therefore, the schematization of the pore-water pressures in these dikes often plays a key role in the calculated flooding probability.

In practice, the schematization guidelines as described in TAW (2005) are typically used, while some water authorities deviate from them. In the past, the effect of extreme precipitation events was often considered as a load scenario for the primary flood defences in the Netherlands. However, nowadays, the way the phreatic line is

schematized depends on 1) the subsurface of the dikes and 2) the maximum outside water level. This means that, in practice, the probability of the phreatic line solely depends on the statistics of outside water levels, while the effects of heavy precipitation and coincidence with the high water event, as well as the duration of high water levels, are not accounted for. The importance of these effects are illustrated by presenting some measurements in primary dikes along Lake Marken and the Oosterschelde.

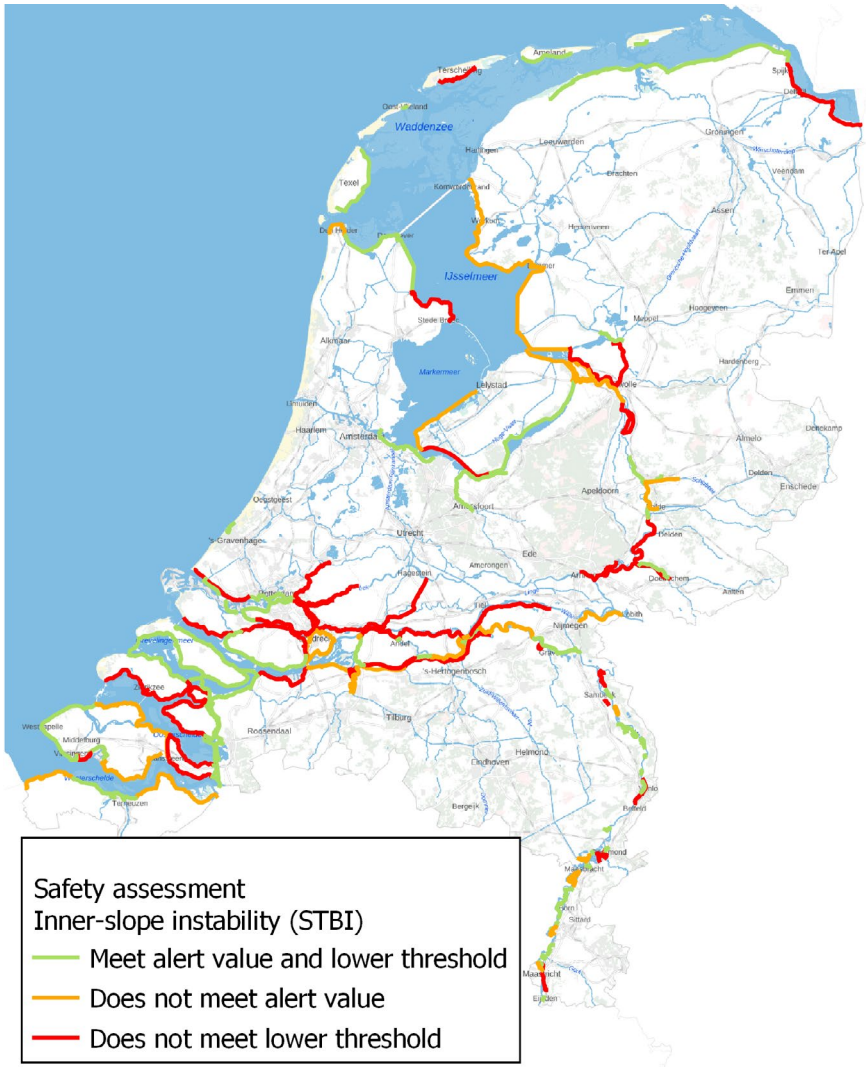


Figure 6.11 A map of primary flood defences where trajectories do not meet the safety standards regarding the failure mechanism, inner-slope instability, where the colours indicate whether it does not meet the alert value or the lower threshold. The lower threshold is the minimum level of protection that the flood defense structure must offer, while the alert value indicates at what point the process to reinforce the dike should start and lies higher than the lower threshold.

6.4.2 *Dikes along Lake Marken*

Lake Marken is located in the centre of the Netherlands and functions as an important freshwater storage. The water levels in the lake can rise due to direct precipitation on the Lake, the river Eem discharging water into the lake, and pumping stations discharging excess water from polders and other water systems into the lake. Water from the Lake can be drained away via Lake IJssel and the Amsterdam-Rhine Canal and the North Sea Canal, depending on the water levels at these systems. During the winter 2023-2024, a prolonged period of precipitation, combined with storm conditions that reduced drainage possibilities at Lake IJssel, caused the water levels at Lake Marken to rise to levels that occur, on average, about every 50–100 years, with peaks occurring in the first weeks of January 2024. During this high water event, the hydraulic heads at several monitoring sites were measured by the water authority Hollands Noorderkwartier. In this section, measurement series at various cross-sections located along Lake Marken were analysed, of which the geographic locations are shown in Figure 6.12.

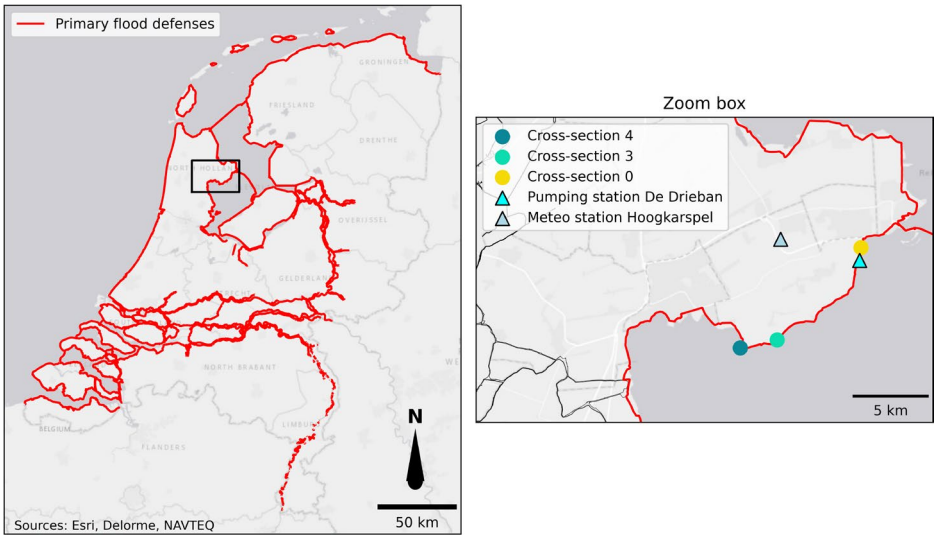


Figure 6.12 Location of the study area within the Dutch primary flood defense system. The zoom box highlights the monitored dike sections along Lake Marken, with nearby pumping and meteorological stations (De Drieban and Hoogkarspel) used for water level and precipitation measurements.

Figure 6.13 shows the dike profiles at these cross-sections (right panel), including the location of the standpipes, as well as the measured hydraulic heads (left panel). As can be seen, the fluctuations in the observed head can be significant and up to two meters. These fluctuations are primarily caused by precipitation and its absence, rather than fluctuations in the outside water level. **The phreatic head levels in the dike body, both prior to and after the peak high water levels, are approximately two to three**

meters higher than the daily target water level. They are even higher than the peak outside water levels that occurred during the high water events, as well as the design water level. This means that the intrusion of outside water will have only a limited effect on the pore-water pressures within the dike body, and shows the importance of precipitation in estimating pore-water pressure levels for dike stability calculations.

Comparing the maximum measured head levels with the typical schematizations of the phreatic line under normative conditions (left panel in Figure 6.13) shows that the observed heads are generally higher than the schematized line. Raising the schematized phreatic line would likely lead to more dike sections being classified as insufficiently stable, increasing the need for reinforcements. However, the dikes have remained stable despite already high phreatic lines caused by heavy rainfall, raising the question of whether such reinforcements are truly necessary.

Estimating the probability of phreatic lines while accounting for both precipitation and outside water levels, including their coincidence, can result in higher absolute values, but also in typical yearly levels that are closer to normative conditions. In other words, the frequency curve of the heads becomes higher and flatter. Therefore, monitoring potentially provides valuable information on the dike's actual performance and resistance, which can be incorporated through reliability updating, as discussed in Chapter 5 of this dissertation.

6.4.3 Sea dikes along Oosterschelde

Dike trajectory 27-2 protects the island of Tholen, located along the Eastern Scheldt. In the mouth of the Eastern Scheldt, the Eastern Scheld barrier is located, which closes during high water levels at the North Sea, reducing high water levels in the Eastern Scheldt. As a result, the normative water levels with exceedance frequencies of 1/10,000 per year lie about 75 cm higher than yearly occurring levels. For example, at Stavenisse, the water level with an exceedance frequency of 1/10,000 per year is NAP+3.57, while a water level of NAP+2.83m occurs, on average, every year. Wave run-up and overtopping can increase water infiltration during storms, but this is limited by the high crest levels, approximately between NAP+5m and NAP+8m, and the relatively short high-water durations, which often last only a few days.

The safety assessment of the sea dike along Tholen identified about 30km of dike not meeting the required safety for inner-slope instability. Failures slip planes are mainly small and shallow, where the shear strength is influenced by phreatic pore-water pressures, making the way the phreatic line are schematized essential. Phreatic heads were monitored for 1.5 to 3 years along the trajectory. An example of a monitoring site is shown in Figure 6.14. At this site, the highest heads occurred at the beginning of

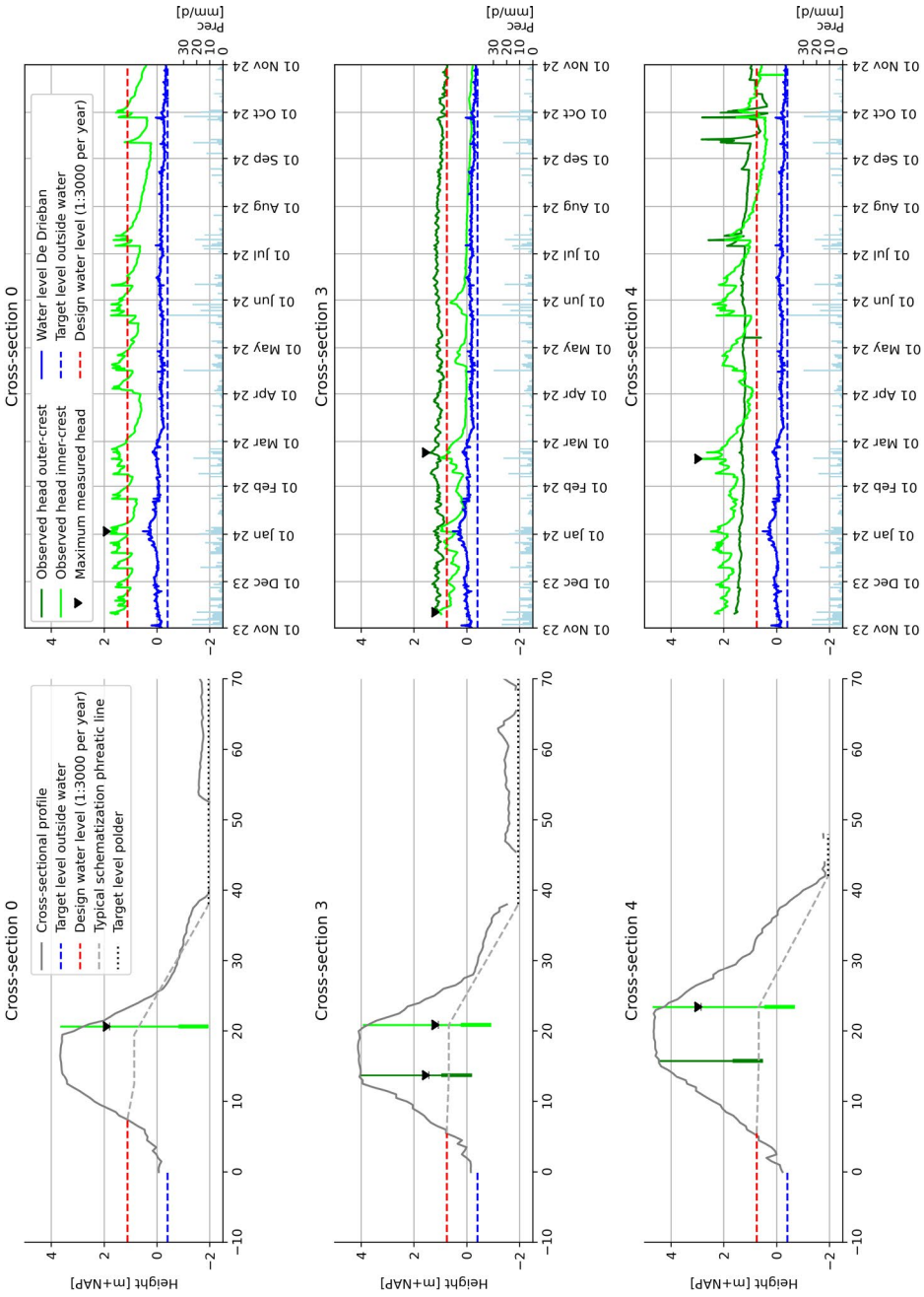


Figure 6.13 Cross-sectional profiles (left) and observed hydraulic heads (right) for three monitored dike sections. The left panels show the dike profiles with schematic phreatic lines according to dike safety assessments and target water levels. The right panels display the corresponding time series of observed inner- and outer-crest heads, as well as the measured water level at the pumping station and recorded precipitation at the Meteo station.

2024, during the exceptionally wet winter of 2023-2024, and the heads in the crest and inner-berm fluctuated throughout the year by about 60 cm. These variations are driven by precipitation and evaporation, with minimal influence from the tide, which fluctuates between NAP-2m and NAP+2m but does not penetrate deeply into the dike.

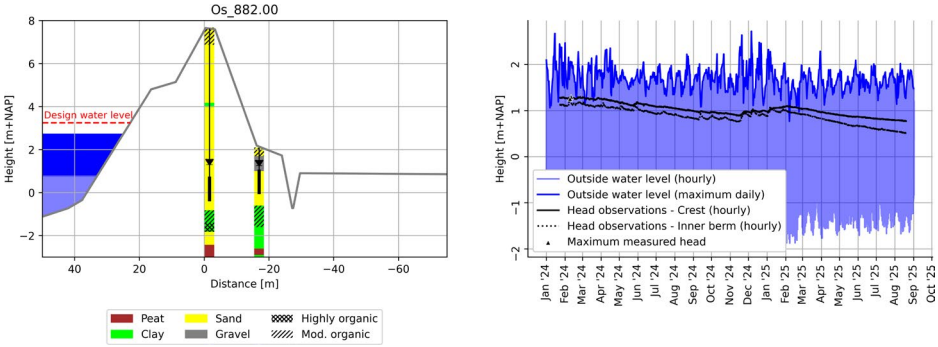


Figure 6.14 Cross-section of the sea dike along the Oosterschelde protecting the island of Tholen. The left panel shows the dike profile and soil layers. The right panel shows outside water levels (hourly and maximum daily) and groundwater heads measured at the crest and inner berm. The shaded blue area in the cross-section marks the range of observed outside water levels.

In total, 26 piezometers between the crest and toe of the dike, the zone most relevant for inner-slope instability, were used for analysis. For each piezometer, a time series model was developed to translate the effects of outside water levels and precipitation and evaporation into head fluctuations using impulse response functions. This approach separates the contributions of outside water levels and precipitation–evaporation to the observed head variations at each location, as illustrated in Figure 6.15. It can be seen that both outside water levels and precipitation-evaporation contribute to head fluctuations, but the latter generally has a stronger effect. At most piezometers, the tidal signal is absent, yet slower variations in the average outside water level still influence the heads. Overall, both drivers play a role in the phreatic heads occurring during high waters.

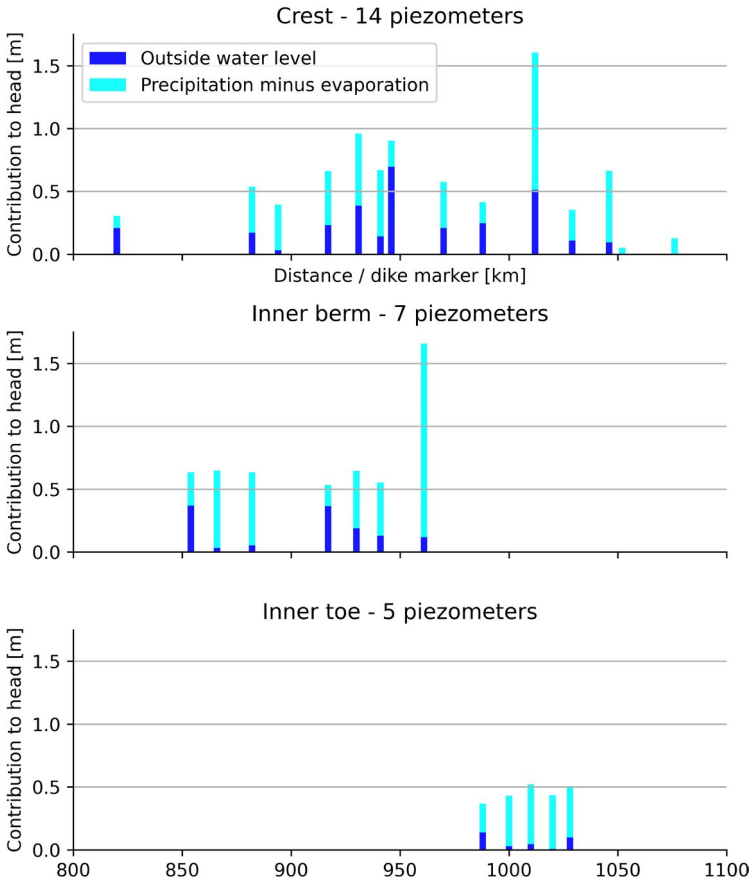


Figure 6.15 Contribution of outside water level and precipitation minus evaporation to groundwater head at different locations in the dike. The three panels show results for piezometers at the crest, inner berm, and inner toe, grouped by distance along the dike.

7

Conclusions and recommendations

7.1 Conclusions

The aim of this thesis is to enhance the geohydrological understanding of canal dikes and improve the reliability assessment of their stability and the dike system through observation-based analysis. Therefore, this study generated and investigated two unique datasets: 1) detailed subsurface water measurements at ten canal dikes, including soil moisture and hydraulic head data, and 2) 108 head observations from 48 canal dike monitoring sites across the Netherlands. These datasets were analysed using data-analysis techniques, time series modelling, and probabilistic modelling. The aim was divided into four research questions, and the key findings are presented in this chapter.

What is the geohydrological response of canal dikes with a focus on drought periods, and what are effective meteorological indicators that reflect this response?

Analysing a novel three-year monitoring data set of soil moisture and hydraulic heads in canal dikes between 2020 and 2023 showed that hydraulic heads in the winter can be very shallow and close to soil saturation: approximately half of the 37 monitoring points located on the inner slope and toe of the dike experienced head levels above 20 cm below the surface. The monitoring data, including two extremely dry summers, also showed that the heads can drop down to almost 2 meters during dry summers, including non-hydrostatic behaviours that can be detrimental to dike stability and often not accounted for in safety assessments for drought situations. These variations occurred while the canal water levels were almost constant, with fluctuations of centimetres to one or two decimeters, highlighting the importance of meteorological conditions.

The recovery time of soil moisture and groundwater levels, defined as the duration until the average winter situation occurs again, following the exceptionally dry summer of 2022, was up to 4.5 months. This occurred after the onset of the subsequent storm/winter season, and despite heavy precipitation events during the winter (in January 2023), the previously dried-out soils were able to store the water effectively, and high heads did not occur. This shows the importance of the interplay between seasons and extreme subsurface water conditions, with different corresponding failure mechanisms; dry summers can mitigate potential high hydraulic head levels relevant to winter failure mechanisms, and vice versa.

Furthermore, the performance of various meteorological drought indicators was assessed as proxies for the geohydrological response, and the precipitation deficit appeared to be the most effective indicator for both soil moisture and groundwater droughts. Standardised indices like *SPI* and *SPEI* are less suitable for safety assessments

because they filter out seasonality and express drought in relative terms, while absolute head levels matter most for dike stability. Direct soil moisture and head measurements, possibly combined with models, remain essential.

What is the variability in peak head responses and peak head statistics across Dutch canal dikes and the impact of climate change, and can these variations be explained by dike characteristics?

Using a unique dataset of head observations across the Dutch canal dike system, time series models were calibrated and used to characterize peak head responses. Four clusters of dikes were identified, each showing similar peak head responses within the cluster and distinct responses between clusters. The differentiating factor was the response times of the upper regime of these dikes, where longer response times caused peak heads to be driven by more prolonged rainfall events. The average response times of the clusters differ by roughly a factor of two. The identified dike clusters do not exhibit a clear spatial pattern and showed a limited relationship with dike characteristics. While the subsurface material and dike width appeared to influence the variation in head response across clusters, their presence across multiple clusters suggests they do not yield a definitive outcome. This suggests that head responses depend on a combination of dike characteristics, that the quantified dike characteristics are subjected to considerable uncertainty, or that additional factors or local anomalies, such as disturbed soil layers, control the response.

Furthermore, hydraulic head level statistics were quantified, with decimate heights ranging from around 5 to 50 cm, with a median value of 15 cm. These values indicate that extreme and yearly occurring load conditions are, on average, relatively close to each other. Considering the impact of climate change, driving higher winter precipitation and higher summer evaporation, it was found that head levels that have a return period of 100 years in the current climate are expected to occur about three times less frequently to seven times more frequently by 2100, depending on the climate scenario and the type of canal dike. Dryer summers can reduce the frequency of extreme peak heads by lowering head levels during summer, thereby increasing dike water storage capacity when rainfall returns, but this is counterbalanced by a wetter winter trend. The varying impacts of climate change on dikes can be largely attributed to the lower regime's response times: dikes with longer response times in the lower regime appear less affected by climate change, as they experience greater drying during drier summers.

How much more frequently do extreme hydraulic head peaks occur at larger spatial scales than at the section level, for Dutch canal dikes?

The frequency of extreme head peaks occurring can increase when moving from a section to larger spatial scales, and the ratio between these scales is known as the

length-effect factor. At the polder scale, for canal dike lengths of a few tens of kilometres and protecting areas ranging from tens to hundreds of square kilometres, natural spatial variability in the weather roughly doubles the occurrence of extreme heads compared to a single section. In contrast, variation in head responses can add a factor of four to five. At the scale of the entire canal dike system, the length-effect factor is dominated by the spatial variability in the weather. This spatial variability differs among head responses: dikes with shorter response or drainage times are more sensitive to short-duration rainfall, which has smaller spatial extents and leads to more extreme peak heads to occur within a region. On average, hydraulic head levels with a 10-year exceedance frequency at the section level occurred about 20 times more often at the scale of the entire dike system, for the homogeneous head response configuration. In contrast to polder scales, the length-effect factor at the scale of the entire dike system increases with smaller exceedance frequencies of peak heads, consistent with findings from other studies showing that the spatial extent of extreme precipitation decreases with increasing intensity: the length effect factor reaches 40 for head levels with an exceedance frequency of 1/100 per year under homogeneous head response. Assuming that the head response is heterogeneous, again, the length-effect factor increases by about a factor of four to five, compared to the homogeneous configuration.

What is the role of load variations in assessing credible failure probabilities?

The role of load variations in reliability updating and assessing credible failure probabilities was assessed using conceptual analyses and case studies, in which both a range of load variation and fragility curve combinations were examined. First, the conceptual analysis using hypothetical dikes showed that the impact of reliability updating increases when load variations are smaller, regardless of the prior failure probabilities. The reason is that only a relatively small load range contributes to the calculated failure probability, which, in absolute terms, is close to the load levels that occur every year. By updating the fragility curve, the conditional failure probabilities closest to the survived load level are reduced the most, and these levels have a larger contribution to the failure probabilities when load variations are small. Therefore, reliability updating has a larger impact on the failure probability for situations with relatively small load variations. Second, these findings were confirmed through case studies of two canal dikes and two river dikes in the Netherlands, considering the dike's probability of instability, which varied with fragility curves and load variations. The two canal dikes with small load variations, with decimate heights of the phreatic line below 10cm, benefited most from reliability updating, while the updated failure probabilities of the river dikes changed only slightly. This is because the load variations of the two river dikes are larger with decimate heights of the phreatic line of approximately 30 and 60 cm, while the shapes of the fragility curve were similar. In the moderate-strength

scenario for the canal dikes, reliability updating with a load occurring once every two years reduces failure probabilities by a factor of three to nine relative to the initial estimates. However, for the considered river dikes, even extreme load levels with a 1/100 probability per year, as occurred during the 2021 floods along the Meuse River, contribute little when it comes to improving failure probabilities through reliability updating.

Based on the conceptual analysis, a credibility check for safety assessments was proposed to identify situations in which the failure probability is dominated by load levels occurring every 10 years or more frequently. The contribution of frequent load levels to the total failure probability exceeds 50% when the inverse gradient of the fragility curve is about 1.5 times greater than the decimate height of the load. This would imply that small, relatively common occurring load levels contribute more to failure than rare, extreme events for which dikes are designed. This outcome can contradict the observed stability of dikes over many decades, and the load variations and strength uncertainties are out of balance, resulting in incredible failure probability estimates. Situations of small load variation occur widely in the Netherlands, for example, along much of the over 10,000 kilometres of canal dikes and in primary flood defences in lake and delta regions where the decimate height of the water level is relatively small. In these situations, reducing the uncertainties related to the dike's strength will lower the inverse gradient of the fragility curve, resulting in more credible failure probability estimates. This can be achieved, for example, by monitoring load levels in combination with reliability updating methods, or by gathering more detailed data on soil strength.

7.2 Recommendations for research

The following topics need more work to improve failure probability estimates of dike stability by a deeper understanding of the geohydrology:

- For dike stability, estimates of exceedance probabilities of pore-water pressures form the important basis of a probability estimate of dike instability. This dissertation aimed to improve the geohydrological understanding of canal dikes using measurements and modelling techniques, including quantifying the probability that a particular phreatic line can occur. However, in the Dutch dike safety practice, available piezometer measurements are limited, and guidelines are used to estimate the probability of the phreatic line, which is often assumed to be conditional on the water level. Although the phreatic line is uncertain, these uncertainties are often not considered, and typically, a deterministic estimate is used, of which it is unknown whether this is a conservative estimate. **It is recommended that existing guidelines be verified using both measurements and models, and that uncertainties in pore-water pressure fields be quantified more statistically**.

The guidelines call the estimates "conservative," yet they rely on limited underlying data. Nowadays, more measurements are available, and these estimates can be refined and supported by more statistically robust analyses.

- In this thesis, the impact of climate change effects on extreme hydraulic head levels has been quantified by considering changes in precipitation and potential evaporation, assuming that the head response itself remains unchanged. Nevertheless, **second-order effects, such as changes in soil properties, were not considered, and their impact on Dutch canal dikes is recommended to be investigated.** Changes in the hydraulic conductivity and water retention capacity of soils can be initiated by cyclic wetting and drying. These processes lead to swelling and shrinkage, consolidation, and gradual alterations in the dike structure over time (Stirling et al., 2021; Azizi et al., 2020). As climate change is expected to intensify dry-wet cycles, these processes may become more pronounced, potentially affecting the stability of dikes. The monitoring network described in this dissertation (Chapter 2), which continues to operate in December 2025, also showed a **summer–winter dependency pattern: dry summers lead to lower winter head levels, while wet summers lead to higher ones**. This could not be explained solely by the fact that dried-out dikes can store more water and result in lower heads in the winter, but a second-order effect appears to play a role as well, in which the head responses seem to change.
  - Quantifying all the effects of future climate scenarios is challenging, as both the hydraulic and mechanical behaviours of soils are intertwined and impacted, and are site-specific. **To better understand these processes and quantify changes in head responses within dikes, continuous long-term monitoring, preferably exceeding 10 years, is essential.** Such datasets are needed to separate climate-driven trends from short-term variability and local effects.
- While this dissertation focused on Dutch canal dikes, where heavy precipitation events primarily drive phreatic pore-water pressures, for other types of dikes, such as the Dutch primary flood defences, the role of precipitation in extreme pore-water pressures can also be significant, as shown in Chapter 6. Next to the maximum outside water level, the effects of precipitation coinciding with the flood wave, as well as the duration of the flood wave, are aspects that are limitedly considered in dike safety, which can be both important aspects in estimating extreme head levels at section levels (van Woerkem et al., 2023). **Therefore, it is recommended to develop a method to estimate pore-water pressure probabilities for different primary flood defences, using groundwater models and statistical analyses to evaluate the influence and coincidence of the main driving processes.**
- **Estimate the system-level consequences of knowledge uncertainties by quantifying how many kilometres of dikes would need reinforcement to ensure that inherently unsafe sections meet the safety standard.** Quantifying epistemic

uncertainties is often based on analysing variability observed in laboratory tests or field data. For example, soil shear strength is typically derived from laboratory datasets containing many samples from a larger region, which are statistically analysed and then used to estimate failure probabilities at the section level. As a result, the most likely shear strength at failure is represented by an average or weak realisation, while strong realisations will not. This results in average or weak realisations across sections of a dike system, despite both weak and strong soils being known to be present. However, there is insufficient information to estimate the strength at the section level. This results in a reinforcement task that is larger than is inherently needed due to knowledge uncertainties. The magnitude of these resulting "false positives" relative to "true positives" is unknown. It is therefore recommended to investigate this by analysing multiple local realisations and quantifying the system-level consequences of this modelling approach.

7.3 Recommendations for practice

Several findings from this dissertation can be directly applied in the management, assessment, and design of dikes:

- **Use soil moisture and hydraulic head monitoring in combination with the continuous precipitation deficit and time series models to signal extreme subsurface water conditions of canal dikes and potentially hazardous situations.** Emergency response, such as inspections and emergency measures, can help prevent dike failures and flooding, but only if we know when hazardous situations are about to occur. These situations can occur both in summer and winter. For dike stability, it is recommended to use monitoring systems that measure soil moisture and hydraulic head levels, such as those set up in this study, which can be used directly to identify extreme loading conditions. However, there are also challenges regarding proper monitoring, and it is unfeasible to monitor everywhere; therefore, models or proxy variables are useful. This dissertation showed that the continuous precipitation deficit is the most reliable meteorological drought indicator for subsurface water conditions in canal dikes. However, it has limitations, as geohydrology involves more than just meteorology, though it can still provide a proper estimate of subsurface conditions during droughts. Whether this results in hazardous situations depends, among others, on the dike's strength. Furthermore, it is recommended to use the time-series model set developed in this study to estimate heads from precipitation and evaporation at places where heads are not monitored, and to focus on their relative extremity rather than absolute levels, in line with the groundwater drought indicator as shown in Chapter 6. While this is developed for droughts, this approach can also be used for wet conditions.

- **Include non-hydrostatic head profiles in safety assessments to assess dike stability during drought situations, where soils are losing weight, but pore-water pressures react delayed**. This research highlighted the presence of non-hydrostatic pore water pressures in dikes during droughts, which can negatively influence dike safety; the top soil layers dry out, reducing soil weights, while the reduction of deeper pore-water pressures is delayed. This can result in situations where effective stresses are lower than those assumed under hydrostatic conditions. To accurately evaluate dike stability during droughts, safety assessments must take into account these non-hydrostatic conditions. Adapt safety assessment methods to account for delayed pressure dissipation and upward flows during dry periods, which can reduce effective stresses.
- **Place time-varying pore-water pressures at the centre of stability analyses by using fragility curves conditional on pore-water-pressure scenarios, even when their uncertainties appear to be irrelevant in failure probability analyses**. The probability of dike instability can be insensitive to variations in pore pressures. This can be due to its relative unimportance compared to other uncertainties, such as the soil shear strength, which can be highly uncertain, or to the fact that stability is physically only moderately influenced by changes in pore-water pressures. This latter case can arise when the shear strength of impermeable soils, such as clay and peat, is modelled using the SHANSEP material model. In that case, the question arises: if pore-water pressure changes play a minor role in dike stability, which other time-varying parameter will trigger a dike to lose its stability? Considering the failure path of a dike for the failure mechanism of inner-slope instability, increasing pore-water pressures are key. Therefore, it is recommended to put this central in safety assessments by using the fragility curve conditional on pore-water pressure scenarios. By placing it more central, monitoring and modelling can be stimulated to improve the schematisations of pore-water pressures, rather than relying on a conservative estimate, substantiated by little data and limited statistical analyses.
- **Introduce a credibility check in dike safety assessments to verify the balance between load variability and strength uncertainty and avoid dikes being assessed as unsafe because of a lack of knowledge**. Almost all dikes have remained stable for decades without visible damage or degradation, and often even gain strength over time due to soil ageing. For such dikes, it is considered incredible if loads with return periods shorter than about 10 years contribute more than 50% to the total failure probability. This situation typically arises when the inverse gradient of the fragility curve exceeds 1.5 times the decimate height of the load. For canal dikes, as well as other flood defences along, for example, Markermeer and the Dutch delta areas with storm surge barriers, the decimate height of the load can be relatively small. In these areas, this check can be especially beneficial for identifying incredible

failure probabilities and it is recommended to implement this check in dike safety assessments.

- **Make a distinction between different types of uncertainties in dike safety assessments, like spatial vs temporal and epistemic vs inherent uncertainties, to better understand the calculated failure probability.** In dike systems with high failure probabilities and epistemic uncertainty dominates, it may be assumed that only part of the dikes are actually unsafe and need to be reinforced. By breaking down the failure probability by accounting for epistemic uncertainties, $P(Z < 0)$, and based on the expected value of epistemic stochastic variables, $P(Z < 0 \mid I = \mu(i))$, the failure probability can be interpreted in a better way: are dikes truly inherently unsafe, or is it challenging to demonstrate adequate safety with the available data and models? These kinds of analyses are not an argument for using the expected value of epistemic stochastic variables, since these uncertainties are present. However, it can be helpful to understand the contributions of inherent and epistemic uncertainties, as also noted by Vrouwenvelder & Vrijling (2000).
- **Treat monitoring and field data collection as a long-term commitment to understand the physics of dike behaviour and its uncertainties**. Models can support the understanding of dike behaviour, but they should never dictate the dike safety assessments, because their conclusions are only as reliable as the observations on which they are based. Therefore, these models need to be aligned with observations, especially for stability analysis where available information is often scares and the challenge lies in dealing with uncertainties. For example, observations show large spatial variability, and models are unable to capture the failure mechanisms observed in the field, as in the recent (near) failures of canal dikes at Reeuwijk (2021) and Ommoord (2025). Although more and more advanced modelling techniques are becoming available, without field and monitoring data, we risk building ever more advanced models on foundations we do not truly understand. Additionally, applying these models in a relatively data-scarce environment, like a few point measurements every hundred meters of dike, can be risky. Therefore, monitoring and field data are essential for failure-probability estimates, and we need to focus on integrating these data with physical and probabilistic models.

References

Aanstoos, J. V., Hasan, K., O'Hara, C. G., Prasad, S., Dabbiru, L., Mahrooghy, M., Doelling, I., Casey, M., Henkel, T., & Metcalf, J. (2011). Earthen levee monitoring with synthetic aperture radar. In *2011 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)* (pp. 1–6). IEEE. https://doi.org/10.1109/AIPR.2011.6176384

Actueel Hoogtebestand Nederland. (2022). http://www.ahn.nl

AghaKouchak, A., Huning, L. S., Sadegh, M., Qin, Y., Markonis, Y., Vahedifard, F., Moftakhari, H., & Mazdiyasni, O. (2023). Toward impact-based monitoring of drought and its cascading hazards. *Nature Reviews Earth and Environment, 4*(8), 582–595. https://doi.org/10.1038/s43017-023-00457-2

Arsenault, R., Brissette, F., & Martel, J.-L. (2018). The hazards of split-sample validation in hydrological model calibration. *Journal of Hydrology, 566*, 346–362. https://doi.org/10.1016/j.jhydrol.2018.09.027

Azizi, A., Musso, G., & Jommi, C. (2020). Effects of repeated hydraulic loads on microstructure and hydraulic behaviour of a compacted clayey silt. *Canadian Geotechnical Journal, 57*(1), 100–114. https://doi.org/10.1139/cgj-2018-0505

Bachmann, D., Huber, N. P., Johann, G., & Schüttrumpf, H. (2013). Fragility curves in operational dike reliability assessment. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 7*(1), 49–60. https://doi.org/10.1080/17499518.2013.767664

Bakker, M., & Schaars, F. (2019). Solving groundwater flow problems with time series analysis: You may not even need another model. *Groundwater, 57*(6), 826–833. https://doi.org/10.1111/gwat.12927

Bayes, T. (1763). LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS Communicated by Mr. Price, in a letter to John Canton, AMFR S. *Philosophical Transactions of the Royal Society of London, 53*, 370–418. https://doi.org/10.1098/rstl.1763.0053

Bedford, T., & Cooke, R. (2001). *Probabilistic risk analysis: Foundations and methods*. Cambridge University Press.

Beersma, J. J., & Buishand, T. A. (2004). Joint probability of precipitation and discharge deficits in The Netherlands. *Water Resources Research*, *40*(12), W12508. https://doi.org/10.1029/2004WR003265

Beersma, J. J., & Buishand, T. A. (2007). Drought in the Netherlands–Regional frequency analysis versus time series simulation. *Journal of Hydrology, 347*(3–4), 332–346. https://doi.org/10.1016/j.jhydrol.2007.09.042

Bengfort, B., & Bilbro, R. (2019). Yellowbrick: Visualizing the Scikit-Learn model selection process. *Journal of Open Source Software, 4*(35), 1075. https://doi.org/10.21105/joss.01075

Berendrecht, W. L., Heemink, A. W., Van Geer, F. C., & Gehrels, J. C. (2006). A non-linear state space approach to model groundwater fluctuations. *Advances in Water Resources, 29*(6), 959–973. https://doi.org/10.1016/j.advwatres.2005.08.009

Berendsen, H. J. A., & Stouthamer, E. (2002). Paleogeographic evolution and avulsion history of the Holocene Rhine-Meuse delta, The Netherlands. *Netherlands Journal of Geosciences, 81*(1), 97–112. https://doi.org/10.1017/S0016774600020606

Bezuijen, A., Kruse, G. A., & Van, M. A. (2005). Failure of peat dikes in The Netherlands. In *Proceedings of the 16th International Conference on Soil Mechanics and Geotechnical Engineering* (pp. 1857–1860). IOS Press.

Bishop, A. W. (1955). The use of the slip circle in the stability analysis of slopes. *Géotechnique, 5*(1), 7–17. https://doi.org/10.1680/geot.1955.5.1.7

Bouw, R. (2008). *Decimate heights TMR2006* (Report No. 373295). Witteveen+Bos. https://open.rijkswaterstaat.nl/open-overheid/onderzoeksrapporten/@26517/decimeringshoogten-tmr2006/

Bouwer, L. M. (2018). Observed and projected impacts from extreme weather events: Implications for loss and damage. In R. Mechler, L. M. Bouwer, T. Schinko, S. Surminski, & J. Linnerooth-Bayer (Eds.), *Loss and damage from climate change: Concepts, methods and policy options* (pp. 63–82). Springer International Publishing. https://doi.org/10.1007/978-3-319-72026-5_3

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day.

Briggs, K. M., Helm, P. R., Smethurst, J. A., Smith, A., Stirling, R., Svalova, A., Fraccica, A., Glendinning, S., Dijkstra, T., Dixon, N., Crosby, C., Cunningham, M., El-Hamalawi, A., Fullen, M., Wackrow, R., Alshawabkeh, A., Ball, J., Freeborough, K., Gunn, D., …Wilkinson, P. (2023). Evidence for the weather-driven deterioration of ageing transportation earthworks in the UK. *Transportation Geotechnics, 43*, 101130. https://doi.org/10.1016/j.trgeo.2023.101130

Calle, E. (1996). *Characteristic values of geotechnical parameters* [Lecture notes]. Foundation Post Academic Education.

Calle, E. O. F., Kanning, W., & Schweckendiek, T. (2021). *Characteristic values of soil properties in Dutch codes of practice – Theoretical backgrounds and assumptions* (Report No. 11206883-014-GEO-0001). Deltares.

Casciati, F., & Faravelli, L. (1991). *Fragility analysis of complex structural systems*. Research Studies Press.

CEN. (2004). *Eurocode 7: Geotechnical design. Part 1: General rules, EN1997-1*. European Committee for Standardization.

Cheng, L., & AghaKouchak, A. (2014). Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific Reports, 4*(1), 7093. https://doi.org/10.1038/srep07093

Chlaib, H. K., Mahdi, H., Al-Shukri, H., Su, M. M., Catakli, A., & Abd, N. (2014). Using ground penetrating radar in levee assessment to detect small scale animal burrows. *Journal of Applied Geophysics, 103*, 121–131. https://doi.org/10.1016/j.jappgeo.2014.01.011

Clark, C. (2002). Measured and estimated evaporation and soil moisture deficit for growers and the water industry. *Meteorological Applications, 9*(1), 85–93. https://doi.org/10.1017/S1350482702001093

Collenteur, R. A., Bakker, M., Caljé, R., Klop, S. A., & Schaars, F. (2019). Pastas: Open source software for the analysis of groundwater time series. *Groundwater, 57*(6), 877–885. https://doi.org/10.1111/gwat.12925

Collenteur, R. A., Bakker, M., Klammler, G., & Birk, S. (2021). Estimation of groundwater recharge from groundwater levels using nonlinear transfer function noise models and comparison to lysimeter data. *Hydrology and Earth System Sciences, 25*(5), 2931–2949. https://doi.org/10.5194/hess-25-2931-2021

Coronese, M., Lamperti, F., Keller, K., Chiaromonte, F., & Roventini, A. (2019). Evidence for sharp increase in the economic damages of extreme natural disasters. *Proceedings of the National Academy of Sciences, 116*(43), 21450–21455. https://doi.org/10.1073/pnas.1907826116

Cundill, S. (2016). *Investigation of remote sensing for dike inspection* [Doctoral dissertation, University of Twente]. https://doi.org/10.3990/1.9789036540360

De Bruin, H. A. R. (1987). From Penman to Makkink. In *Evaporation and weather* (Technical Meeting 44, pp. 5–31). TNO Committee on Hydrological Research.

De Bruin, H. A. R., & Stricker, J. N. M. (2000). Evaporation of grass under non-restricted soil moisture conditions. *Hydrological Sciences Journal*, *45*(3), 391–406. https://doi.org/10.1080/02626660009492337

De Koker, N., Viljoen, C., Lenner, R., & Jacobsz, S. W. (2020). Updating structural reliability efficiently using load measurement. *Structural Safety*, *84*, 101939. https://doi.org/10.1016/j.strusafe.2020.101939

Dekker, L. W., & Ritsema, C. J. (2000). Wetting patterns and moisture variability in water repellent Dutch soils. *Journal of Hydrology*, *231–232*, 148–164. https://doi.org/10.1016/S0022-1694(00)00191-8

De Lange, W. J., Prinsen, G. F., Hoogewoud, J. C., Veldhuizen, A. A., Verkaik, J., Essink, G. H. O., van Walsum, P. E. V., Delsman, J. R., Hunink, J. C., Massop, H. T. L., & Kroon, T. (2014). An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument. *Environmental Modelling & Software*, *59*, 98–108. https://doi.org/10.1016/j.envsoft.2014.05.009

Den Haan, E. J., & Kruse, G. A. M. (2007). Characterisation and engineering properties of Dutch peats. In T. S. Tan, K. K. Phoon, D. W. Hight, & S. Leroueil (Eds.), *Characterisation and engineering properties of natural soils* (Vol. 3, pp. 2101–2133). Taylor & Francis.

Der Kiureghian, A., & Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, *31*(2), 105–112. https://doi.org/10.1016/j.strusafe.2008.06.020

Deyer, M., Utili, S., & Zielinski, M. (2009). Field survey of desiccation fissuring of flood embankments. *Proceedings of the Institution of Civil Engineers - Water Management*, *162*(3), 221–232. https://doi.org/10.1680/wama.2009.162.3.221

Elia, G., Cotecchia, F., Pedone, G., Vaunat, J., Vardon, P. J., Pereira, C., Springman, S. M., Rouainia, M., van Esch, J., Koda, E., Josifovski, J., Nocilla, A., Askariazad, A., Stirling, R. A., Helm, P., Loveridge, F., & Toll, D. G. (2017). Numerical modelling of slope–vegetation–atmosphere interaction: An overview. *Quarterly Journal of Engineering Geology and Hydrogeology*, *50*(3), 249–270. https://doi.org/10.1144/qjegh2016-079

European Commission, Joint Research Centre, van Den Eijnden, B., Knuuti, M., Lesny, K., Löfman, M., Mavritsakis, A., Roubos, A., Schweckendiek, T., Sciarretta, F., Ebener, A., Escher, K., Spross, J., Commend, S., Hehenkamp, M., Arnold, P., Wilhelm, S., Ene, A., & Rimoldi, L. (2024). *Reliability background of the Eurocodes: Support to the implementation, harmonization and further development of the Eurocodes*. Publications Office of the European Union. https://doi.org/10.2760/134254

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Wiley. https://doi.org/10.1002/9780470977811

Frank, R., Bauduin, C., Driscoll, R. M., Kavvadas, M., Ovesen, N. K., Orr, T., & Schuppener, B. (2004). *Designers' guide to Eurocode 7: Geotechnical design: Designers' guide to EN 1997-1, Eurocode 7: Geotechnical design—General rules* (Vol. 17). Thomas Telford.

Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. *Earth-Science Reviews*, *162*, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Geerse, C. P. M. (2011). *Hydra-zoet for the fresh water systems in the Netherlands—Probabilistic model for the assessment of dike heights* (Report No. PR2168). HKV Consultants.

Gildeh, H. K., Hosseini, P., Zhang, H., Riaz, M., & Acharya, M. (2019). Canal embankment failure mechanism, breach parameters and outflow predictions. In *Sustainable and safe dams around the world/Un monde de barrages durables et sécuritaires* (pp. 61–74). CRC Press. https://doi.org/10.1201/9780429319778-6

Green, C. H., Parker, D. J., & Tunstall, S. M. (2000). *Assessment of flood control and management options* (Technical Report). Middlesex University Flood Hazard Research Centre.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology, 31*(4), 337–350. https://doi.org/10.1007/s10654-016-0149-3

Guha-Sapir, D., Below, R., & Hoyois, P. (2016). *EM-DAT: The CRED/OFDA international disaster database*. Université Catholique de Louvain.

Hall, J., & Blöschl, G. (2018). Spatial patterns and characteristics of flood seasonality in Europe. *Hydrology and Earth System Sciences, 22*(7), 3883–3901. https://doi.org/10.5194/hess-22-3883-2018

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 28*(1), 100–108. https://doi.org/10.2307/2346830

Hemel, M. J., Peters, D. J., Schweckendiek, T., & Jonkman, S. N. (2024). Reliability updating for lateral failure of historic quay walls. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 18*(4), 882–903. https://doi.org/10.1080/17499518.2024.2302141

Hillel, D. (2003). *Introduction to environmental soil physics*. Elsevier.

Hisdal, H., Stahl, K., Tallaksen, L. M., & Demuth, S. (2001). Have streamflow droughts in Europe become more severe or frequent? *International Journal of Climatology, 21*(3), 317–333. https://doi.org/10.1002/joc.619

Hoogheemraadschap Hollands Noorderkwartier. (2025). *Administrative progress report on regional flood defences 2024* (Version 1.4; Registration number 25.0347462). [In Dutch]

Hooghoudt, S. B. (1940). *Bijdragen tot de kennis van eenige natuurkundige grootheden van den grond: Algemeene beschouwing van het probleem van de detailontwatering en de infiltratie door middel van parallel loopende drains, greppels, slooten en kanalen* (Vol. 46, 14B). Algemeene Landsdrukkerij.

Huang, W., Loveridge, F. A., Briggs, K. M., Smethurst, J. A., Saffari, N., & Thomson, F. (2024). Forecast climate change impact on porewater pressure regimes for the design and assessment of clay earthworks. *Quarterly Journal of Engineering Geology and Hydrogeology, 57*(1), qjegh2023-015. https://doi.org/10.1144/qjegh2023-015

Huang, W.-C., Liu, W.-C., & Liu, H.-M. (2025). Uncertainty analysis of overflow due to sea dike failure during typhoon events. *Journal of Marine Science and Engineering, 13*(3), 573. https://doi.org/10.3390/jmse13030573

Hubble, T., De Carli, E., & Airey, D. (2014). Geomechanical modeling of the Murray's Millennium Drought river bank failures: A case of the unexpected consequences of slow drawdown, soft bank materials and anthropogenic change. In *Proceedings of the 7th Australian Stream Management Conference* (pp. 278–284).

IPCC. (2021). *Climate change 2021: The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou, Eds.). Cambridge University Press. https://doi.org/10.1017/9781009157896

IPCC. (2022). *Climate change 2022: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama, Eds.). Cambridge University Press. https://doi.org/10.1017/9781009325844

Jamalinia, E., Vardon, P. J., & Steele-Dunne, S. C. (2019). The effect of soil–vegetation–atmosphere interaction on slope stability: A numerical study. *Environmental Geotechnics*, *8*(7), 430–441. https://doi.org/10.1680/jenge.18.00201

Jamalinia, E., Vardon, P. J., & Steele-Dunne, S. C. (2020). The impact of evaporation induced cracks and precipitation on temporal slope stability. *Computers and Geotechnics*, *122*, 103506. https://doi.org/10.1016/j.compgeo.2020.103506

Jasim, F. H., Vahedifard, F., Ragno, E., AghaKouchak, A., & Ellithy, G. (2017). Effects of climate change on fragility curves of earthen levees subjected to extreme precipitations. In *Geo-Risk 2017: Reliability-Based Design and Code Developments* (pp. 498–507). American Society of Civil Engineers. https://doi.org/10.1061/9780784480717.047

Jevrejeva, S., Jackson, L. P., Grinsted, A., Lincke, D., & Marzeion, B. (2018). Flood damage costs under the sea level rise with warming of 1.5°C and 2°C. *Environmental Research Letters*, *13*(7), 074014. https://doi.org/10.1088/1748-9326/aacc76

Jongejan, R. B., Diermanse, F., Kanning, W., & Bottema, M. (2020). Reliability-based partial factors for flood defenses. *Reliability Engineering & System Safety*, *193*, 106589. https://doi.org/10.1016/j.ress.2019.106589

Jongejan, R. B., & Maaskant, B. (2015). Quantifying flood risks in the Netherlands. *Risk Analysis*, *35*(2), 252–264. https://doi.org/10.1111/risa.12285

Jonkman, S. N., Curran, A., & Bouwer, L. M. (2024). Floods have become less deadly: An analysis of global flood fatalities 1975–2022. *Natural Hazards*, *120*(7), 6327–6342. https://doi.org/10.1007/s11069-024-06504-4

Jonkman, S. N., Jongejan, R., & Maaskant, B. (2011). The use of individual and societal risk criteria within the Dutch flood safety policy—Nationwide estimates of societal risk and policy applications. *Risk Analysis*, *31*(2), 282–300. https://doi.org/10.1111/j.1539-6924.2010.01502.x

Jonkman, S. N., Kok, M., & Vrijling, J. K. (2008). Flood risk assessment in the Netherlands: A case study for dike ring South Holland. *Risk Analysis*, *28*(5), 1357–1374. https://doi.org/10.1111/j.1539-6924.2008.01103.x

Kanning, W. (2012). *The weakest link—Length effects for piping* [Doctoral dissertation, Delft University of Technology]. https://doi.org/10.4233/uuid:5fb7b121-dc00-48aa-bda2-b163f10513bf

Kchouk, S., Melsen, L. A., Walker, D. W., & van Oel, P. R. (2021). A review of drought indices: Predominance of drivers over impacts and the importance of local context. *Natural Hazards and Earth System Sciences Discussions*, 1–28.

Kind, J. M. (2014). Economically efficient flood protection standards for the Netherlands. *Journal of Flood Risk Management*, *7*(2), 103–117. https://doi.org/10.1111/jfr3.12026

Klerk, W. J., Kanning, W., Kok, M., Bronsveld, J., & Wolfert, A. R. M. (2023). Accuracy of visual inspection of flood defences. *Structure and Infrastructure Engineering*, *19*(8), 1076–1090. https://doi.org/10.1080/15732479.2021.2001543

Klijn, F., Asselman, N., & Mosselman, E. (2019). Robust river systems: On assessing the sensitivity of embanked rivers to discharge uncertainties, exemplified for the Netherlands' main rivers. *Journal of Flood Risk Management*, *12*(S2), e12511. https://doi.org/10.1111/jfr3.12511

Knotters, M., & de Gooijer, J. G. (1999). TARSO modeling of water table depths. *Water Resources Research*, *35*(3), 695–705. https://doi.org/10.1029/1998WR900049

Knotters, M., & van Walsum, P. E. V. (1997). Estimating fluctuation quantities from time series of water-table depths using models with a stochastic component. *Journal of Hydrology*, *197*(1–4), 25–46. https://doi.org/10.1016/S0022-1694(96)03278-7

Kok, M., Jongejan, R., Nieuwjaar, M. W. C., & Tánczos, I. C. (2017). *Fundamentals of flood protection*. ENW Expertisenetwerk Waterveiligheid.

Kong, L. W., Bai, W., & Guo, A. G. (2012). Effects of cracks on the electrical conductivity of a fissured laterite: A combined experimental and statistical study. *Geotechnical Testing Journal*, *35*(6), 870–878. https://doi.org/10.1520/GTJ20120070

Ladd, C. C., & Foott, R. (1974). New design procedure for stability of soft clays. *Journal of the Geotechnical Engineering Division*, *100*(7), 763–786. https://doi.org/10.1061/AJGEB6.0000066

Lendering, K., Schweckendiek, T., & Kok, M. (2018). Quantifying the failure probability of a canal levee. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, *12*(3), 203–217. https://doi.org/10.1080/17499518.2018.1426865

Liu, L., Gudmundsson, L., Hauser, M., Qin, D., Li, S., & Seneviratne, S. I. (2019). Revisiting assessments of ecosystem drought recovery. *Environmental Research Letters*, *14*(11), 114028. https://doi.org/10.1088/1748-9326/ab4c61

Lloyd-Hughes, B. (2014). The impracticality of a universal drought definition. *Theoretical and Applied Climatology*, *117*(3–4), 607–611. https://doi.org/10.1007/s00704-013-1025-7

Manh, N. V., Merz, B., & Apel, H. (2013). Sedimentation monitoring including uncertainty analysis in complex floodplains: A case study in the Mekong Delta. *Hydrology and Earth System Sciences*, *17*(8), 3039–3057. https://doi.org/10.5194/hess-17-3039-2013

Martín-Antón, M., Negro, V., del Campo, J. M., López-Gutiérrez, J. S., & Esteban, M. D. (2016). Review of coastal land reclamation situation in the world. *Journal of Coastal Research*, *75*(sp1), 667–671. https://doi.org/10.2112/SI75-133.1

McDonald, M. G., & Harbaugh, A. W. (2003). The history of MODFLOW. *Groundwater*, *41*(2), 280–283. https://doi.org/10.1111/j.1745-6584.2003.tb02591.x

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology* (Vol. 17, pp. 179–183). American Meteorological Society.

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C., Bates, P., Bertola, M., Kemter, M., Kreibich, H., Lall, U., & Macdonald, E. (2021). Causes, impacts and patterns of disastrous river floods. *Nature Reviews Earth & Environment*, *2*(9), 592–609. https://doi.org/10.1038/s43017-021-00195-3

Meter Group. (2021). *TEROS 11/12 manual web*. https://publications.metergroup.com/Manuals/20587_TEROS11-12_Manual_Web.pdf

Moore, R., Carey, J. M., & McInnes, R. G. (2010). Landslide behaviour and climate change: Predictable consequences for the Ventnor Undercliff, Isle of Wight. *Quarterly Journal of Engineering Geology and Hydrogeology, 43*(4), 447–460. https://doi.org/10.1144/1470-9236/08-086

Morton, L. W., & Olson, K. R. (2018). The pulses of the Mekong River Basin: Rivers and the livelihoods of farmers and fishers. *Journal of Environmental Protection, 9*(4), 431–459. https://doi.org/10.4236/jep.2018.94027

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J.-N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data, 13*(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Nationaal Georegister. (2024). Toets- en veiligheidsoordelen primaire waterkeringen (WBI). Metadata Date 0–2024:4–09.

Neuzil, C. E. (1986). Groundwater flow in low-permeability environments. *Water Resources Research, 22*(8), 1163–1195. https://doi.org/10.1029/WR022i008p01163

Nicholls, R. J., & Cazenave, A. (2010). Sea-level rise and its impact on coastal zones. *Science, 328*(5985), 1517–1520. https://doi.org/10.1126/science.1185782

Nofal, O. M., van de Lindt, J. W., & Do, T. Q. (2020). Multi-variate and single-variable flood fragility and loss approaches for buildings. *Reliability Engineering & System Safety, 202*, 106971. https://doi.org/10.1016/j.ress.2020.106971

O'Dell, J., Nienhuis, J. H., Cox, J. R., Edmonds, D. A., & Scussolini, P. (2021). A global open-source database of flood-protection levees on river deltas (openDELvE). *Natural Hazards and Earth System Sciences Discussions*, 1–16. https://doi.org/10.5194/nhess-2021-26

Oude Essink, G. H. P., Van Baaren, E. S., & De Louw, P. G. (2010). Effects of climate change on coastal groundwater systems: A modeling study in The Netherlands. *Water Resources Research, 46*(10), W00F04. https://doi.org/10.1029/2009WR008719

Özer, I. E., Rikkert, S. J., van Leijen, F. J., Jonkman, S. N., & Hanssen, R. F. (2019a). Sub-seasonal levee deformation observed using satellite radar interferometry to enhance flood protection. *Scientific Reports, 9*(1), 2646. https://doi.org/10.1038/s41598-019-39474-x

Özer, I. E., van Damme, M., & Jonkman, S. N. (2019b). Towards an international levee performance database (ILPD) and its use for macro-scale analysis of levee breaches and failures. *Water, 12*(1), 119. https://doi.org/10.3390/w12010119

PBL Netherlands Environmental Assessment Agency. (n.d.). *Correction wording flood risks for the Netherlands in IPCC report*. Retrieved May 21, 2025, from https://www.pbl.nl/en/correction-wording-flood-risks-for-the-netherlands-in-ipcc-report

Philip, S. Y., Kew, S. F., van der Wiel, K., Wanders, N., & van Oldenborgh, G. J. (2020). Regional differentiation in climate change induced drought trends in The Netherlands. *Environmental Research Letters, 15*(9), 094081. https://doi.org/10.1088/1748-9326/ab97ca

Phoon, K. K. (2016). Chapter 1 Reliability as a basis for geotechnical design. In K. K. Phoon & J. Ching (Eds.), *Reliability of geotechnical structures in ISO2394* (pp. 1–32). CRC Press. https://doi.org/10.1201/9781315364179-2

Phoon, K. K. (2023). What geotechnical engineers want to know about reliability. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *9*(2), 03123001. https://doi.org/10.1061/AJRUA6.RUENG-1002

Phoon, K. K., & Kulhawy, F. H. (1999). Characterization of geotechnical variability. *Canadian Geotechnical Journal*, *36*(4), 612–624. https://doi.org/10.1139/t99-038

Pigott, P. T., Hanrahan, E. T., & Somers, N. (1992). Major canal reconstruction in peat areas. *Water Maritime and Energy*, *96*(3), 141–152. https://doi.org/10.1680/iwtme.1992.21082

Pleijster, E. J., van der Veeken, C., Jongerius, R., & Luiten, E. (2015). *Dijken van Nederland*. Nai010 uitgevers.

Poussin, J., Bubeck, P., Aerts, J., & Ward, P. (2012). Potential of semi-structural and non-structural adaptation strategies to reduce future flood risk: Case study for the Meuse. *Natural Hazards and Earth System Sciences*, *12*(11), 3455–3471. https://doi.org/10.5194/nhess-12-3455-2012

Province of South Holland. (2023). *Summary of progress reports on regional flood defences: South Holland water authorities 2023*. [In Dutch]

Rentschler, J., Salhab, M., & Jafino, B. A. (2022). Flood exposure and poverty in 188 countries. *Nature Communications*, *13*(1), 3527. https://doi.org/10.1038/s41467-022-30727-4

Rianna, G., Zollo, A., Tommasi, P., Paciucci, M., Comegna, L., & Mercogliano, P. (2014). Evaluation of the effects of climate changes on landslide activity of Orvieto clayey slope. *Procedia Earth and Planetary Science*, *9*, 54–63. https://doi.org/10.1016/j.proeps.2014.06.017

Ridley, A., McGinnity, B., & Vaughan, P. (2004). Role of pore water pressures in embankment stability. *Proceedings of the Institution of Civil Engineers - Geotechnical Engineering*, *157*(4), 193–198. https://doi.org/10.1680/geng.2004.157.4.193

Rikkert, S. J. H. (2022). *A system perspective on flood risk in polder drainage canal systems* [Doctoral dissertation, Delft University of Technology]. https://doi.org/10.4233/uuid:61e6645a-f275-4b69-aa29-1150d3cda2b1

Rikkert, S. J. H., Kok, M., Lendering, K., & Jongejan, R. (2022). A pragmatic, performance-based approach to levee safety assessments. *Journal of Flood Risk Management*, *15*(4), e12836. https://doi.org/10.1111/jfr3.12836

Robinson, J. D., & Vahedifard, F. (2016). Weakening mechanisms imposed on California's levees under multiyear extreme drought. *Climatic Change*, *137*(1–2), 1–14. https://doi.org/10.1007/s10584-016-1649-6

Robinson, J. D., Vahedifard, F., & AghaKouchak, A. (2017). Rainfall-triggered slope instabilities under a changing climate: Comparative study using historical and projected precipitation extremes. *Canadian Geotechnical Journal*, *54*(1), 117–127. https://doi.org/10.1139/cgj-2015-0602

Rouainia, M., Helm, P., Davies, O., & Glendinning, S. (2020). Deterioration of an infrastructure cutting subjected to climate change. *Acta Geotechnica*, *15*(11), 2997–3016. https://doi.org/10.1007/s11440-020-00965-1

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Schmertmann, J. H. (1991). The mechanical aging of soils. *Journal of Geotechnical Engineering*, *117*(9), 1288–1330. https://doi.org/10.1061/(ASCE)0733-9410(1991)117:9(1288)

Schulte, R. P. O., Diamond, J., Finkele, K., Holden, N. M., & Brereton, A. J. (2005). Predicting the soil moisture conditions of Irish grasslands. *Irish Journal of Agricultural and Food Research, 44*(1), 95–110.

Schultz, B. (2024). Sustainable development of the Netherlands lowlands in cultural context. In *Proceedings of the 14th International Drainage Workshop* (pp. 21–28). International Commission on Irrigation and Drainage. https://icid-ciid.org/icid_data_web/Proceedings-14thIDW2024.pdf#page=21

Schultz, M., Gouldby, B., Simm, J., & Wibowo, J. L. (2010). *Beyond the factor of safety: Developing fragility curves to characterize system reliability* (ERDC SR-10-1). U.S. Army Corps of Engineers. https://doi.org/10.21236/ADA525580

Schweckendiek, T. (2010). *Reassessing reliability based on survived loads*. In *Coastal Engineering Proceedings, 32*(structures), 23. https://doi.org/10.9753/icce.v32.structures.23

Schweckendiek, T. (2014). *On reducing piping uncertainties: A Bayesian decision approach* [Doctoral dissertation, Delft University of Technology]. https://doi.org/10.4233/uuid:f9be2f7e-7009-4c73-afe5-8b4bb16e956f

Schweckendiek, T., Slomp, R., & Knoeff, H. (2015). New safety standards and assessment tools in the Netherlands. In *Proceedings of the 5th Siegener Symposium „Sicherung von Dämmen, Deichen und Stauanlagen*" (pp. 19–20).

Schweckendiek, T., van der Krogt, M. G., Teixeira, A., Kanning, W., Brinkman, R., & Rippi, K. (2017). Reliability updating with survival information for dike slope stability using fragility curves. In *Geo-Risk 2017: Reliability-Based Design and Code Developments* (pp. 494–503). American Society of Civil Engineers. https://doi.org/10.1061/9780784480717.046

Schweckendiek, T., & Vrouwenvelder, A. C. W. M. (2013). Reliability updating and decision analysis for head monitoring of levees. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 7*(2), 110–121. https://doi.org/10.1080/17499518.2013.791034

Schweckendiek, T., Vrouwenvelder, A. C. W. M., & Calle, E. O. F. (2014). Updating piping reliability with field performance observations. *Structural Safety, 47*, 13–23. https://doi.org/10.1016/j.strusafe.2013.10.002

Schwalm, C. R., Anderegg, W. R., Michalak, A. M., Fisher, J. B., Biondi, F., Koch, G., Litvak, M., Ogle, K., Shaw, J. D., Wolf, A., Huntzinger, D. N., Schaefer, K., Cook, R., Wei, Y., Fang, Y., Hayes, D., Huang, M., Jain, A., & Tian, H. (2017). Global patterns of drought recovery. *Nature, 548*(7666), 202–205. https://doi.org/10.1038/nature23021

Senatilleke, U., Sirisena, J., Gunathilake, M. B., Muttil, N., & Rathnayake, U. (2023). Monitoring the meteorological and hydrological droughts in the largest river basin (Mahaweli River) in Sri Lanka. *Climate, 11*(3), 57. https://doi.org/10.3390/cli11030057

Shao, W., Bogaard, T. A., Bakker, M., & Greco, R. (2015). Quantification of the influence of preferential flow on slope stability using a numerical modelling approach. *Hydrology and Earth System Sciences, 19*(5), 2197–2212. https://doi.org/10.5194/hess-19-2197-2015

Sharp, M., Wallis, M., Deniaud, F., Hersch-Burdick, R., Tourment, R., Matheu, E., Seda-Sanabria, Y., Wersching, S., Veylon, G., Durand, E., & Simm, J. (2013). *The international levee handbook*. CIRIA.

Shen, H., Tolson, B. A., & Mai, J. (2022). Time to update the split-sample approach in hydrological model calibration. *Water Resources Research, 58*(3), e2021WR031523. https://doi.org/10.1029/2021WR031523

Šimůnek, J., Šejna, M., & van Genuchten, M. T. (1999). *The HYDRUS-2D software package for simulating the two-dimensional movement of water, heat, and multiple solutes in variably-saturated media: Version 2.0*. U.S. Salinity Laboratory, Agricultural Research Service, U.S. Department of Agriculture.

Simpson, B., Pappin, J. W., & Croft, D. D. (1981). An approach to limit state calculations in geotechnics. *Ground Engineering*, *14*(6), 21–28.

Slomp, R., Knoeff, H., Bizzarri, A., Bottema, M., & de Vries, W. (2016). Probabilistic flood defence assessment tools. In *E3S Web of Conferences* (Vol. 7, p. 03015). EDP Sciences. https://doi.org/10.1051/e3sconf/20160703015

Stafleu, J., Maljers, D., Busschers, F. S., Gunnink, J. L., Schokker, J., Dambrink, R. M., Hummelman, H. J., & Schijf, M. L. (2012). *GeoTop modellering* (TNO Report 10991). TNO Geological Survey of the Netherlands.

Steenbergen, C. M., Reh, W., Nijhuis, S., & Pouderoijen, M. T. (2009). *The polder atlas of The Netherlands: Pantheon of the lowlands*. Uitgeverij Thoth.

Steenbergen, H. M. G. M., Lassing, B. L., Vrouwenvelder, A. C. W. M., & Waarts, P. H. (2004). Reliability analysis of flood defence systems. *Heron*, *49*(1), 51–73.

Stirling, R. A., Toll, D. G., Glendinning, S., Helm, P. R., Yildiz, A., Hughes, P. N., & Asquith, J. D. (2021). Weather-driven deterioration processes affecting the performance of embankment slopes. *Géotechnique*, *71*(11), 957–969. https://doi.org/10.1680/jgeot.19.SIP.038

STOWA. (2023). *Soil parameters for the assessment and design of flood defences* (STOWA Report 2023-37). Foundation for Applied Water Research.

Straub, D. (2011). Reliability updating with equality information. *Probabilistic Engineering Mechanics*, *26*(2), 254–258. https://doi.org/10.1016/j.probengmech.2010.08.003

Straub, D., & Papaioannou, I. (2015). Bayesian updating with structural reliability methods. *Journal of Engineering Mechanics*, *141*(3), 04014134. https://doi.org/10.1061/(ASCE)EM.1943-7889.0000839

Strijker, B. (2023). *Multi-year monitoring data of subsurface water conditions for Dutch canal dikes* (Version 1) [Dataset]. 4TU.ResearchData. https://doi.org/10.4121/136aa5df-1907-43ac-a5b0-e0ea8f2dedf3.v1

Strijker, B. (2024). *Dataset in support of the dynamics of peak head responses at Dutch canal dikes and the impact of climate change* [Dataset]. 4TU.ResearchData. https://doi.org/10.4121/4004f445-b71b-4996-bd7d-10b1fafbc86b

Strijker, B., Asselman, N., de Jong, J., & Barneveld, H. (2023). The 2021 flood event in the Dutch Meuse and tributaries from a hydraulic and morphological perspective. *Journal of Coastal and Riverine Flood Risk*, *2*, 6. https://doi.org/10.59490/jcrfr.2023.0006

Strijker, B., Heimovaara, T. J., Jonkman, S. N., & Kok, M. (2024). Exploring subsurface water conditions in Dutch canal dikes during drought periods: Insights from multiyear monitoring. *Water Resources Research*, *60*(9), e2023WR036046. https://doi.org/10.1029/2023WR036046

Strijker, B., & Kok, M. (2023). The joint impact of rainfall events on water- and dike systems in Dutch polders. In *EGU General Assembly 2023* (EGU23-9048). Copernicus. https://doi.org/10.5194/egusphere-egu23-9048

Strijker, B., & Kok, M. (2025). The dynamics of peak head responses at Dutch canal dikes and the impact of climate change. *Natural Hazards and Earth System Sciences*, *25*(9), 3355–3379. https://doi.org/10.5194/nhess-25-3355-2025

TAW – Technical Advisory Committee on Flood Defences. (1989). *Guideline for the design of river dikes: Part 2, Tidal River Area* (ISBN 9021231689). [In Dutch]

TAW – Technical Advisory Committee on Flood Defences. (2004). *Technisch rapport waterspanningen bij dijken*. [In Dutch]

Terzaghi, K., Peck, R. B., & Mesri, G. (1996). *Soil mechanics in engineering practice* (3rd ed.). John Wiley & Sons.

Tran, D. D., & Weger, J. (2018). Barriers to implementing irrigation and drainage policies in An Giang Province, Mekong Delta, Vietnam. *Irrigation and Drainage, 67*(S1), 81–95. https://doi.org/10.1002/ird.2172

Trenberth, K. E. (2011). Changes in precipitation with climate change. *Climate Research, 47*(1–2), 123–138. https://doi.org/10.3354/cr00953

Triet, N. V. K., Dung, N. V., Fujii, H., Kummu, M., Merz, B., & Apel, H. (2017). Has dyke development in the Vietnamese Mekong Delta shifted flood hazard downstream? *Hydrology and Earth System Sciences, 21*(8), 3991–4010. https://doi.org/10.5194/hess-21-3991-2017

UNISDR. (2011). *Global assessment report on disaster risk reduction: Revealing risk, redefining development*. United Nations International Strategy for Disaster Reduction Secretariat.

U.S. Army Corps of Engineers. (2003). *Slope stability—Engineering manual* (EM 1110-2-1902). https://www.publications.usace.army.mil/Portals/76/Publications/EngineerManuals/EM_1110-2-1902.pdf

Vahedifard, F., Robinson, J. D., & AghaKouchak, A. (2016). Can protracted drought undermine the structural integrity of California's earthen levees? *Journal of Geotechnical and Geoenvironmental Engineering, 142*(6), 02516001. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001465

Van Baars, S. (2005). The horizontal failure mechanism of the Wilnis peat dyke. *Géotechnique, 55*(4), 319–323. https://doi.org/10.1680/geot.55.4.319.65491

Van Baars, S., & Van Kempen, I. M. (2009). The causes and mechanisms of historical dike failures in the Netherlands. *E-Water Official Publication of the European Water Association*, 1–14. https://doi.org/10.2175/193864702784164505

Van der Krogt, M. G., Klerk, W. J., Kanning, W., Schweckendiek, T., & Kok, M. (2022). Value of information of combinations of proof loading and pore pressure monitoring for flood defences. *Structure and Infrastructure Engineering, 18*(4), 505–520. https://doi.org/10.1080/15732479.2020.1857794

Van der Krogt, M. G., Schweckendiek, T., & Kok, M. (2019). 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*(3), 226–236. https://doi.org/10.1080/17499518.2018.1554820

Van der Meer, J. W., Ter Horst, W. L. A., & Van Velzen, E. H. (2009). Calculation of fragility curves for flood defence assets. In *Flood Risk Management: Research and Practice* (pp. 567–573). Taylor & Francis. https://doi.org/10.1201/9780203883020.ch65

Van der Meer, M. T., Kapinga, H. S. O., Jonkman, L. N., Lam, K. S., & Calle, E. O. F. (2009). *Material factors for canal dikes* (STOWA Report 2009-05; ISBN 978.90.5773.420.5). [In Dutch]

Van der Meij, R. (2023). *D-Stability user manual: Version 2024.01* (Document ID 2024.01; Doc. version 2023.2). Deltares. [In Dutch]

van der Meulen, M. J., van der Spek, A. J., de Lange, G., Gruijters, S. H., van Gessel, S. F., Nguyen, B. L., Maljers, D., Schokker, J., Mulder, J. P. M., & van der Krogt, R. A. A. (2007). Regional sediment

deficits in the Dutch lowlands: Implications for long-term land-use options. *Journal of Soils and Sediments*, *7*(1), 9–16. https://doi.org/10.1065/jss2006.12.199

Van der Wiel, K., Beersma, J., van den Brink, H., Krikken, F., Selten, F., Severijns, C., Siegmund, P., Bakker, A., Beersma, J., Bessembinder, J., Bloemendaal, N., Brotons Blanes, H., Drijfhout, S., Groenland, R., Haarsma, R., Homan, C., Keizer, I., Le Bars, D., Lenderink, G., …van Dorland, R. (2024). KNMI'23 climate scenarios for the Netherlands: Storyline scenarios of regional climate change. *Earth's Future*, *12*(2), e2023EF003983. https://doi.org/10.1029/2023EF003983

Van Dorland, R., Beersma, J., Bessembinder, J., Bloemendaal, N., van den Brink, H., Brotons Blanes, H., Drijfhout, S., Groenland, R., Haarsma, R., Homan, C., Keizer, I., Krikken, F., Le Bars, D., Lenderink, G., van Meijgaard, E., Meirink, J. F., Overbeek, B., Reerink, T., Selten, F., …van der Wiel, K. (2023). *KNMI national climate scenarios 2023 for The Netherlands* (Scientific Report WR-23-02). Royal Netherlands Meteorological Institute.

Van Esch, J. M. (2012). Modeling groundwater flow through embankments for climate change impact assessment. In *Proceedings, 19th International Conference on Computational Methods in Water Resources (CMWR 2012)*, University of Illinois at Urbana-Champaign.

Van Essen. (2016). *Diver product manual*. https://www.vanessen.com/wp-content/uploads/2022/04/TD-Diver-DI8xx-ProductManual-en.pdf

Van Etten, R. J. G., & Förster, U. (2006). *Exploration of peat dikes: Changes over time and characteristic profiles* (Report No. DWW-2005-087). Rijkswaterstaat, Directorate General for Public Works and Water Management. [In Dutch]

Van Gelder, P. H. A. J. M. (2000). *Statistical methods for the risk-based design of civil structures* [Doctoral dissertation, Delft University of Technology].

Van Loon, A. F. (2015). Hydrological drought explained. *Wiley Interdisciplinary Reviews: Water*, *2*(4), 359–392. https://doi.org/10.1002/wat2.1085

Van, M. A., Koelewijn, A. R., & Barends, F. B. (2005). Uplift phenomenon: Model, validation, and design. *International Journal of Geomechanics*, *5*(2), 98–106. https://doi.org/10.1061/(ASCE)1532-3641(2005)5:2(98)

Vanmarcke, E. H. (1977). Reliability of earth slopes. *Journal of the Geotechnical Engineering Division*, *103*(11), 1247–1265. https://doi.org/10.1061/AJGEB6.0000518

van Woerkom, T., van Beek, R., Middelkoop, H., & Bierkens, M. F. (2021). Global sensitivity analysis of groundwater related dike stability under extreme loading conditions. *Water*, *13*(21), 3041. https://doi.org/10.3390/w13213041

van Woerkom, T., van Beek, R., Middelkoop, H., & Bierkens, M. F. (2022). Assessing lithological uncertainty in dikes: Simulating construction history and its implications for flood safety assessment. *Journal of Flood Risk Management*, *15*(4), e12848. https://doi.org/10.1111/jfr3.12848

Vardon, P. J. (2015). Climatic influence on geotechnical infrastructure: A review. *Environmental Geotechnics*, *2*(3), 166–174. https://doi.org/10.1680/envgeo.13.00055

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, *23*(7), 1696–1718. https://doi.org/10.1175/2009JCLI2909.1

von Asmuth, J. R., Bierkens, M. F., & Maas, K. (2002). Transfer function-noise modeling in continuous time using predefined impulse response functions. *Water Resources Research*, *38*(12), 1287. https://doi.org/10.1029/2001WR001136

von Asmuth, J. R., Maas, K., Bakker, M., & Petersen, J. (2008). Modeling time series of ground water head fluctuations subjected to multiple stresses. *Groundwater, 46*(1), 30–40. https://doi.org/10.1111/j.1745-6584.2007.00382.x

Vrijling, J. K. (2001). Probabilistic design of water defense systems in The Netherlands. *Reliability Engineering & System Safety, 74*(3), 337–344. https://doi.org/10.1016/S0951-8320(01)00082-5

Vrijling, J. K., & van Gelder, P. H. A. J. M. (2002). *Probabilistic design in hydraulic engineering*. Delft University of Technology. https://resolver.tudelft.nl/uuid:0ba44475-700e-436a-8a2d-7df89b85e19a

Vrijling, J. K., Schweckendiek, T., & Kanning, W. (2011). Safety standards of flood defenses. In H. P. Hong, D. Frangopol, & M. Su (Eds.), *Proceedings of the 11th International Conference on Structural Safety and Reliability* (pp. 67–84).

Vrouwenvelder, A. C. W. M. (2003). Keynote lecture: Uncertainty analysis for flood defence systems in the Netherlands. In R. Cooke & M. Mendel (Eds.), *Proceedings of ESREL 2003* (pp. 11–18). CRC Press/Balkema.

Vrouwenvelder, A. C. W. M., & Calle, E. (2003). Measuring spatial correlation of soil properties. *Heron, 48*(4), 297–311.

Vrouwenvelder, T., Dimova, S., Sousa, L., Marková, J., Mancini, G., Kuhlmann, U., Taras, A., Croce, P., Matos, J., Clifton, C., Priebe, M., Spaethe, G., Unterweger, H., Mensinger, M., Bouchair, A., Faber, M. H., Kote, P., Lenzi, C., Sykora, M., …Gulvanessian, H. (2024). *Reliability background of the Eurocodes: Support to the implementation, harmonization and further development of the Eurocodes*. Publications Office of the European Union. https://doi.org/10.2760/948283

Wageningen Environmental Research. (2023). *Bodemkaart van Nederland 1:50 000 (versie 2023-01)* [Dataset]. Basisregistratie Ondergrond. https://www.pdok.nl/

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B., Janssen, P., & Krayer von Krauss, M. P. (2003). Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment, 4*(1), 5–17. https://doi.org/10.1076/iaij.4.1.5.16466

Warner, J. F., van Staveren, M. F., & van Tatenhove, J. (2018). Cutting dikes, cutting ties? Reintroducing flood dynamics in coastal polders in Bangladesh and The Netherlands. *International Journal of Disaster Risk Reduction, 32*, 106–112. https://doi.org/10.1016/j.ijdrr.2018.03.020

Williams, E. (1957). Some observations of Leonardo, Galileo, Mariotte and others relative to size effect. *Annals of Science, 13*(1), 23–29. https://doi.org/10.1080/00033795700200147

Wojciechowska, K. A. (2015). *Advances in operational flood risk management in the Netherlands* [Doctoral dissertation, Delft University of Technology]. https://doi.org/10.4233/uuid:5d719beb-bbbf-4fde-8e10-526785315fd1

Wolff, R. G., & Olsen, H. W. (1968). Piezometer for monitoring rapidly changing pore pressures in saturated clays. *Water Resources Research, 4*(4), 839–843. https://doi.org/10.1029/WR004i004p00839

Wolters, E., Hakvoort, H., Bosch, S., Versteeg, R., Bakker, M., Heijkers, J., Talsme, M., & Peerdeman, K. (2013). Meteobase: Online neerslag-en referentiegewasver-dampingsdatabase voor het Nederlandse waterbeheer. *Meteorologica, 1*, 15–18.

Wong, T. E., Batjes, D. A. J., & de Jager, J. (2007). *Geology of the Netherlands*. Royal Netherlands Academy of Arts and Sciences.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, *20*, 557–585.

Yang, Z., Yin, C., Li, X., Jiang, S., & Li, D. (2024). Efficient slope reliability and sensitivity analysis using quantile-based first-order second-moment method. *Journal of Rock Mechanics and Geotechnical Engineering*, *16*(11), 4192–4203. https://doi.org/10.1016/j.jrmge.2024.04.007

Yang, Z., Li, D., Jiang, S., Sheng, D., & Ching, J. (2025). Bayesian back analysis of geotechnical parameters for slope reliability updating. *Engineering Geology*, *330*, 107419. https://doi.org/10.1016/j.enggeo.2024.107419

Zhang, J., Zhang, L. M., & Tang, W. H. (2011). Slope reliability analysis considering site-specific performance information. *Journal of Geotechnical and Geoenvironmental Engineering*, *137*(3), 227–238. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000422

Zhang, J. M., Luo, Y., Zhou, Z., Chong, L., Victor, C., & Zhang, Y. F. (2021). Effects of preferential flow induced by desiccation cracks on slope stability. *Engineering Geology*, *288*, 106164. https://doi.org/10.1016/j.enggeo.2021.106164

Appendices

A Supplementary information Chapter 2

Figure A1 shows frequency distributions of measured hydraulic heads using violin plots. Figure A2 displays the relation between measured head ranges and various dike characteristics. Figures A3, A4 and A5 give time series of soil moisture and head levels at the mid-slope of the canal dikes. Figures A6 and A7 visualize the relationship between different meteorological drought indicators and subsurface water measurements in dikes and their performance based on different statistical measures. The figures A8 until A11 show the relationships between *PD* and *SPEI-6* and the subsurface water measurements in more detail by considering scatterplots

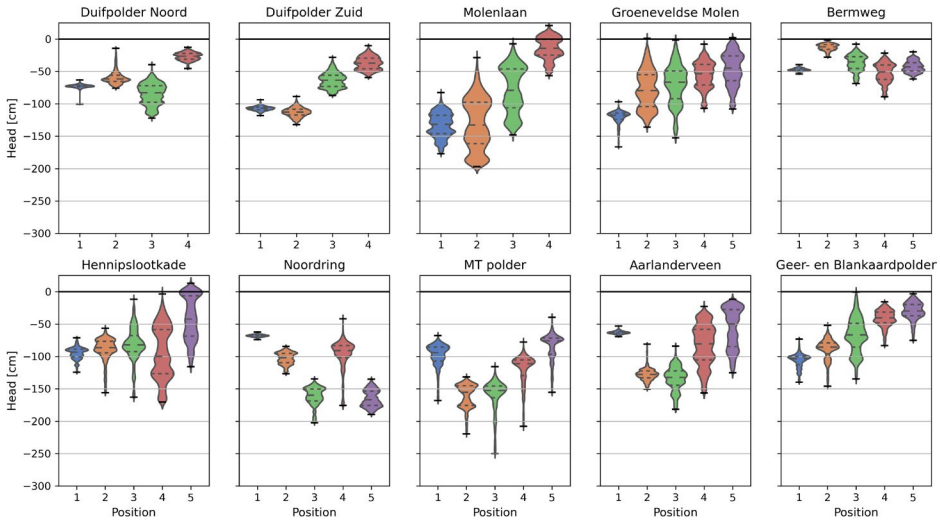


Figure A 1 The violin plots show the variation of the groundwater levels at different monitoring points within the cross-sectional profile: position nr1 is near the crest and nr5 near the toe. The violins indicate the shape of a data set by using the probability density and are scaled in such a way that the width is the same for every monitoring point. The groundwater levels are with respect to the ground surface elevation at the piezometer.

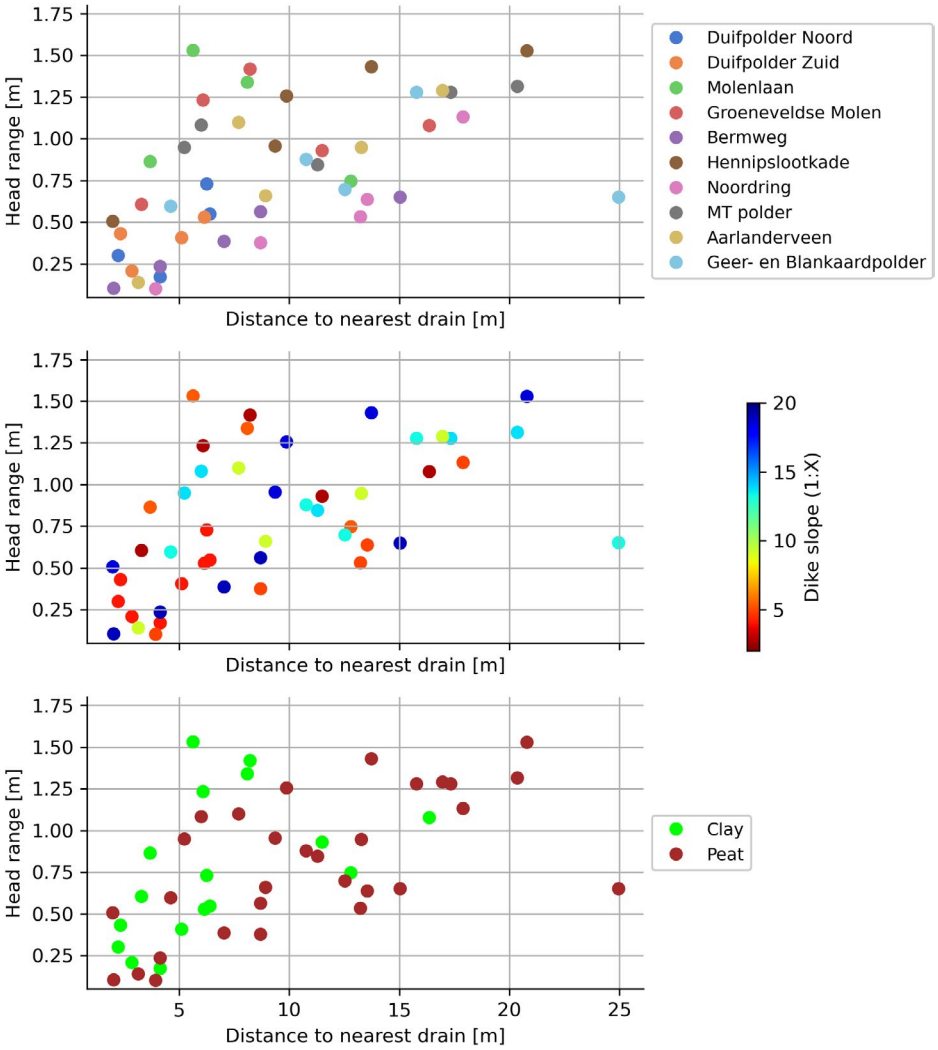


Figure A 2 The measured head range in relation to the distance to the nearest drainage. The top graph illustrates the corresponding monitoring sites, while the middle and lower graphs use colors to indicate the dike slope and predominant soil type (clay or peat).

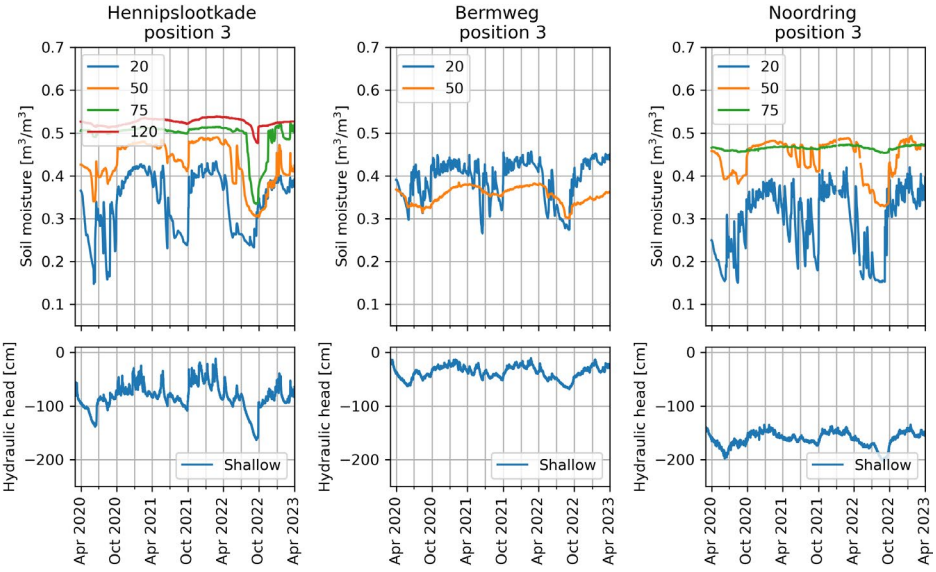


Figure A 3 The time series of soil moisture levels at different depths (upper graphs) and head levels at the talud of the canal dikes w.r.t. the local ground surface level. Only measurements taken on the inner slope / talud are shown at all locations (1/3).

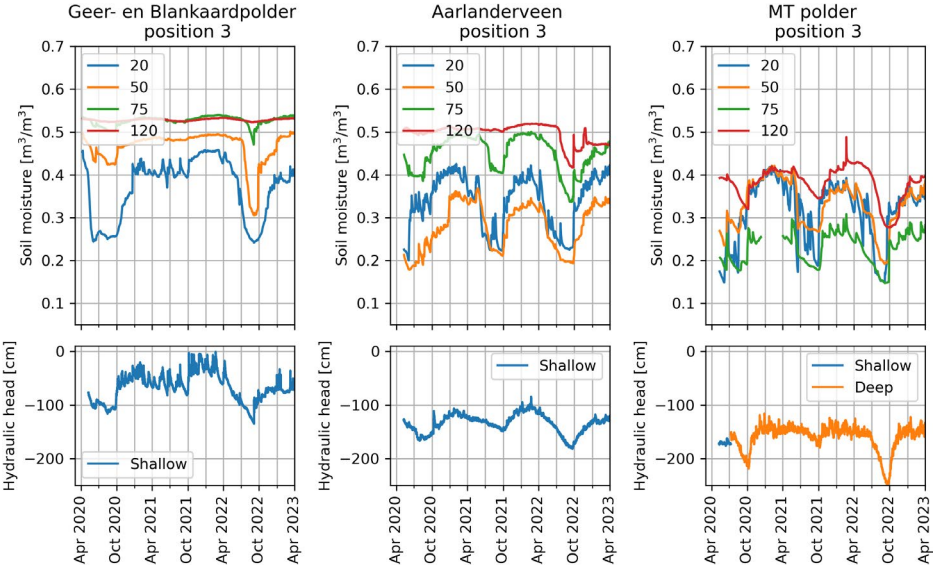


Figure A 4 The time series of soil moisture levels at different depths (upper graphs) and head levels at the talud of the canal dikes w.r.t. the local ground surface level. Only measurements taken on the inner slope / talud are shown at all locations (2/3).

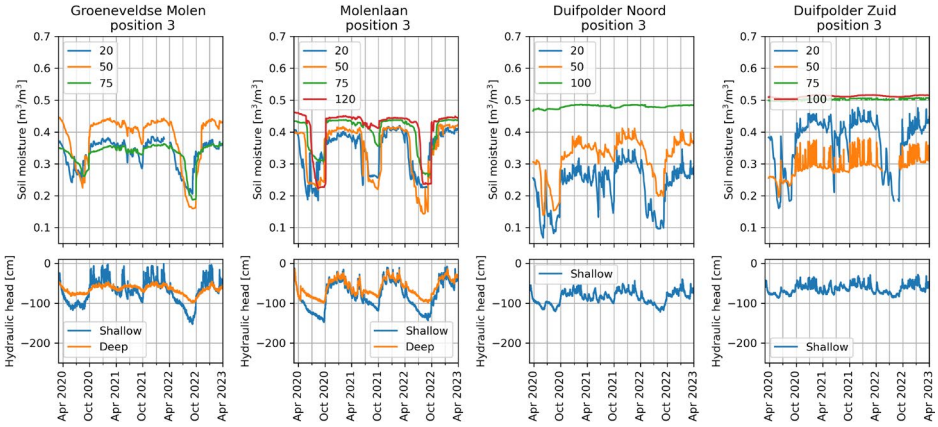


Figure A 5 The time series of soil moisture levels at different depths (upper graphs) and head levels at the talud of the canal dikes w.r.t. the local ground surface level. Only measurements taken on the inner slope / talud are shown at all locations (3/3).

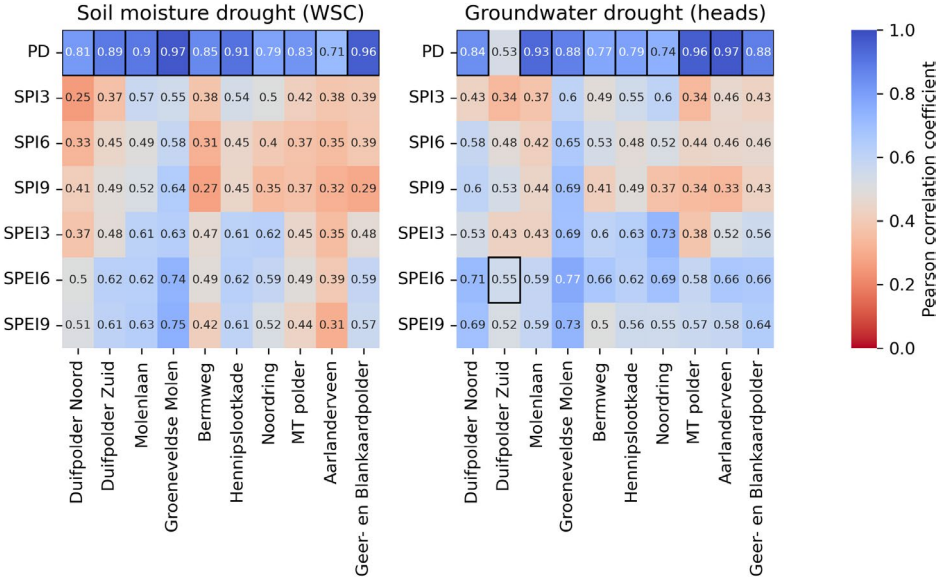


Figure A 6 The relationship between different meteorological drought indicators and subsurface water measurements in dikes was evaluated using the Pearson correlation coefficient.

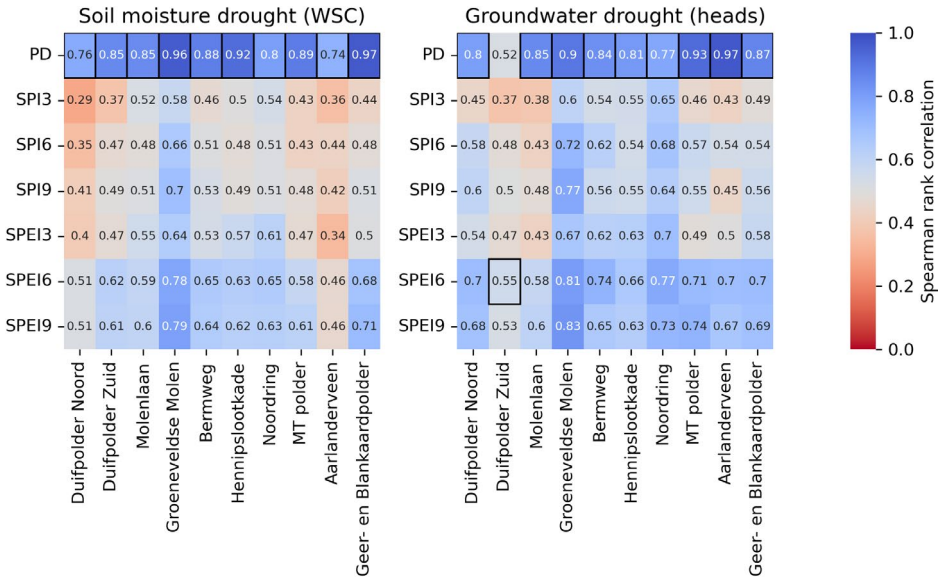


Figure A 7 The relationship between different meteorological drought indicators and subsurface water measurements in dikes was evaluated using the Spearman rank correlation.

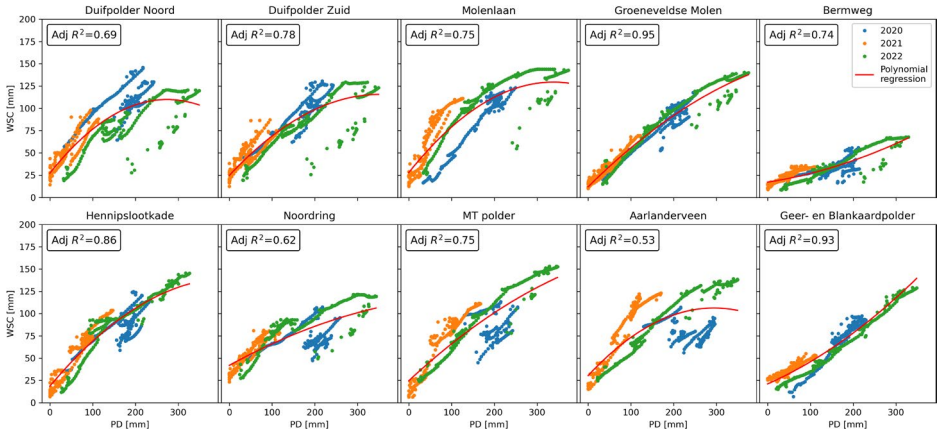


Figure A 8 The relation between precipitation deficit (PD) and water storage capacity (WSC), where distinction has been made between different years. A polynomial regression (red line) has been fitted, based on all years, and the adjusted R-squared is used to evaluate the goodness of fit.

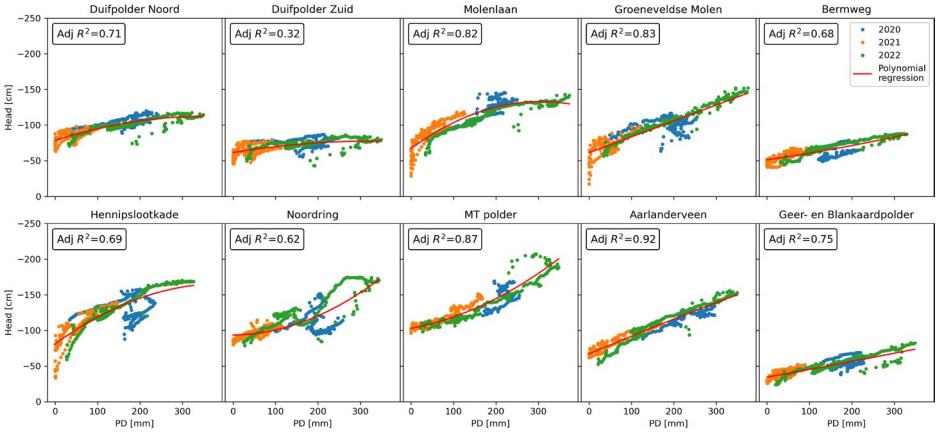


Figure A 9 The relation between precipitation deficit (PD) and the hydraulic head level (in cm below local surface level), where distinction has been made between different years. A polynomial regression (red line) has been fitted, based on all years, and the adjusted R-squared is used to evaluate the goodness of fit.

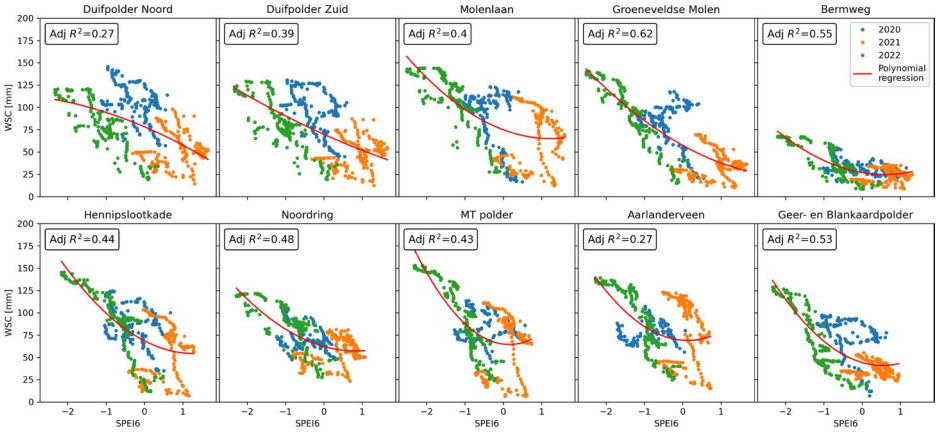


Figure A 10 The relation between SPEI-6 and water storage capacity (WSC), where distinction has been made between different years. A polynomial regression (red line) has been fitted, based on all years, and the adjusted R-squared is used to evaluate the goodness of fit.

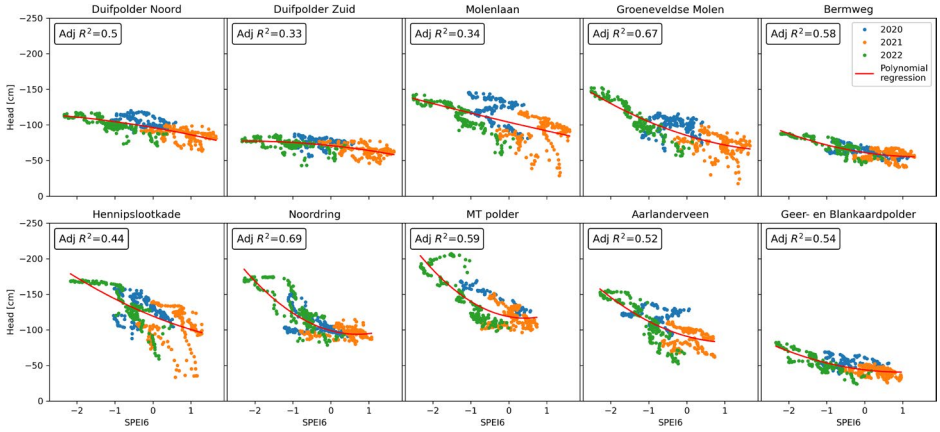


Figure A 11 The relation between SPEI-6 and the hydraulic head level (in cm below local surface level), where distinction has been made between different years. A polynomial regression (red line) has been fitted, based on all years, and the adjusted R-squared is used to evaluate the goodness of fit.

B Supplementary information Chapter 3

B.1 Time series models

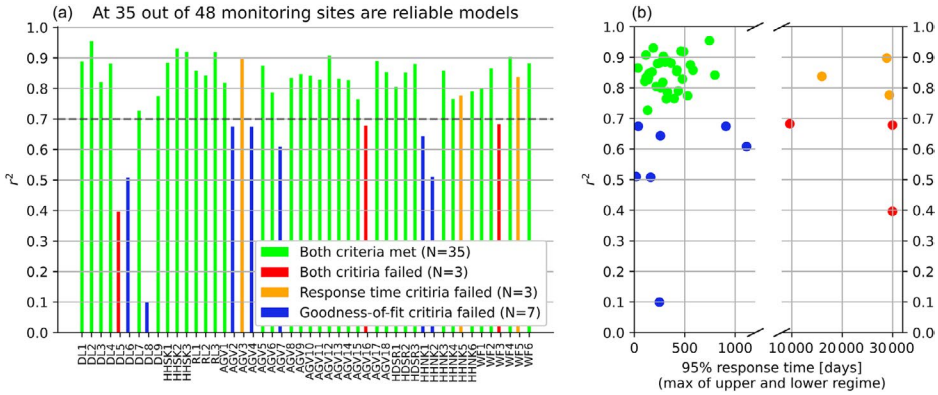


Figure B 1 (a) r^2 values of the best-performing TARSO models at each monitoring site, with colours indicating whether the models meet the reliability criteria. (b) Corresponding response times (maximum for both the upper and lower regimes) associated with the R^2 values.

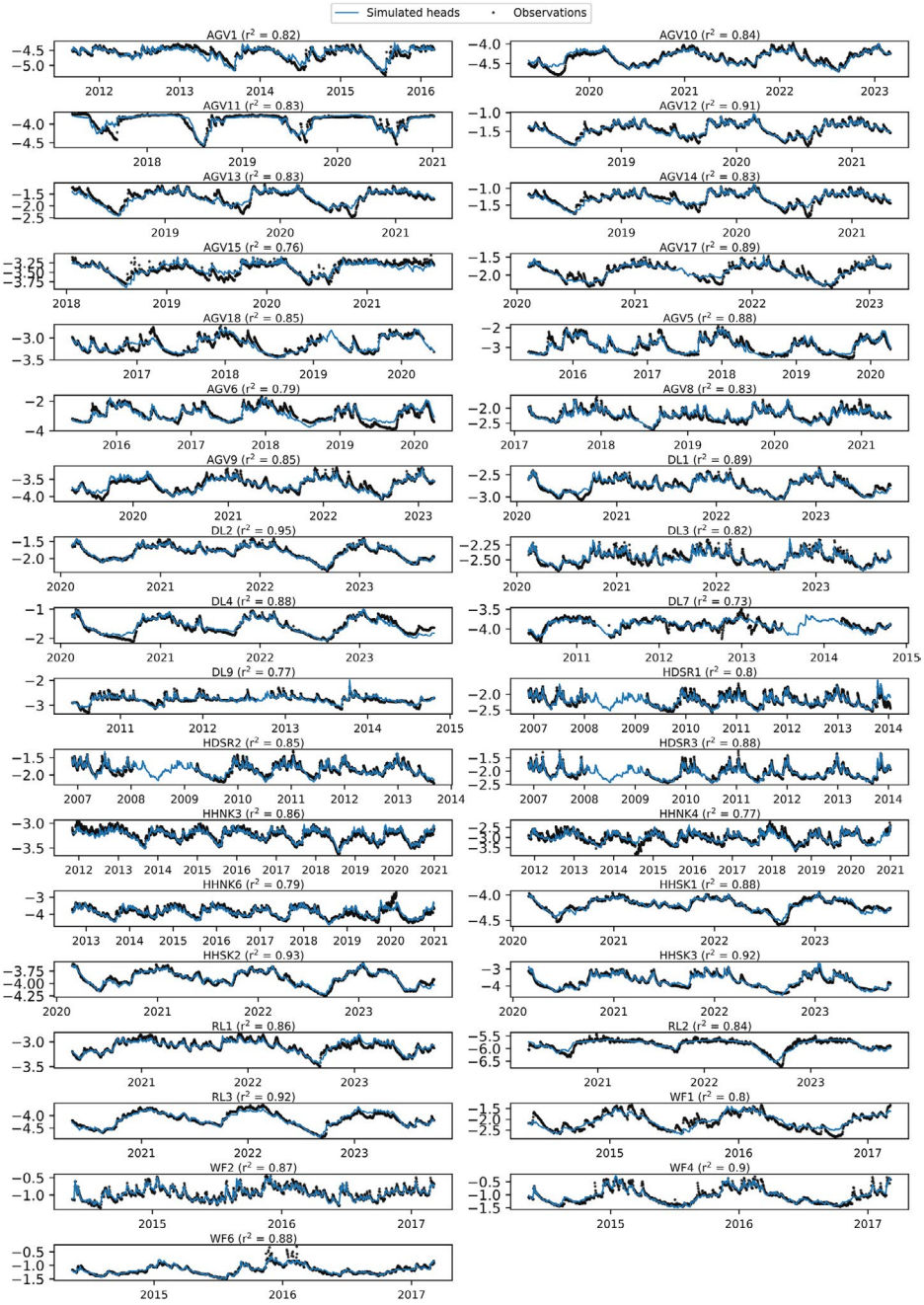


Figure B 2 The observed and simulated heads for all monitoring wells.

B.2 Seasonality

The dynamics of the peak heads were analysed by quantifying the seasonality of the dikes, which is measured by using the average timing and temporal concentration of the selected head peaks. The method to determine the average timing of head peaks involves circular statistics, and it is extensively described in Hall and Blöschl (2018). The average timing of the head peaks is the average date on which peaks have occurred during the time series. The average timing of head peaks can be the result of a wide range of peak dates during the year; therefore, the temporal concentration of peaks occurrence within the year is considered using the concentration index. The concentration index of peak dates around the average timing serves as a measure of how well the seasonality is defined for a specific dike, with 0 indicating evenly distributed peaks during the year and 1 indicating that all peaks occur on the same date.

Seasonality varies across the dikes, but the average timing of these peaks occurs in the winter half-year, as shown in the right graph in Fig. A2. The average timing shifts further into the winter from cluster A to cluster D, with increasing temporal concentrations. This behaviour is also illustrated by the probability density distributions of the peaks of four dikes in the left graph in Fig. A2. The vertical dashed line indicates the average timing, which moves further into the winter, while also the density functions become narrower.

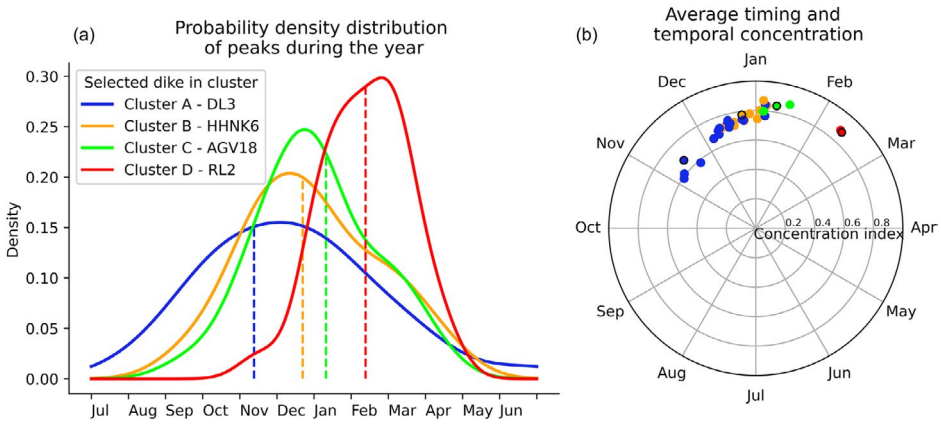


Figure B 3 (b) The average timing and temporal concentrations for the considered dikes, with colours representing the dike clusters (refer to the legend in the left graph). **(a)** The probability density distribution of the peaks throughout the year for four selected dikes from different clusters, with the vertical dashed line indicating the average timing. These four dikes are highlighted with a black edge around the circle in the right graph.

B.3 Spatial patterns in dike clusters

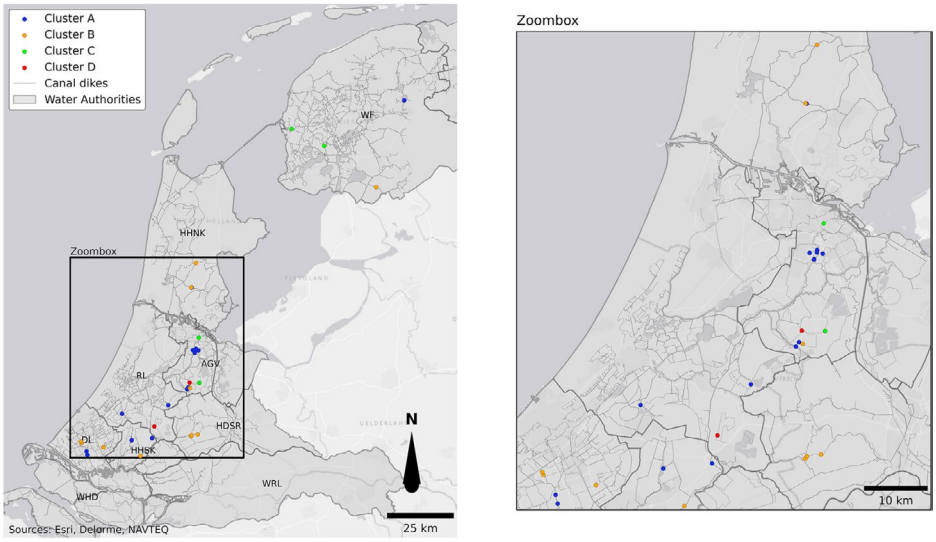


Figure B 4 Mapping of clustered canal dikes across the Netherlands. The left panel shows the full study area, with dikes grouped into four clusters (A–D) based on their peak head response to weather events. The right panel provides a zoomed-in view of the highlighted region, offering a more detailed look at the cluster distribution.

B.4 Physical dike characteristics and their relation to head responses

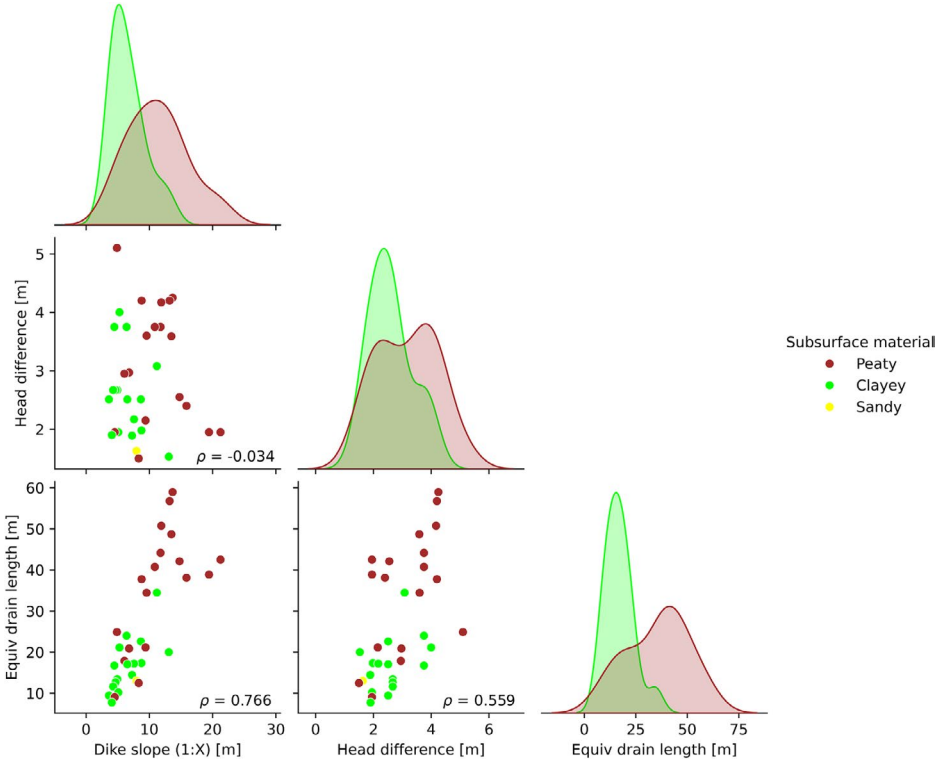


Figure B 5 The relationships between the considered physical dike characteristics are illustrated by scatterplots, with the subsurface material of the dike indicated by colour. The Pearson correlation coefficients between the variables are displayed in the bottom right corner. The diagonal plots show the univariate distributions, highlighting the marginal distribution of each variable, with distinctions made based on soil type.

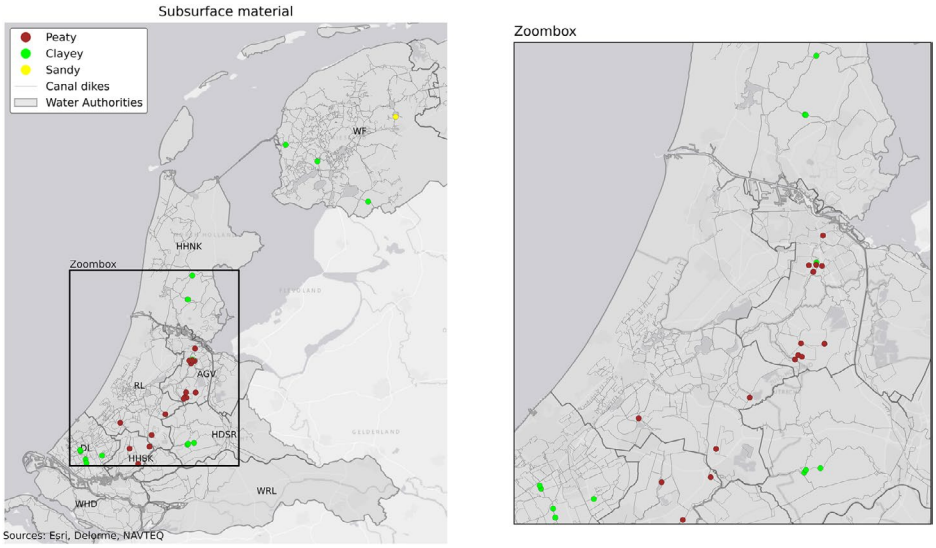


Figure B 6 Mapping of subsurface material (peat, clay and sand) in the 35 canal dikes for which reliable models were developed across the Netherlands.

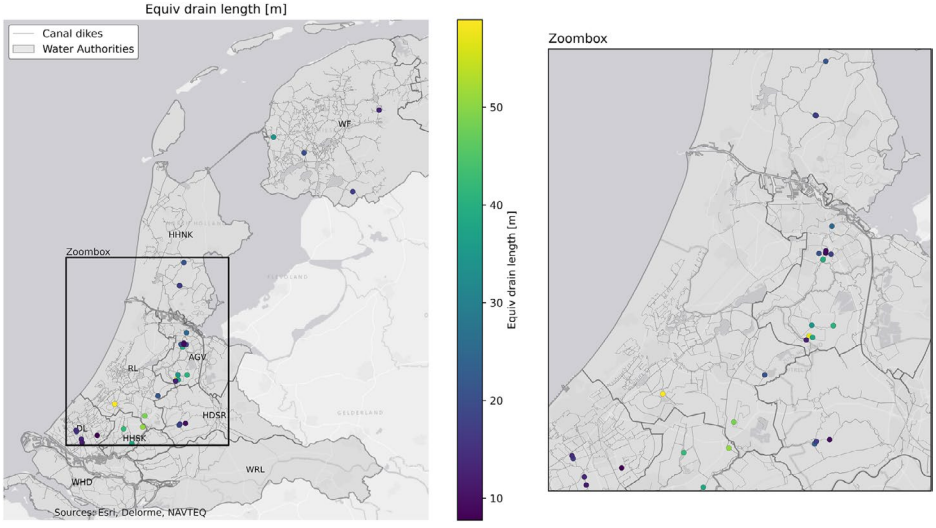


Figure B 7 Mapping of the equivalent drainage length in the 35 canal dikes for which reliable models were developed across the Netherlands.

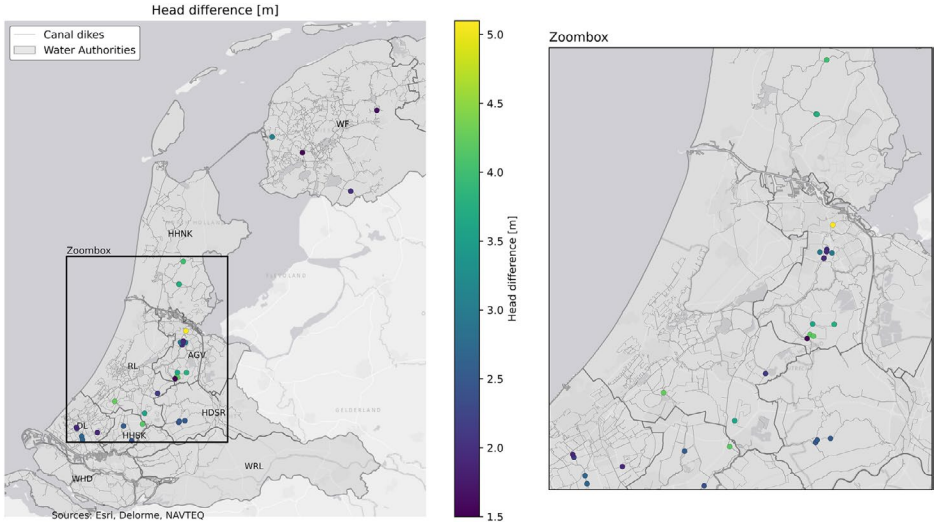


Figure B 8 Mapping of the head differences in the 35 canal dikes for which reliable models were developed across the Netherlands

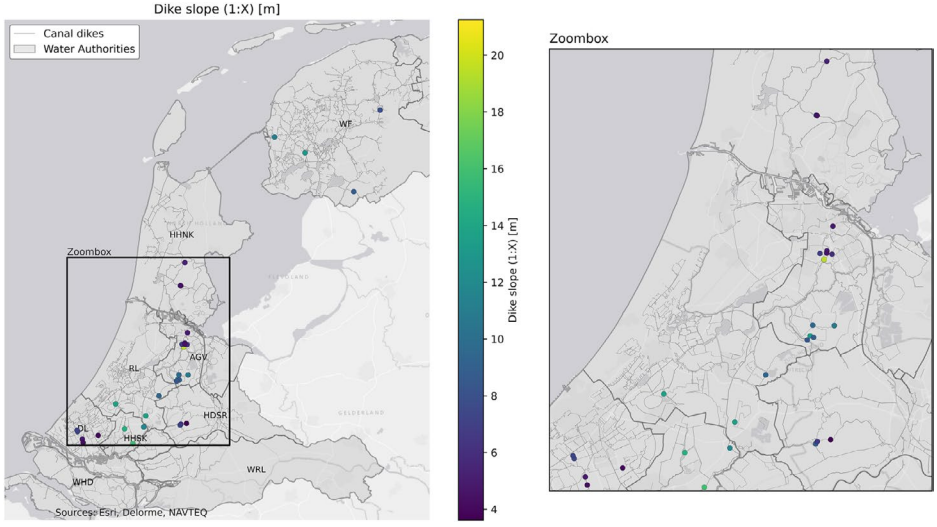


Figure B 9 Mapping of the dike slopes in the 35 canal dikes for which reliable models were developed across the Netherlands.

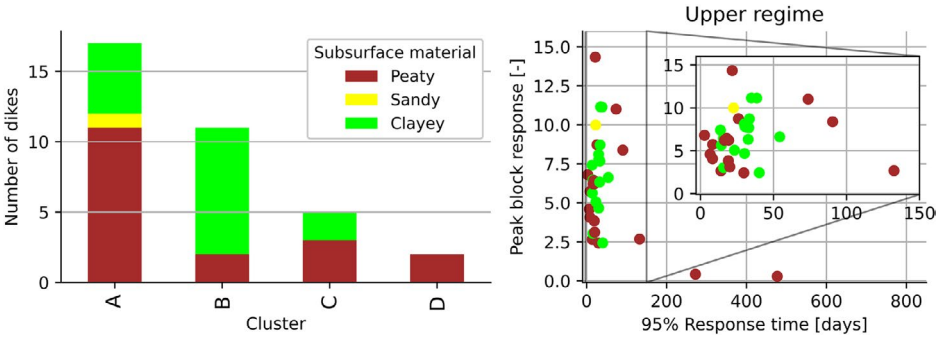


Figure B 10 Stacked bars of the subsurface material of the dike body for the three clusters of dikes. Right: the characteristics of the impulse response functions (95% response time and peak block response) where the colours indicate the subsurface material.

B.5 Spatial variation in decimate heights

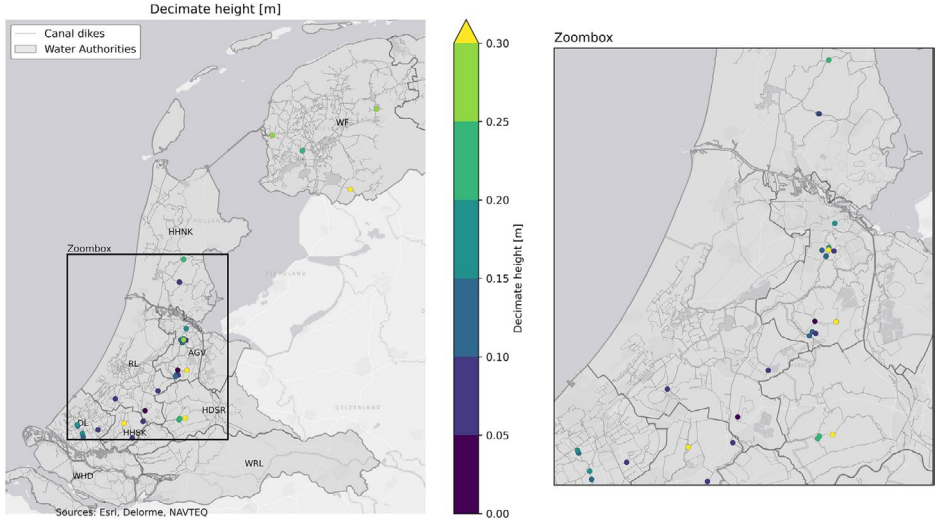


Figure B 11 Spatial variation of decimate height across canal dikes in the Netherlands. The left panel shows the full study area, with different water authorities labelled. The right panel provides a zoomed-in view of the highlighted region. The colour scale represents decimate height values. No clear spatial pattern is observed across different regions.

C Supplementary information Chapter 4

C.1 Rainfall regimes in the Netherlands

To derive the peak head statistics, homogenised long-term measurement series of rainfall and evaporation were used in addition to historical data. These series represent present-day climatic conditions and are available for several precipitation regimes across the Netherlands.The regimes distinguish between extreme rainfall patterns that occur throughout the year and those specific to the winter season. Three annual regimes are defined: L (Low), R (Reference), and H (High). For the winter season, an additional LL (Extra Low) regime is included. The probability of extreme precipitation increases from regime L to H. Because the spatial pattern of rainfall differs between summer and winter, each location can experience a unique combination of summer and winter regimes. This results in ten possible combinations of annual and seasonal regimes across the country. For each combination, a regional long-term series of rainfall and evaporation has been derived to characterize local climatic conditions.

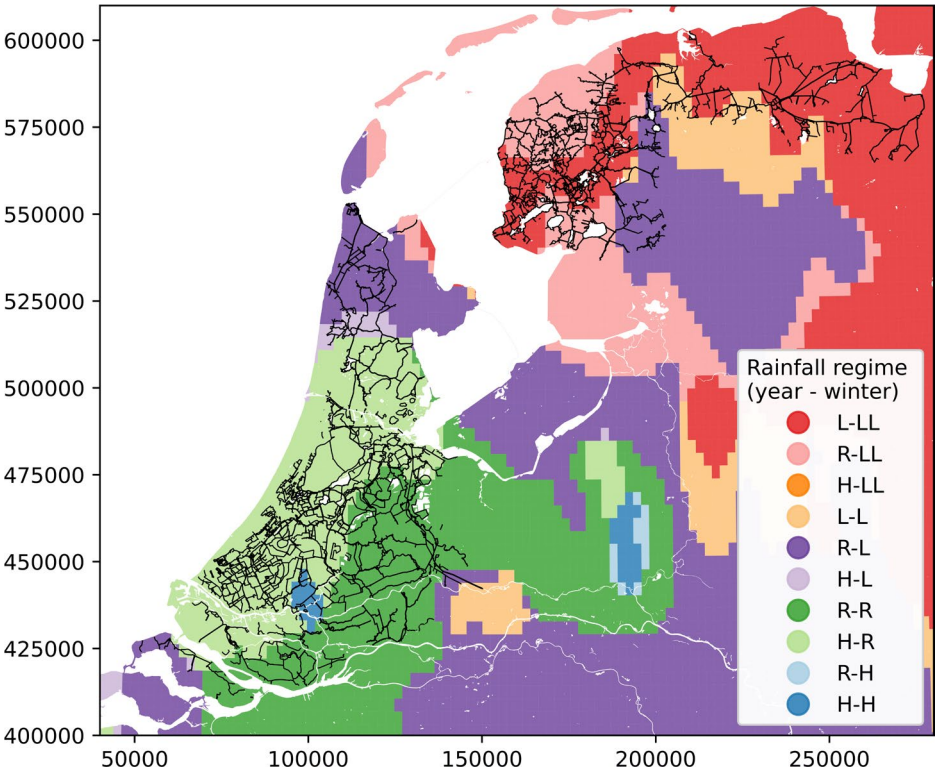
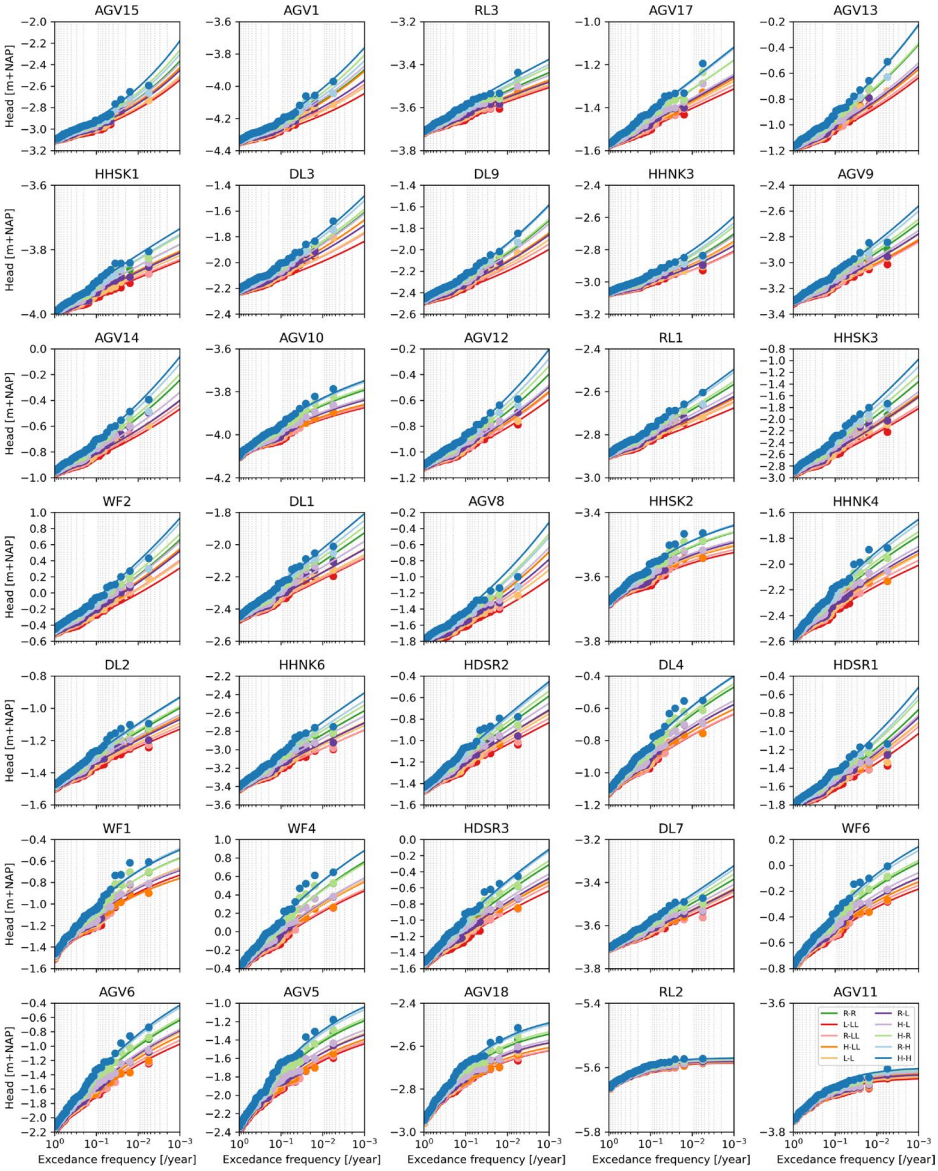


Figure 0.1 Rainfall regimes in the Netherlands based on long-term precipitation and evaporation series. Each color indicates a combination of annual and winter regimes (year–winter), where L, R, and H denote Low, Reference, and High annual regimes, and LL denotes the Extra Low winter regime.

C.2 Head statistics of time series models



D Supplementary information Chapter 5

D.1 Load scenarios in case studies

For every case, at least five load situations are considered, ranging from yearly occurring levels to a situation with a probability of occurrence of 1/10,000 per year, discretized with load levels that differ by a maximum probability factor of 10.

The water level statistics of the various case studies are as follows.

Return period [years]	Aarlanderveen	Duifpolder	Meers	Gennep
1	-2.10	-0.45	33.88	9.23
10	-1.95	-0.375	36.01	11.99
100	-1.80	-0.3	37.37	13.28
1,000	-1.65	-0.225	38.13	14.14
10,000	-1.50	-0.15	38.66	14.82

The head level at the inner-crest in the various case studies is as follows:

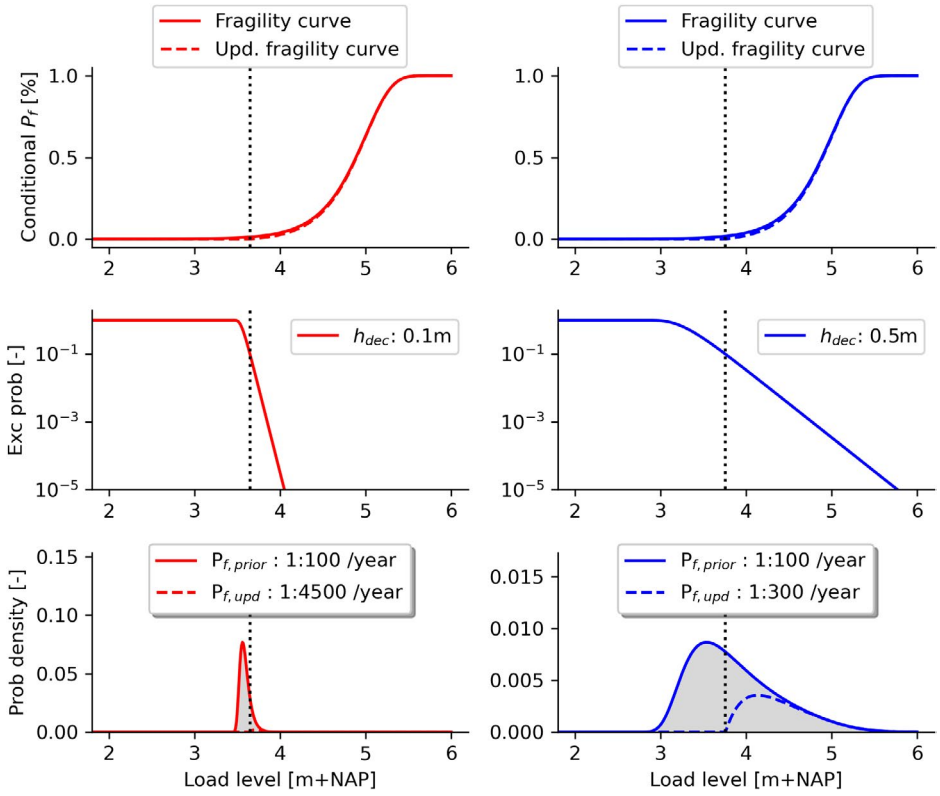
Return period [years]	Aarlanderveen	Duifpolder	Meers	Gennep
1	-2.53	-0.68	33.88	9.8
10	-2.43	-0.61	34.51	10.49
100	-2.33	-0.54	35.87	11.78
1,000	-2.23	-0.47	36.63	12.64
10,000	-1.40	-0.15	37.16	13.92

In three of the four cases, the contribution of the load scenarios with a return period of 10,000 years is small, except for the dike at Meers. Therefore, two additional scenarios were added with average return periods of 30,000 and 100,000 years.

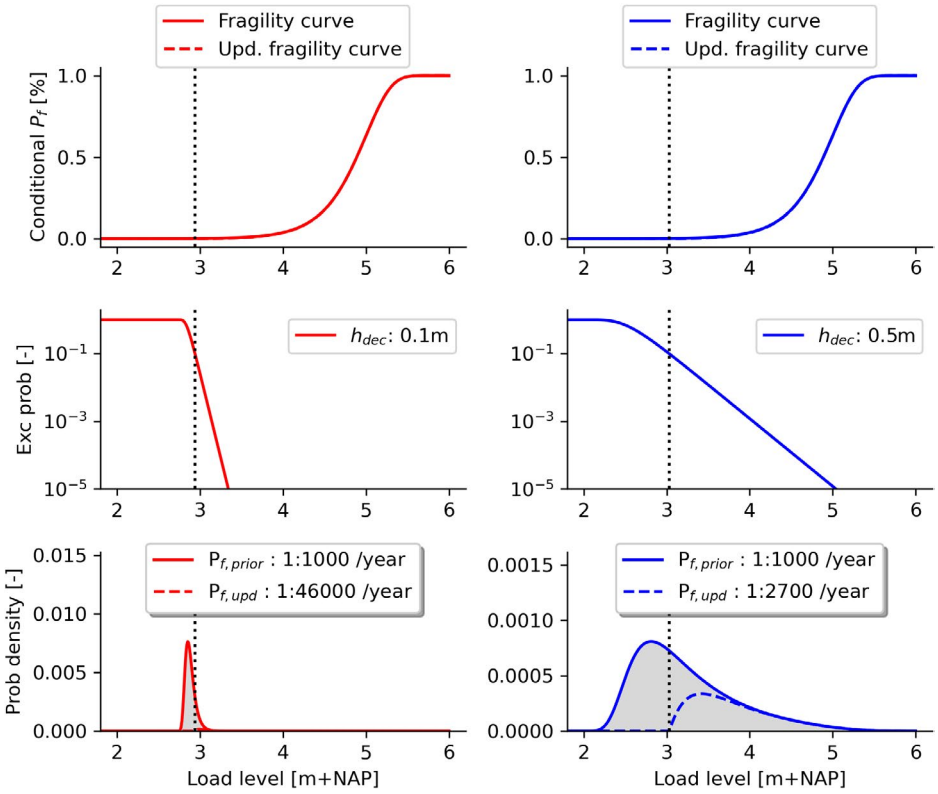
Furthermore, for the Gennep case, two additional load scenarios with return periods of 30 and 300 years were added, because uplift occurred there. This caused a bending fragility curve that was better captured by including these extra scenarios.

D.2 Illustrative example of two dikes

Let's look at two dikes with the same fragility curve and an initial failure probability of 1/100 per year. The key difference lies in the load statistics: one dike faces small load variations (left), while the other faces large load variations (right). Both dikes have survived a load level corresponding to a 10-year return period. After updating, the failure probability is lower for the dike with smaller load variations.

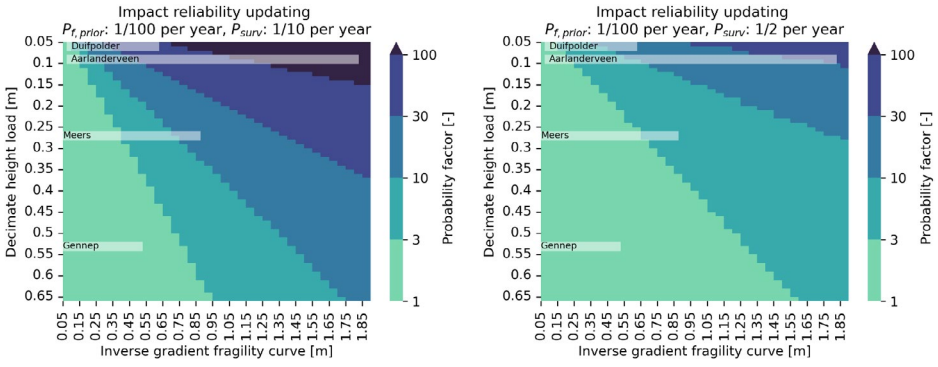


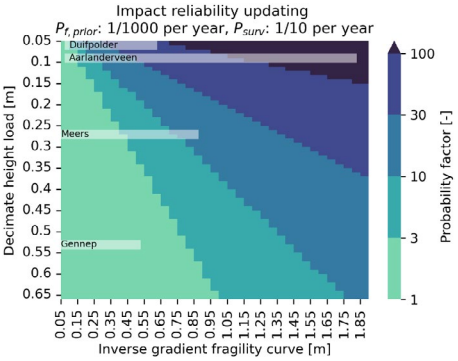
Now, consider two new dikes with similar load statistics and inverse fragility curve gradients, but with a different mean strength: the fragility curve is shifted to the right. The initial failure probabilities are now 1/1,000 per year—ten times smaller than in the first example. Again, they have survived a load level corresponding to a 10-year return period. The failure probability distributions have the same shape, but the scale of the y-axis decreased by a factor of ten. The effect of updating is similar in terms of the ratio between the prior and the updated failure probabilities.

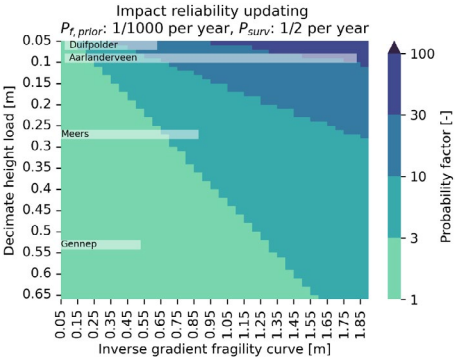


D.3 *Additional figures for conceptual analysis*

The figures below show the impact of reliability updating under different prior failure probabilities (1/100 and 1/1,000 per year) and survival load levels (levels with exceedance probabilities of 1/2 and 1/10 per year).




Impact reliability updating
$P_{f,prior}$: 1/1000 per year, P_{surv}: 1/10 per year
Duifpolder
Aarlanderveen
Meers
Gennep
Decimate height load [m]
Inverse gradient fragility curve [m]
Probability factor [-]
100
30
10
3
1


Impact reliability updating
$P_{f,prior}$: 1/1000 per year, P_{surv}: 1/2 per year
Duifpolder
Aarlanderveen
Meers
Gennep
Decimate height load [m]
Inverse gradient fragility curve [m]
Probability factor [-]
100
30
10
3
1

E Assessing the impact of climate change on extreme hydraulic head levels and dry wet cycles of Dutch canal dikes

Canal dikes in low-lying polders worldwide are critical earth structures for flood protection. Historically, inner-slope instabilities have been triggered by extreme dry and wet subsurface water conditions. Additionally, cyclic wetting and drying affect the dikes' resistance through mechanisms such as swelling and shrinkage, shear-strength reduction and soil consolidation. To accurately estimate failure probabilities of canal dikes, quantifying extreme hydraulic head levels and the magnitude of dry-wet cycles is crucial. However, there is limited understanding of changes in extreme head levels and the dry-wet cycles under climate change. This study assessed the impact of climate change on extreme hydraulic head levels and dry-wet cycles of canal dikes and how it varies across different dikes. This is done by analyzing dry-wet cycles and extreme head levels at three canal dikes in the Netherlands using both observations and models. Head observations from three monitoring sites provided data for developing time series models. By forcing 30 years of precipitation and evaporation to these models, extreme head levels and dry-wet cycles were derived. This was done under different climate scenarios to quantify the impact of climate change. It was found that the impact of climate change has a more pronounced effect on extreme low head levels than on high ones for the dikes considered. By 2100, extreme low levels could decrease by up to half a meter in drying scenarios, potentially shifting events from a 100-year frequency to every 2 years. Climate change also amplifies dry-wet cycles in canal dikes, with increases of up to several decimeters. In general, dikes with higher peak head responses and longer response times are more affected by climate change, emphasizing the importance of considering site-specific factors when evaluating canal dike stability under future climate conditions.

This appendix is based on Strijker, B. and Kok, M. (2025). Assessing the impact of climate change on extreme hydraulic head levels and dry wet cycles of Dutch canal dikes, 9th International Symposium on Geotechnical Safety and Risk, ISGSR 2025.

E.1 Introduction

Many people live in low-lying polders around the world. These places are characterized by their unique water management, as they lie below sea or river levels and are separated from the surrounding hydrological regime by dikes. Due to this separation, the rainfall that falls in the polders has to be drained out to prevent flooding. This is done using inland drainage canals. Water levels in these canals are controlled and can rise above the polder elevation due to soil subsidence, causing earthen embankments to form along the canals. People in these polders live with a continuous threat, as the water levels in these canals can be several meters higher than their houses. In the event of a dike breach, their houses can be flooded. Therefore, the performance of these dikes

and how they evolve over time is of key interest to society. Their performances are periodically assessed and if the dikes' performance does not meet the safety standard, they are reinforced. These reinforcements have to ensure reliable performance for the decades to come. But do these reinforcements account properly for the changes in strength and load that can be expected in the future?

Canal dikes often consist of organic materials, like clay and peat. These soils are continuously exposed to environmental interactions, especially to variations in atmospheric conditions. This resulted in several inner-slope instabilities over time, see for example Figure E1. Relevant load variables for dike instabilities include the pore-water pressures, or head levels, within the dike body, as both extreme high and low head levels can lead to failure. Low head levels are accompanied by weight loss in the unsaturated zone, potentially causing deformations and uplift that may trigger failure. High head levels, on the other hand, reduce effective stresses and shear strength, resulting in reduced stability. Next to extremes, cyclic wetting and drying also affect the dikes' resistance. Soils may lose shear strength due to swelling and shrinkage (Stirling et al., 2021). Effects that improve the strength can also be expected, as saturated soils consolidate over time. This process can be amplified by climate change. Lower head levels increase effective stresses, potentially reaching levels higher than ever experienced. This, in turn, contributes to undrained shear strength (Ladd & Foott, 1974). Depending on the subsoil and site-specific conditions, changing extreme head levels and dry-wet cycling can improve or worsen dike stability. However, there is little known about how the dynamics of head levels might evolve in the future. This paper dives into the dry-wet cycles and extreme head levels at three canal dikes in the Netherlands and tries to answer the research question: What is the impact of climate change on extreme hydraulic head levels and dry-wet cycles of canal dikes and how it varies across different dikes?

Figure E 1 Left: The dike breach at Terbregge, The Netherlands, in 2003. Right: the dike breach at Reeuwijk, The Netherlands, in 2021. In both cases, the canal is on the left side and the polder is on the right side. The canals were rather shallow and up to about one meter deep, limiting the development of the breach and the water that could flow into the polder.

Answering this question involves analyzing dry-wet cycles and extreme head levels at three canal dikes in the Netherlands using both observations and models. Head observations from three monitoring sites provided data for developing time series models. These models aim to explain the observed heads by other observed variables, like precipitation and evaporation. By forcing 30 years of precipitation and evaporation to these models, extreme head levels and dry-wet cycles were derived. This was done under different climate scenarios to quantify the impact of climate change.

E.2 Monitoring sites

The Netherlands has around 10,000 km of canal dikes, of which the majority of the dikes consist of soft organic soils, like peat and clay (Bezuijen et al., 2005). This study considered three dikes located at three different polders in the Western part of the Netherlands. These dikes are characterized by gentle slopes (1:10) and small head differences (2-4 meters), see Figure E2. At each dike, the phreatic head levels in the dike body are measured with five piezometers for over four years. The head level fluctuations are mainly driven by precipitation and evaporation under both daily and extreme conditions, as the fluctuations in the canal water levels are small (see bottom of Figure E2).

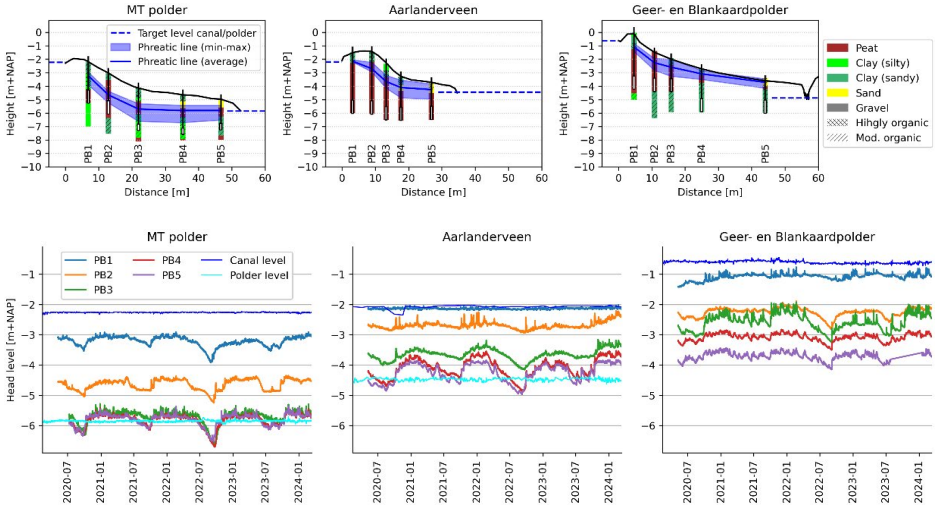


Figure E 2 Top: cross-sectional profile of the three canal dikes and the variation of the head levels. Bottom: evolution of the observed head levels and canal/polder water levels.

E.3 Methodology

Modelling head levels using time series modelling

To assess the impact of climate change, models are needed to analyze head responses in dikes under different scenarios. Numerical groundwater models are commonly used, but they struggle to reproduce observed head levels in dikes. This is due to uncertainties in subsurface conceptualization, spatial heterogeneity, and boundary conditions (Van Esch, 2012). Therefore, data-driven time series models have been used to model head dynamics (Bakker & Schaars, 2019). These models rely on multi-year head observations but do not require detailed subsurface information.

Although the physical processes are not represented, time series models, like the PIRFICT approach, can be insightful in understanding the head dynamics. This approach uses impulse response functions to represent the contributions of various drivers, such as precipitation and evaporation, to the head (Collenteur et al., 2019). Among the available model structures, the TARSO model was chosen (Knotters and Gooijer, 1999). This structure accounts for nonlinearity, with the head responding differently above or below a threshold. This nonlinearity can be found in the head observation, where head levels are capped, which can be caused by surface runoff or different soil properties at the top layer (Strijker & Kok, 2024). Furthermore, the only driver in this model is the net groundwater recharge, representing the combined effects of rainfall and evaporation. This factor is likely the most important because it explains the observed head series very well. To make the most of available data, a separate model was calibrated for each piezometer along the dike profile, using the entire set of observations. An example of one of these calibrated models is shown in Figure E3.

To understand why head levels vary as they do, two key characteristics of the response functions are useful. First, the peak of the block response (A) indicates the maximum effect of an impulse on the head level. Second, the response time (t95) reflects how long it takes for the influence of the impulse to dissipate. These characteristics vary across dikes and can help explain why climate change affects head levels differently from one dike to another.

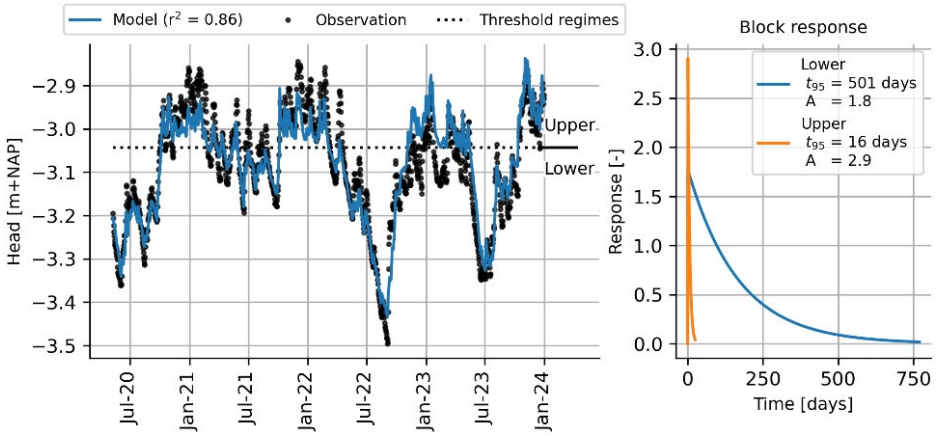


Figure E 3 A time series model that explains the observed head levels for PB4 at Geer- en Blankaardpolder, where the threshold level of the two regimes is indicated (left) with the corresponding response functions (right).

Quantifying the impact of climate change on head dynamics

Answering the research question requires head time series under different climate scenarios and time horizons. Therefore, these models were simulated using 30-year time series of precipitation and evaporation corresponding to different climate scenarios. These scenarios provide plausible and physically consistent storylines or pathways of climate change in the Netherlands, considering uncertainty for regional climate change and global temperature changes driving it. Note that these time series are a representation of the climate and not an estimate of the actual weather (Van Dorland et al., 2023). In total, nine different model simulations were used. One simulation for the current climate and eight for the future climate with combinations of different time horizons (2050 and 2100), emission scenarios (High SSP5-8.5 and Low SSP1-2.6) and regional climate response (wet and dry-trending). These simulations give a range of possible head time series that can be expected in the future, see Figure E4. This study does not account for changes in head responses caused by shifts in hydraulic conductivity and water retention capacity under the impact of climate change (Azizi et al., 2019).

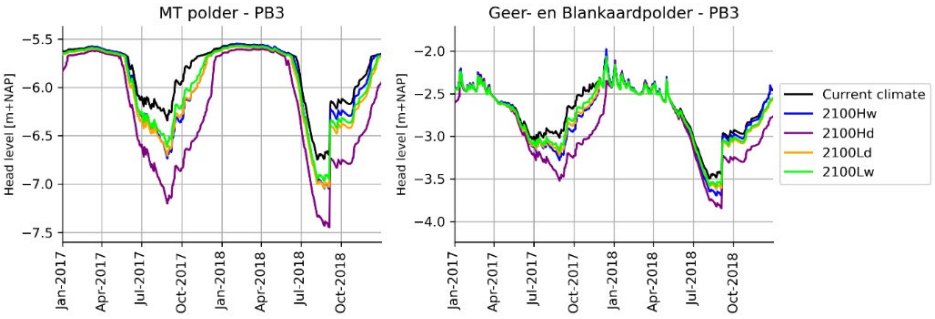


Figure E 4 Simulated head levels under different climate scenarios for two monitoring points at two different dikes

To address the research question, extreme hydraulic head levels and dry-wet cycles must be quantified. This involves analyzing extreme head values across the nine simulations. The extreme levels have two sides; highs and lows. Both were examined by separately fitting extreme value distributions (exponential distributions) to the minima and maxima. The minimum and maximum values were identified using the Peaks-Over-Threshold method, using a 90-day time window between adjacent values to ensure the independence of selected values. To accurately capture extreme values, a threshold was set to select only 15 peaks from the 30-year time series. The fitted extreme value distributions were used to estimate the head levels with a 100-year return period, aligning with required safety standards. Secondly, the extreme value distributions of the low head levels were used to calculate the size of wet-dry cycles. This cycle size is defined as the average distance between the surface level and the low head level occurring once every 10 years, representing the zone where the dry-wet cycle occurs and a decrease in shear strength is likely to occur.

E.4 Results: climate change effects on extreme heads and dry-wet cycles

The findings demonstrate the impact of climate change on canal dikes. This is done by first focusing on extreme hydraulic head levels (highs and lows), then on dry-wet cycles, and finally on the variability of both across different dike profiles.

First, the extreme head levels are discussed, where a clear pattern emerges: the impact of climate change is more pronounced for low levels than for high ones (see Tables E1 and E2). The extreme high head levels increase only slightly up to 2050, and in some cases, even decrease due to increasing drought affecting high levels during winter as well. By 2100, high head levels with a 100-year return period could rise by up to 7 cm in the Hw-scenario. Expressed in terms of changing frequencies, by 2100 and under the Hw-scenario, the extreme head level that now occurs once every 100 years could happen

every 25 years, driven by wetter winters. For extreme low head levels, a drop of 8 to 23 cm is expected by 2050, with further decreases of up to nearly 50 cm possible by 2100. Extreme low head levels become far more frequent, as also expected by decreasing precipitation and increasing evaporation during summers. In 2050, head levels that currently occur once every 100 years might occur every 15-50 years and by 2100 could occur on average every 2 years, in the Hd-scenario.

Table E.1 Increase in the T=100 high head level for different climate scenarios (High/Low emissions; dry/wet regional climate response) in 2050 and 2100. Dash sign indicates no change.

	2050				2100			
	Hd	Hw	Ld	Lw	Hd	Hw	Ld	Lw
MT-polder	-1 cm	-	-1 cm	-1 cm	-1 cm	+3 cm	-1 cm	-1 cm
Aarlanderveen	+1 cm	+1 cm	+1 cm	+1 cm	+7 cm	+5 cm	+1 cm	+1 cm
Geer- en Blankaardpolder	-1 cm	+1 cm	-1 cm	-	-	+7 cm	-1 cm	-

Table E.2 The change in the T=100 low head level for different climate scenarios (High/Low emissions; dry/wet regional climate response) in 2050 and 2100.

	2050				2100			
	Hd	Hw	Ld	Lw	Hd	Hw	Ld	Lw
MT-polder	-23cm	-13 cm	-18 cm	-14 cm	-48 cm	-26 cm	-18 cm	-14 cm
Aarlanderveen	-8 cm	-8 cm	-10 cm	-8 cm	-38 cm	-15 cm	-10 cm	-8 cm
Geer- en Blankaardpolder	-10 cm	-10 cm	-12 cm	-10 cm	-39 cm	-20 cm	-12 cm	-10 cm

Next, the impact of climate change on the dry-wet cycles is analysed. Under the current climate, these cycles already vary in size across the three dikes due to differences in horizontal groundwater flow, influenced by canal water intrusion and drainage capacities. At the MT polder, the cycles are the largest, reaching about 2 meters—twice the size of those at the Geer- and Blankaardpolder, see Figure E5. While the cycle sizes differ across locations, climate change causes a similar absolute increase in cycle size at all sites. By 2100, average cycles, based on different climate scenarios, could rise by 15 cm compared to current levels, with a maximum of up to 40 cm in the Hd-scenario at the MT polder. At Geer- and Blankaardpolder, cycle sizes could increase by just over 30 cm.

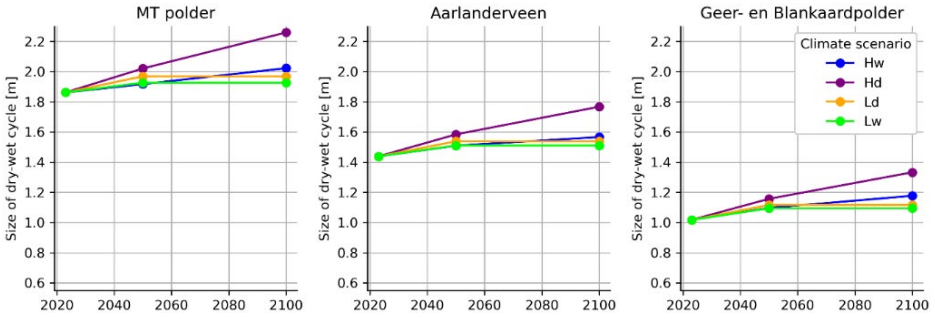


Figure E 5 Evolution of the dry-wet cycles under different climate scenarios.

Lastly, differences between dikes are explained by considering the variations in response functions. For extreme high head levels, climate change has a greater impact when upper-regime head levels respond strongly to precipitation. This is indicated by larger peak block responses and slower drainage, shown by longer 95% response times. Similarly, for extreme low head levels, higher peak block responses and longer response times in the lower regime amplify the impact.

E.5 Qualitative impacts and implications for dike safety

As shown, climate change is expected to result in more extreme high and low head levels, along with enlarged dry-wet cycles. How does this impact dike stability? This depends on the failure mechanism and site-specific characteristics being considered. Let's consider the three canal dikes and instabilities triggered by extreme high pore pressures, as analyzed in the safety assessments of Dutch canal dikes. A first indication can be drawn by examining changes in effective stresses and shear strength along critical or representative slip planes. For the three canal dikes studied, the representative slip planes are mainly seated below the phreatic line with lengths from 15 to 60 meters. The total shear resistance is modest due to the presence of soft, lightweight soils. These light soils, in combination with gentle slopes, also keep the driving moments relatively low. The first effect considered is more extreme head levels, which lead to a decrease in effective stresses along the slip planes and, consequently, a reduction in stability. Secondly, the enlarged dry-wet cycle has two consequences. On one hand, it can induce deterioration processes, which decreases the shear strength of soils in the unsaturated zone. On the other hand, it can cause consolidation, which increases the shear strength of soils in the saturated zone. Whether the latter occurs, depends on the stress history of the soils, which is unknown for the considered locations. The impact of the shear strength reduction of the unsaturated zone, depends on the initial contribution of the shear strength from this zone. This initial contribution is small for the considered slip planes that run largely through the saturated zone where the majority of the resistance comes from. If the contribution of the unsaturated shear

strength is larger, for example for shallower and smaller slip planes, the impact of soil deterioration can be expected to be higher. However, for dike safety, these slip planes are irrelevant, since they are unlikely to cause dike breaches. To summarize, while more extreme head levels and deterioration in the unsaturated zone can slightly reduce the dike stability, consolidation has the potential to counteract these effects.

E.6 Conclusions

This study examined the impact of climate change on 1) extreme hydraulic head levels, 2) the size of dry-wet cycles and 3) the variation of these effects across different canal dikes. Firstly, the findings indicate that climate change has a more pronounced effect on extreme low head levels than on high ones. By 2050, changes in extreme high heads will be modest, with some scenarios even showing decreases due to increasing droughts. Next to limited changes in extreme high head levels, extreme low levels may drop up to several decimeters, with highest drops in dry-trending scenarios. By 2100, the effect on extreme high heads is still moderate, while extreme low levels may decrease further: up to half a meter compared to the current climate in scenarios with a drying trend. Therefore, low levels are expected to occur far more frequently, with some scenarios indicating a shift from a 100-year event to occurrences as frequent as every 2 years.

Secondly, climate change amplifies the size of dry-wet cycles in canal dikes affecting the resistance in both positive and negative ways. These cycles vary in size across the studied dikes under the current climate. The MT polder exhibits the largest cycles, reaching about 2 meters. This is twice the size of the cycles observed at the Geer- and Blankaardpolder. Despite these site-specific differences, climate change leads to a similar absolute increase in cycle sizes across all locations of about 15-20 cm by 2100, with maximum increases of up to 40 cm for the high emission scenario and dry-trending regional climate response.

Thirdly, the variation in how climate change impacts head dynamics across different dikes was analyzed. Dikes with head responses characterized by higher peak block responses and longer response times were found to be generally more susceptible to the effects of climate change.

The impact of changing head dynamics on stability was qualitatively assessed. It revealed that while more extreme head levels and deterioration in the unsaturated zone may slightly weaken dike stability, consolidation has the potential to mitigate these effects. These findings emphasize the need to consider site-specific characteristics when assessing the stability of canal dikes under future climate scenarios.

List of publications

Journal papers

- Strijker, B., Jonkman, S.N., Kok, M. (2026). Estimating spatial dependencies of peak heads in Dutch canal dikes: a national-scale hindcasting analysis. *Journal of Coastal and Riverine Flood Risk*. (submitted)
- Strijker, B., Jonkman, S. N., & Kok, M. (2026). The role of load variations in assessing credible dike failure probabilities: balancing load and strength uncertainties. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 1-22.
- Strijker, B., & Kok, M. (2025). The dynamics of peak head responses at Dutch canal dikes and the impact of climate change. Natural Hazards and Earth System Sciences, 25(9), 3355-3379.
- Strijker, B., Heimovaara, T. J., Jonkman, S. N., & Kok, M. (2024). Exploring subsurface water conditions in Dutch canal dikes during drought periods: Insights from multiyear monitoring. *Water Resources Research*, *60*(9), e2023WR036046.
- Strijker, B., Asselman, N., de Jong, J., & Barneveld, H. (2023). The 2021 floods in the Netherlands from a river engineering perspective. *Journal of Coastal and Riverine Flood Risk*, *2*.

Conference contributions

- Strijker, B., & Kok, M. (2025). Assessing the Impact of Climate Change on Extreme Hydraulic Head Levels and Dry-Wet Cycles Of Dutch Canal Dikes. In *9th International Symposium for Geotechnical Safety and Risk (ISGSR). 25—28 August 2025, Oslo, Norway*
- Strijker, B., & Kok, M. (2023). The joint impact of rainfall events on water-and dike systems in Dutch polders. In *EGU General Assembly Conference Abstracts* (pp. EGU-9048).
- Strijkera, B., & Kok, M. (2022). Vulnerability of villages in the Dutch Geul valley. In *NCR DAYS 2022: Anthropogenic Rivers* (pp. 41-42).

Selection of other projects

- Strijker, B. (2025b). *Approach for schematizing the phreatic line for the primary dike around Tholen* (in Dutch). In collaboration with Waterschap Scheldestromen. December 2025.
- Strijker, B. (2025a). *Impact of the wet winter 2023–2024 on canal dikes: Lessons for dike safety—observed damages and phreatic lines* (in Dutch). PR5177.10. January 2025. ISBN: 978.94.6479.111.2
- Strijker, B., & de Snoo, I. (2024a). *Geohydrology of canal dikes at Rijnland: Extreme phreatic lines in canal dikes and changing climate conditions* (in Dutch). PR5181.10. August 2024.

- Strijker, B. (2024b). *Pore water pressures in the Markermeer Dikes during the high water and wet winter of 2023* (Presentation in Dutch). In collaboration with Hoogheemraadschap Hollands Noorderkwartier. July 2024. (Section 6.4)
- Strijker, B. (2023a). *Time series analysis of canal dikes: Time series analysis of phreatic lines in canal dikes in the management area of HHSK* (in Dutch). On behalf of Hoogheemraadschap van Schieland en de Krimpenerwaard. PR4877.10. June 2023.
- Strijker, B. (2023b). *A new groundwater drought indicator for polder dikes: A statistical measure representing the extremity of groundwater drought in the* (in Dutch). On behalf of STOWA, Rijkswaterstaat, Hoogheemraadschap Hollands Noorderkwartier, and Hoogheemraadschap van Schieland en de Krimpenerwaard. PR5039.10. February 2023.