



# Vergelijking van twee modellen voor het bepalen van economisch optimale overstromingskansen

Eindrapport derdejaars stage Toegepaste Wiskunde binnen de afdeling Water, Verkeer en Leefomgeving van Rijkswaterstaat

29 juli 2022, Delft

E.W.I. van der Klis  
19077211

Begeleiders:  
André Wooning  
Robert Slomp  
Joyce Macduff



Dit eindrapport is op 28 juli 2022 goedgekeurd door



Handwritten signature of Robert Slomp, consisting of the letters 'R', 'S', and 'S' in a stylized, cursive font.

28 Juli 2022, Delft

Robert Slomp, senior adviseur bij Rijkswaterstaat

## Voorwoord

Dit document is het stageverslag van mijn derdejaars stage Toegepaste Wiskunde bij Rijkswaterstaat. In deze stage zijn twee modellen voor het bepalen van economisch optimale overstromingskansen vergeleken. Hiervoor zijn de voor- en nadelen van de modellen op een rijtje gezet aan de hand van de uitgevoerde testen. Ook is een model aangevuld, zodat dit meer dezelfde functionaliteiten heeft als het andere model. Door de modellen zoveel mogelijk gelijk te maken, kan de vergelijking worden gemaakt. Tijdens de stage is er gewerkt aan de opleidingscompetenties onderzoeken, analyseren, professionaliseren, beheren en adviseren. De stageperiode liep van 7 februari 2022 tot en met 29 juli 2022.

Indien u geïnteresseerd bent in de resultaten van de testen kunt u deze bekijken in hoofdstuk 6. De conclusie en aanbevelingen die volgen uit de resultaten zijn te vinden in hoofdstuk 7.

Mijn stageperiode was niet altijd gemakkelijk, zo was ik vaak erg onzeker over mijn werk. Graag wil ik mijn stagebegeleider André Wooning bedanken voor zijn bruikbare adviezen, positiviteit en de motiverende gesprekken die we gevoerd hebben. Daarnaast wil ik docent Joyce Macduff bedanken voor de opbouwende feedback. Verder wil ik de betrokkenen bij het project bedanken, namelijk Robert Slomp en Jarl Kind voor hun medewerking aan de interviews en collegialiteit. Tot slot wil ik de medestagiairs Marlou Wijnans en Jolie van Hummel bedanken voor hun collegialiteit. Het was een plezier om met hen samen te werken.

Ik wens u veel leesplezier toe.

Esmeé van der Klis

Delft, juli 2022

## Samenvatting

Rijkswaterstaat verwacht vanaf 2023 nieuwe berekeningen te gaan maken rondom de economisch optimale overstromingskansen omdat er sinds de laatste berekening in 2011 veel factoren veranderd zijn. Daarvoor moet eerst de kennis over de modellen binnen Rijkswaterstaat worden opgefrist. Zo wordt in deze stage het model OptimaliseRing Discreet uitgebreid, zodat dit ook de economisch optimale overstromingskansen kan bepalen. Ook worden de modellen OptimaliseRing en OptimaliseRing Discreet vergeleken. Dit gebeurt aan de hand de resultaten van de testen aan kwaliteitscriteria, waaruit de voor- en nadelen van de modellen volgen.

Rijkswaterstaat wil weten welk model te gebruiken voor toekomstige berekeningen van de economisch optimale overstromingskansen. Om dit te onderzoeken is er een hoofdvraag opgesteld. De hoofdvraag luidt als volgt: *Wat zijn de voor- en nadelen van het gebruik van OptimaliseRing en OptimaliseRing Discreet voor de toekomstige berekening van de economisch optimale overstromingskansen?* Deze hoofdvraag is beantwoord door middel van meerdere deelvragen. Zo is er literatuuronderzoek gedaan naar de modellen en naar de statistische methoden, zijn de kwaliteitscriteria opgesteld met bijbehorende testen, is het model OptimaliseRing Discreet aangevuld, is er een statistische analyse uitgevoerd en zijn de testen uitgevoerd. Dit leverde uiteindelijk voor- en nadelen voor beide modellen op.

Het vergelijken van de modellen OptimaliseRing en OptimaliseRing Discreet gebeurt aan de hand van kwaliteitscriteria. De kwaliteitscriteria zijn opgesteld door interviews af te nemen met experts op het gebied van overstromingskansen en/of de modellen OptimaliseRing en OptimaliseRing Discreet. De kwaliteitscriteria zijn gecategoriseerd in de volgende categorieën: juistheid, verklaarbaarheid, stabiliteit, aanpasbaarheid, gebruiksvriendelijkheid, reproduceerbaarheid en rekentijd. Aan de kwaliteitscriteria zijn testen gekoppeld en deze testen zijn uitgevoerd. Een onderdeel van deze testen is de statistische analyse. Uit de statistische analyse volgt dat er een statistisch significant verschil in de economisch optimale overstromingskansen is tussen OptimaliseRing en OptimaliseRing Discreet,  $p < 0,001$ . Met behulp van meervoudige robuuste regressie is er een relatie gevonden tussen het verschil in overstromingskans en  $\psi$  en tussen het verschil in overstromingskans en  $\zeta$ . Dit zijn twee belangrijke parameters ten aanzien van klimaatverandering en potentiële schade.

Uit de resultaten van de testen volgen de voor- en nadelen voor beide modellen. De voordelen van OptimaliseRing zijn: de uitkomsten van het model zijn correct, er is uitgebreide informatie beschikbaar over het model, het programma heeft een gebruikersschil en heeft een kortere rekentijd dan OptimaliseRing Discreet. Ook zijn er wat voordelen in de code achter OptimaliseRing. De nadelen van OptimaliseRing zijn dat er vrijwel geen commentaar in de code staat en het programma gebouwd is in een taal waar een licentie voor nodig is. De voordelen van OptimaliseRing Discreet zijn: het model biedt flexibiliteit met de invoer data en het model is gebouwd in een taal die open source is. Ook zijn er wat voordelen in de code. De nadelen van OptimaliseRing Discreet zijn: de uitkomsten wijken sterk af van OptimaliseRing, er zit een systematische fout in de uitkomsten, er is weinig documentatie beschikbaar, het programma heeft geen gebruikersschil en het programma heeft een langere rekentijd dan OptimaliseRing. Ook zijn er wat nadelen in de code.

## Inhoudsopgave

Voorwoord.....	2
Samenvatting .....	3
H1: Inleiding.....	6
H1.1: Aanleiding en achtergronden.....	6
H1.2: Doelstelling.....	6
H1.3: Hoofdvraag en deelvragen .....	7
H1.4: Methode .....	7
H1.5: Afbakening .....	7
H1.6: Leeswijzer .....	8
H2: Literatuuronderzoek .....	9
H2.1: Geschiedenis.....	9
H2.2: Optimaliseren voor dijkversterking.....	10
H2.3: OptimaliseRing .....	11
H2.4: OptimaliseRing Discreet .....	16
H2.5: Statistische methoden .....	21
H3: Testen aan kwaliteitscriteria .....	25
H3.1: Juistheid .....	25
H3.2: Verklaarbaarheid .....	25
H3.3: Stabiliteit.....	26
H3.4: Aanpasbaarheid .....	26
H3.5: Gebruiksvriendelijkheid .....	26
H3.6: Reproduceerbaarheid.....	27
H3.7: Rekeningtijd .....	27
H3.8: Kosten .....	27
H4: Bepalen economisch optimale overstromingskansen .....	28
H4.1: Input data genereren .....	28
H4.2: Berekening implementeren.....	29
H4.3: Resultaten OptimaliseRing Discreet.....	29
H5: Statistische analyse .....	31
H5.1: Voorbewerking data.....	31
H5.2: Gemiddelde procentuele verschil .....	31
H5.3: Verschiltoetsen .....	32
H5.4: Regressie .....	34

H6: Resultaten testen aan kwaliteitscriteria .....	37
H6.1: Juistheid .....	37
H6.2: Verklaarbaarheid .....	38
H6.3: Stabiliteit .....	38
H6.4: Aanpasbaarheid .....	38
H6.5: Gebruiksvriendelijkheid .....	39
H6.6: Reproduceerbaarheid .....	41
H6.7: Rekentijd .....	41
H6.8: Kosten .....	42
H7: Conclusie, aanbevelingen en discussie .....	43
H7.1: Conclusie .....	43
H7.2: Aanbevelingen .....	44
H7.3: Discussie .....	45
Literatuurlijst .....	46
Bijlagen .....	48
Bijlage 1: Schematische weergave code .....	48
Bijlage 2: Dictionary's .....	50
Bijlage 3: Resultaten OptimaliseRing Discreet .....	51
Bijlage 4: Validatie regressiemodel .....	54

# H1: Inleiding

## H1.1: Aanleiding en achtergronden

Rijkswaterstaat beheert en ontwikkelt de rijkswegen, -vaarwegen en -wateren en zet in op een duurzame leefomgeving. Dit onderzoek is uitgevoerd binnen de dienst water, verkeer en leefomgeving en daarbinnen de afdeling waterkeringen. Deze afdeling focust zich op het toetsen, versterken en in stand houden van waterkeringen. Het doel van het onderzoek is om twee modellen te vergelijken voor het bepalen van economisch optimale overstromingskansen.

Er zijn meerdere modellen om economisch optimale overstromingskansen te bepalen. Deze modellen zijn zowel wiskundig als qua software anders opgebouwd. Onlangs is OptimaliseRing Discreet ontwikkeld, dit model is in open source software geschreven. In deze stage worden de modellen OptimaliseRing en OptimaliseRing Discreet vergeleken.

De optimale overstromingskansen zijn berekend in 2011 met het model OptimaliseRing en sindsdien zijn er verschillende factoren veranderd. Zo is de economie gegroeid en zijn er andere prognoses voor het jaar 2050, voor de economie en voor het klimaat. Ook zijn de kosten van dijkversterkingen veranderd en is de discontovoet<sup>1</sup> aangepast. Daarom is het belangrijk om opnieuw de economisch optimale overstromingskansen te berekenen. Rijkswaterstaat verwacht vanaf 2023 dat er nieuwe berekeningen gemaakt gaan worden rondom de economisch optimale overstromingskansen. Daarvoor moet eerst de kennis over de modellen binnen Rijkswaterstaat worden opgefrist. De bevindingen uit 2011 zijn vastgelegd in het rapport Maatschappelijke kosten-batenanalyse Waterveiligheid 21e eeuw (Kind, 2011). Het model OptimaliseRing Discreet moet worden uitgebreid, zodat deze ook de economisch optimale overstromingskansen kan berekenen. Verder is het belangrijk om de twee modellen te vergelijken en te bepalen aan de hand van testen aan de kwaliteitscriteria wat de voor- en nadelen zijn van het gebruik van de modellen voor toekomstige berekeningen.

## H1.2: Doelstelling

Het model OptimaliseRing Discreet wordt uitgebreid en aangepast zodat dit programma kan draaien en de berekening voor de economisch optimale overstromingskansen wordt geïmplementeerd in dit model. De economisch optimale overstromingskansen worden in deze stage bepaald met het model OptimaliseRing Discreet op basis van dezelfde data als gebruikt voor de bepaling van de economisch optimale overstromingskansen met OptimaliseRing. Rijkswaterstaat kan mede op basis van dit rapport een onderbouwde keuze maken over welk model te gebruiken voor de toekomstige berekeningen van de economisch optimale overstromingskansen op basis van de gegeven voor- en nadelen.

---

<sup>1</sup> De discontovoet is een jaarlijks percentage waarmee verwachte kosten en baten in de toekomst worden teruggerekend naar het basisjaar van het project (Ministerie van Infrastructuur en Waterstaat, 2022).

### H1.3: Hoofdvraag en deelvragen

Om de doelstelling te bereiken, wordt de opdracht uitgevoerd aan de hand van onderzoeksvragen. Naar aanleiding van de hoofdvraag zijn er een aantal deelvragen opgesteld om de hoofdvraag te kunnen beantwoorden.

Hoofdvraag:

Wat zijn de voor- en nadelen van het gebruik van OptimaliseRing en OptimaliseRing Discreet voor de toekomstige berekening van de economisch optimale overstromingskansen?

Deelvragen:

1. Hoe zijn de modellen OptimaliseRing en OptimaliseRing Discreet opgebouwd en welke wiskunde wordt in deze modellen toegepast?
2. Wat zijn de kwaliteitscriteria voor de modellen OptimaliseRing en OptimaliseRing Discreet?
3. Wat zijn de testmogelijkheden om de modellen OptimaliseRing en OptimaliseRing Discreet aan de kwaliteitscriteria te toetsen?
4. Wat zijn de economisch optimale overstromingskansen gebruik makend van het model OptimaliseRing Discreet?
5. Wat is een statistisch onderbouwde verklaring voor de verschillen in de uitkomsten tussen de modellen OptimaliseRing en OptimaliseRing Discreet?
6. In welke mate voldoen de modellen OptimaliseRing en OptimaliseRing Discreet aan de kwaliteitscriteria?

### H1.4: Methode

De eerste deelvraag is beantwoord met behulp van literatuuronderzoek. Door interviews af te nemen met collega's met inhoudelijke kennis over dit onderwerp zijn er kwaliteitscriteria opgesteld voor de modellen OptimaliseRing en OptimaliseRing Discreet en hiermee is de tweede en derde deelvraag beantwoord. Het model OptimaliseRing Discreet is aangevuld, zodat hier de bestaande data mee kan worden ingelezen en er voor elk dijkkringdeel voor elk jaar een overstromingskans kan worden berekend. Hiermee is de vierde deelvraag beantwoord. Er zijn verschiltoetsen uitgevoerd op de uitkomsten van de modellen OptimaliseRing en OptimaliseRing Discreet om erachter te komen of er een statistisch significant verschil is. Vervolgens zijn de verschillen in uitkomsten verklaard met behulp van robuuste regressie en daarmee is de vijfde deelvraag beantwoord. Als laatste zijn alle testen uitgevoerd op de modellen OptimaliseRing en OptimaliseRing Discreet en zijn de resultaten besproken met collega's om hierover een conclusie te kunnen trekken, waarmee de zesde deelvraag beantwoord is. Uiteindelijk zijn alle bevindingen gerapporteerd in dit rapport en is er een advies uitgebracht aan Rijkswaterstaat.

### H1.5: Afbakening

De stage is gestart op 7 februari 2022 en eindigde op 29 juli 2022. In de stage zijn de economische optimale overstromingskansen bepaald met het model OptimaliseRing Discreet. Dit is gedaan voor dezelfde dijkkringdelen waarvan eerder de economisch optimale overstromingskansen zijn bepaald met OptimaliseRing. Hiervoor is de input data gegenereerd en de berekening van de economisch optimale overstromingskansen geïmplementeerd in OptimaliseRing Discreet.



De verklaring van de verschillen in uitkomsten tussen de modellen zal gebeuren met de economisch optimale overstromingskansen die al bepaald zijn met OptimaliseRing en de economisch optimale overstromingskansen die zijn bepaald met OptimaliseRing Discreet. Wat buiten dit onderzoek valt, is onderzoek naar mogelijke andere modellen voor het bepalen van de economisch optimale overstromingskansen. Er is binnen dit onderzoek alleen gekeken naar de modellen OptimaliseRing en OptimaliseRing Discreet. De opdrachtgever heeft de toegang geregeld tot de modellen en benodigde software. Ook heeft de opdrachtgever de toegang tot rekencomputers geregeld, die deze modellen relatief snel kunnen draaien. Verder waren er mensen bij het project betrokken met inhoudelijke kennis over de modellen, deze mensen waren ook beschikbaar voor vragen.

#### H1.6: Leeswijzer

In hoofdstuk 2 wordt het literatuuronderzoek toegelicht. In hoofdstuk 3 worden de opgestelde kwaliteitscriteria en bijbehorende testen uitgelegd. Hoe het model is aangevuld om de economisch optimale overstromingskansen te berekenen is beschreven in hoofdstuk 4. In hoofdstuk 5 wordt de statistische analyse toegelicht. De resultaten van de testen worden besproken in hoofdstuk 6. Tot slot worden in hoofdstuk 7 de conclusie, aanbevelingen en discussie gegeven.

## H2: Literatuuronderzoek

Een beknopte geschiedenis over waterveiligheid in Nederland en het ontstaan van de modellen Optimalisering en Optimalisering Discreet is beschreven in hoofdstuk 2.1. In hoofdstuk 2.2 wordt het algemene principe van optimaliseren voor dijkversterking uitgelegd. In hoofdstuk 2.3 wordt het model OptimaliseRing toegelicht en in hoofdstuk 2.4 het model OptimaliseRing Discreet. Tot slot worden in hoofdstuk 2.5 de gebruikte statistische methoden uitgelegd.

De modellen OptimaliseRing en OptimaliseRing Discreet zijn ontwikkeld door Deltares en CPB in opdracht van het ministerie van Infrastructuur en Waterstaat. Daarom beschikt Rijkswaterstaat over de nodige literatuur over de modellen OptimaliseRing en OptimaliseRing Discreet. Bijvoorbeeld over een technische documentatie en gebruikershandleiding. Deze aangeleverde literatuur is gebruikt in dit rapport om achtergrondinformatie te vergaren over de modellen. Het gaat hier ook regelmatig om niet-openbaar-gepubliceerde literatuur. Om kennis op te doen over de modellen is deze literatuur doorgelezen en zijn relevante stukken verwerkt in dit hoofdstuk.

### H2.1: Geschiedenis

In 1953 vond in Nederland de watersnoodramp plaats, wat duidelijk maakte dat Nederland onvoldoende beschermd was tegen het water. Ten gevolge van deze gebeurtenis werd de Deltacommissie opgericht. De Deltacommissie had als taak om het gewenste beschermingsniveau voor de dreigingen van het water vanuit de zee vast te stellen. Rijkswaterstaat bouwde als antwoord op de watersnoodramp de Deltawerken, om Nederland te beschermen tegen hoogwater vanuit de Noordzee (Ministerie van Infrastructuur en Waterstaat, z.d.). Prof. D. van Dantzig heeft voor deze commissie een besliskundig model gemaakt voor Centraal Holland. Hierin zijn de kosten voor het versterken van waterkeringen (investeringskosten) afgewogen tegen de vermindering van de verwachte schade (baten). Hierbij is vastgesteld bij welke dijkhoogte de totale kosten (dus de som van investeringskosten en de verwachte schade) minimaal zijn en wat het daarbij behorende optimale veiligheidsniveau is (Duits, Kind & Bak, 2011, p. 3)

Vooruitlopend op het programma "Ruimte voor de Rivier", wat in 2006 gestart is, is door het CPB een kosten-batenanalyse uitgevoerd. In het eerste deel zijn de economisch optimale overstromingskansen in het rivierengebied van de Rijn berekend. Hierbij is het model van Dantzig als startpunt genomen. Eijgenraam heeft dit model verbeterd, waardoor er rekening kan worden gehouden met veranderende omstandigheden in de tijd, zoals klimaatverandering en bodemdaling. In het model van Eijgenraam zijn dijkkringdelen geschematiseerd als één uniform traject, waardoor het model analytisch oplosbaar is (Duits et al., 2011, pp. 3-4).

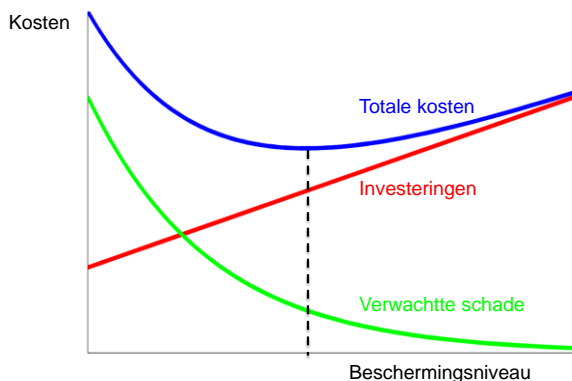
Het model van Eijgenraam is uitgebreid ten behoeve van de maatschappelijke kosten-batenanalyse waterveiligheid 21<sup>e</sup> eeuw (Kind, 2011). Met de uitbreiding konden meerdere trajecten binnen een dijkkring of dijkkringdeel worden onderscheiden. Door deze uitbreiding is het model niet meer analytisch op te lossen, maar wel met een numerieke oplossingsmethode. Daarom is het model OptimaliseRing gebouwd. Het minimalisatie-algoritme van OptimaliseRing is in AIMMS geïmplementeerd. OptimaliseRing bepaald de optimale investeringsstrategie. Dat wil zeggen dat het

model de optimale tijdstippen en verhogingen van de dijktrajecten bepaalt. Ook berekent het model per dijkkringdeel de economisch optimale overstromingskans.

OptimaliseRing Discreet bepaalt ook de optimale investeringsstrategie zoals OptimaliseRing, maar gebruikt discrete wiskunde in plaats van numerieke wiskunde. Het model is omgezet van GAMS naar Python. Hier is voor gekozen, omdat Python open source is. Dit heeft zowel voordelen qua kosten als voor toekomstig beheer.

## H2.2: Optimaliseren voor dijkversterking

Het principe van het optimaliseren voor dijkversterking komt uit de economische analyse die de eerste Deltacommissie liet uitvoeren. In figuur 1 is dit principe schematisch weergegeven. Voor het versterken en onderhouden van dijken zijn investeringskosten nodig. Hoe hoger het gewenste beschermingsniveau van een dijk, hoe hoger deze investeringskosten zijn. Daarom is dit een stijgende lijn. Tegelijkertijd resulteert een hoger beschermingsniveau tot een afname in verwachte overstromingsschade. Deze neemt asymptotisch af, omdat bij hogere dijkhoogtes het risico steeds minder afneemt. In de verwachte overstromingsschade zit de materiële schade, maar ook de schade voor getroffen en dodelijke slachtoffers. De totale kosten voor dijkversterking bestaan dus uit de investeringskosten opgeteld bij de kosten voor de verwachte schade. Er wordt net zo lang geïnvesteerd in een dijk, totdat de investeringskosten niet meer opwegen tegen de afname in de verwachte schade. In de kosten-baten analyse wordt vastgesteld wat het beschermingsniveau van een dijk is waarvoor de totale kosten minimaal zijn.

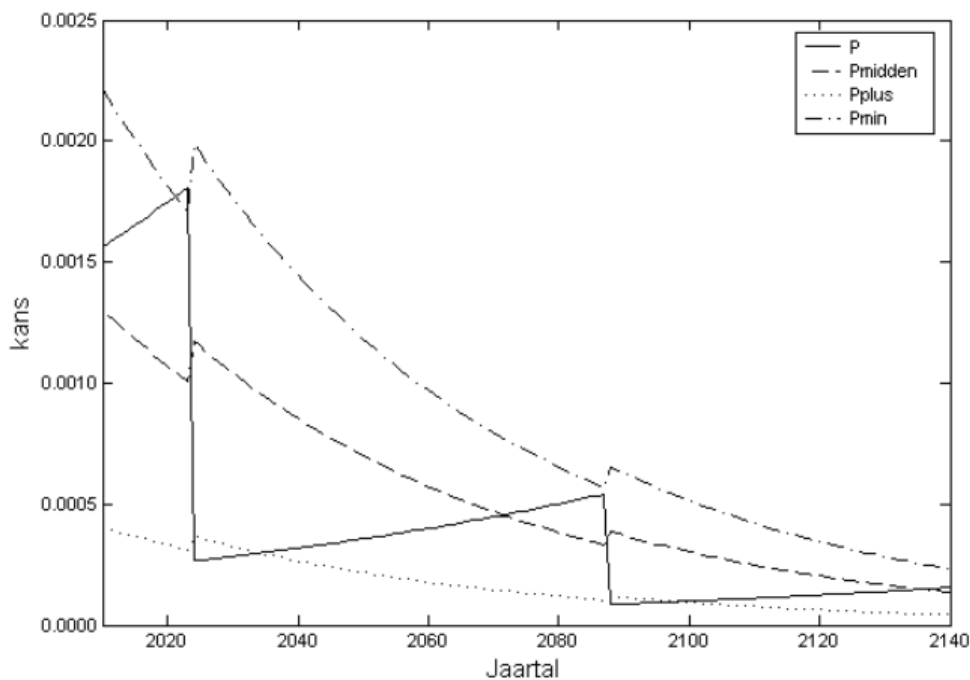


Figuur 1: Schematische weergave van de wisselwerking tussen beschermingsniveau enerzijds en kosten voor investeringen en verwachte overstromingsschade anderzijds. De stippellijn geeft aan bij welk beschermingsniveau de totale kosten minimaal zijn. Overgenomen uit (P. de Grave, persoonlijke communicatie, 14 maart 2022).

De investeringskosten bestaan uit een deel vaste kosten en een deel variabele kosten. De vaste kosten zijn af te lezen in figuur 1, namelijk de kosten op het punt waar de lijn investeringen de y-as snijdt. Uit figuur 1 is af te lezen wat het optimale beschermingsniveau is en welke kosten hierbij horen. In de praktijk is dit lastiger, omdat er een samenhang is van meerdere waterkeringen die mogelijk deels hetzelfde gebied beschermen en/of elkaars faalkans kunnen beïnvloeden. Bovendien is er een groot aantal mogelijkheden om op verschillende tijdstippen te versterken. In de praktijk is het aantal keuzes enorm, waardoor het bepalen van een optimum complex is. OptimaliseRing en OptimaliseRing Discreet zijn twee modellen die gebruik maken van rekensoftware om dit probleem op te lossen.

## De middenkans

In OptimaliseRing en OptimaliseRing Discreet is de economisch optimale overstromingskans de middenkans  $P_{midden}$  (zie figuur 2). Er wordt van de middenkans uitgegaan om het moment om actie te gaan ondernemen met betrekking tot investeringen te bepalen. De middenkans ligt tussen  $P_{min}$  en  $P_{plus}$ .  $P_{min}$  is de maximaal toelaatbare overstromingskans die hoort bij het moment vlak voor de investering.  $P_{plus}$  is de ontwerpnorm die hoort bij het moment vlak na de investering. Er is voor gekozen om  $P_{midden}$  in de wet te zetten, omdat deze kans tussen  $P_{min}$  en  $P_{plus}$  in ligt. Als  $P_{min}$  zou gelden als norm, dan zou de investering te laat plaatsvinden. Verder is  $P_{plus}$  de ontwerpnorm en is  $P$  een zaagtandpatroon, wat niet praktisch is voor in de wet. In de loop van de tijd wordt de economisch optimale overstromingskans kleiner, als gevolg van economische groei. In figuur 2 is  $P$ ,  $P_{midden}$ ,  $P_{plus}$  en  $P_{min}$  weergegeven. De sprongen omhoog in figuur 2 zijn te verklaren doordat de kosten na een eerste investering hoger kunnen zijn. Als de investering heeft plaatsgevonden kan er een hogere schade ontstaan, daardoor kunnen er ook sprongen omlaag voorkomen (Duits et al., 2011, pp. 7-8).



Figuur 2:  $P$ ,  $P_{midden}$ ,  $P_{plus}$  en  $P_{min}$ . Overgenomen uit Maatschappelijke kosten-batenanalyse Waterveiligheid 21e eeuw Bijlage C: Wiskundige achtergrond MKBA model van Deltares, 2011 ([Maatschappelijke kosten-batenanalyse Waterveiligheid 21e eeuw - Rijkswaterstaat Rapportendatabank \(overheid.nl\)](#))

## H2.3: OptimaliseRing

OptimaliseRing is het model dat gebruikt is in de maatschappelijke kosten-batenanalyse waterveiligheid 21<sup>e</sup> eeuw om de optimale investeringsstrategie en economisch optimale overstromingskansen te bepalen.

### Wiskundige model

Het doel van het model is om de optimale investeringsstrategie te kiezen. Met andere woorden, de optimale momenten en verhogingen om te investeren worden berekend, zodat de totale verdisconteerde kosten  $K$  minimaal zijn. Hierbij spelen contante

waarden van alle toekomstige kosten van schade  $S$  en investeringen  $I$  in een dijkkring een centrale rol (Duits et al., 2011, p. 9). De contante waarde is de waarde op dit moment van een bedrag dat in de toekomst beschikbaar komt (Kind, 2011, p. 113). De contante waarde van een som geld in de toekomst is gelijk aan het aantal euro's op dit moment dat dezelfde welvaart oplevert. Het maakt voor de welvaart niet uit of die som geld in de toekomst verkregen wordt of de contante waarde ervan op dit moment. De discontovoet is het jaarlijkse percentage dat toekomstige bedragen terugreken naar de contante waarde.

Formule totale verdisconteerde kosten (Duits et al., 2011, p. 9):

$$K = \int_0^z \frac{S(t)}{(1 + \delta_1)^t} dt + \frac{S(z)}{(1 + \delta_1)^z} \frac{1}{\ln(1 + \delta_1)} + \sum_j \sum_i \frac{I_{ij}}{(1 + \delta_2)^{T_{ij}}} \quad (1)$$

Waarin:

$S(t)$  = verwachte schade op tijdstip  $t$

$I_{ij}$  = investeringskosten voor investering  $i$  en traject  $j$

$\delta_1$  = discontovoet voor de schade

$\delta_2$  = discontovoet voor de investeringen

$i$  = index investering

$j$  = index traject

De contante waarde van de verwachte schade tot tijdstip  $z$  wordt weergegeven in de eerste term in het rechterlid in de formule van de totale verdisconteerde kosten. De tweede term in het rechterlid geeft de kosten door verwachte schade na tijdstip  $z$  weer. De laatste term geeft de som van de verdisconteerde investeringsbedragen weer. De term  $T_{ij}$  geeft het jaar waarin de investering gerealiseerd moet zijn weer (Duits et al., 2011, p. 9).

Formule investeringskosten (Duits et al., 2011, p. 10):

$$I(u, W) = f_k(C + bu)e^{\lambda(u+W)} \quad (2)$$

Waarin:

$I$  = investeringskosten

$u$  = dijkverhoging

$W$  = som van eerdere dijkverhogingen

$C$  = vaste kosten van investeringen uitgaande van de actuele dijkhoogte

$b$  = variabele kosten van de investeringen uitgaande van de actuele dijkhoogte

$\lambda$  = schaalparameter lambda van dijkverhogingen per centimeter

$f_k$  = vermenigvuldigingsfactor voor de kosten (wordt gebruikt in gevoeligheidsanalyses, is normaal gesproken 1)

De investeringskosten bestaan uit een deel vaste kosten en een deel variabele kosten. Hoge vaste kosten maken dat het aantrekkelijk is om een dijk in één keer flink te verhogen. De variabelen kosten stijgen naarmate een dijk verder wordt verhoogd. Het verband tussen investeringskosten en dijkhoogte hoeft niet lineair te zijn, want investeringskosten kunnen meer dan lineair toenemen met de dijkhoogte. Zo kunnen vervolginvesteringen duurder zijn, omdat de goedkoopste maatregelen voor dijkversterking al zijn uitgevoerd (Kind, 2011, p. 19). Om  $I_{ij}$  te berekenen,

worden de investeringskosten verhoogd met de contante waarde van de extra beheer- en onderhoudskosten die het verhogen van de dijk met zich meebrengt, waarvoor een percentage  $\omega$  van de investeringskosten wordt aangenomen.

Formule investeringskosten voor investering  $i$  en traject  $j$  (Duits et al., 2011, p. 10):

$$I_{ij} = I(u, W) \left(1 + \frac{\omega}{\delta_2}\right) \quad (3)$$

De verwachte schade voor een dijkkringdeel wordt berekend door het maximum te nemen van de overstromingskansen per traject vermenigvuldigd met de overstromingsschade.

Formule verwachte schade op tijdstip  $t$  (Duits et al., 2011, p. 11):

$$S(t) = \max_j \{P_j(t)\} V(t) \quad (4)$$

Formule overstromingskans op tijdstip  $t$  (Duits et al., 2011, p. 11):

$$P(t) = P(0) e^{\alpha \eta t} e^{-\alpha(H(t)-H(0))} \quad \text{voor } t \geq 0 \quad (5)$$

Waarin:

$P(t)$  = overstromingskans als functie van tijd

$P(0)$  = overstromingskans op  $t = 0$  (zonder dijkverhoging)

$a$  = schaalparameter alfa

$h_{10}$  = decimeringshoogte

$\eta$  = structurele stijging relatieve waterstand in centimeter

$H(t)$  = dijkhoogte als functie van tijd

$H(0)$  = dijkhoogte op  $t = 0$  (zonder dijkverhoging)

In formule 5 is te zien dat de overstromingskans afneemt bij een toename van de dijkhoogte. Hoeveel deze afneemt is afhankelijk van schaalparameter  $a$ , die de relatie tussen waterstand en overstromingskans bepaalt. Hoe groter de waarde van  $a$  hoe groter de afname in overstromingskans.  $a$  kan worden uitgedrukt in decimeringshoogte van de kruinhoogte. Kruinhoogte is het hoogste punt van een dijk. Decimeringshoogte betekent de dijkverhoging die als gevolg een 10 keer lagere overstromingskans heeft. Stel dat de decimeringshoogte van een dijkkringdeel 30 centimeter is en de overstromingskans  $\frac{1}{500}$  is. Als het dijkkringdeel met 30 centimeter wordt verhoogd, dan wordt de overstromingskans  $\frac{1}{5000}$ . Het kost relatief weinig om een tien keer kleinere overstromingskans te krijgen bij een lage decimeringshoogte ten opzichte van een plek met een hoge decimeringshoogte. Achter stormvloedkeringen zijn de decimeringshoogten over het algemeen laag, langs de kust en in de bovenrivieren zijn ze relatief hoog. Parameter  $\eta$  betreft een relatieve waterstand stijging, omdat waterstand stijging als gevolg van klimaatverandering nog wordt vermeerderd met bodemdaling.

Formule schaalparameter  $a$  (Kind, 2011, p. 20):

$$\alpha = \frac{\ln(10)}{\Delta H_{10}} \approx \frac{2.3}{\Delta H_{10}} \quad (6)$$

De decimeringshoogte is in formule 6 weergegeven als  $\Delta H_{10}$ . De decimeringshoogte is omgekeerd evenredig met schaalparameter  $a$ . Met andere woorden, de effectiviteit van dijkverhoging is groter als de decimeringshoogte kleiner is (Kind, 2011, p. 20).

Formule overstromingsschade op tijdstip  $t$  (Duits et al., 2011, p. 11):

$$V(t) = V(0)e^{\gamma t} e^{\psi \eta t} e^{\zeta(H_{j^*}(t) - H_{j^*}(0))} \quad \text{voor } t \geq 0 \quad (7)$$

Waarin:

$V(t)$  = overstromingsschade als functie van tijd

$V(0)$  = overstromingsschade op  $t = 0$  (zonder dijkverhoging)

$\gamma$  = tempo van de economische groei

$\psi$  = parameter psi voor extra schade door peilstijging

$\zeta$  = parameter zeta voor stijgingstempo schade

$j^*$  = dijktraject waarvoor de dijkhoogte in het horizontale vlak het kleinste is

De overstromingsschade zal in de loop van de tijd toenemen als gevolg van economische en demografische groei (Kind, 2011, p. 21). De overstromingsschade stijgt mee met de jaarlijkse stijging van de waterstand, namelijk  $\psi \cdot \eta$ , omdat een diepere overstroming in meer schade resulteert. De toename van de overstromingsschade is daarmee afhankelijk van klimaatverandering en bodemdaling (Kind, 2011, p. 21). Van tevoren wordt bepaald welk dijktraject in een dijkkringdeel de kleinste dijkhoogte heeft. Voor het berekenen van de overstromingsschade wordt gedurende de gehele tijdsperiode het traject met de initieel kleinste dijkhoogte gebruikt. Er is namelijk de aanname gemaakt dat de dijkhoogte in dit dijktraject gedurende de gehele tijdsperiode het kleinste blijft. Parameter  $\psi$  kan worden toegepast bij zeedijken en dijken langs het IJsselmeer (Duits et al., 2011, p. 11). Parameter  $\zeta$  kan worden toegepast bij alle dijken en brengt tot uiting dat bij een dijk met een hoger beschermingsniveau de omstandigheden waaronder deze faalt extremer zijn (Duits et al., 2011, p. 11).

Formule overstromingsschade op  $t = 0$  (Duits et al., 2011, p. 12):

$$V(0) = f_v \cdot \text{schade} + \text{aantal getroffen} \cdot \text{bedrag per getroffen} + \text{aantal dodelijke slachtoffers} \cdot \text{bedrag per dodelijk slachtoffers} \quad (8)$$

De formules die eerder genoemd zijn, hebben als doel om per dijkkringdeel de optimale investeringsstrategie te bepalen waarbij de som van de verdisconteerde investeringskosten en verwachte schades minimaal is. De overstromingsschade kan berekend worden door de kans op schade te vermenigvuldigen met de verwachte schade. Deze berekening kan worden omgeschreven en met de volgende formule kan de economisch optimale overstromingskans  $P_{midden}$  berekend worden.

Formule economisch optimale overstromingskans  $P_{midden}$  (Duits et al., 2011, p. 12):

$$P_{midden}(t) = \frac{S_t^{midden}}{V(t)} \quad \text{voor } t \geq 0 \quad (9)$$

Waarin:

$P_{midden}(t)$  = middenkans op tijdstip  $t$

$S_t^{midden}$  = logaritmisches gemiddelde van de verwachte schade direct voor en na

*investeren*

$V(t)$  = *overstromingsschade op tijdstip  $t$*

Formule logaritmisch gemiddelde van de verwachte schade direct voor en na investeren (Duits et al., 2011, p. 12):

$$S_{k+1}^{midden} = \frac{S_{k+1}^- - S_{k+1}^+}{\ln(S_{k+1}^-) - \ln(S_{k+1}^+)} \quad \text{voor } t \in [t_k, t_{k+1}) \quad \text{en } k = 0, 1, 2, \dots \quad (10)$$

Waarin:

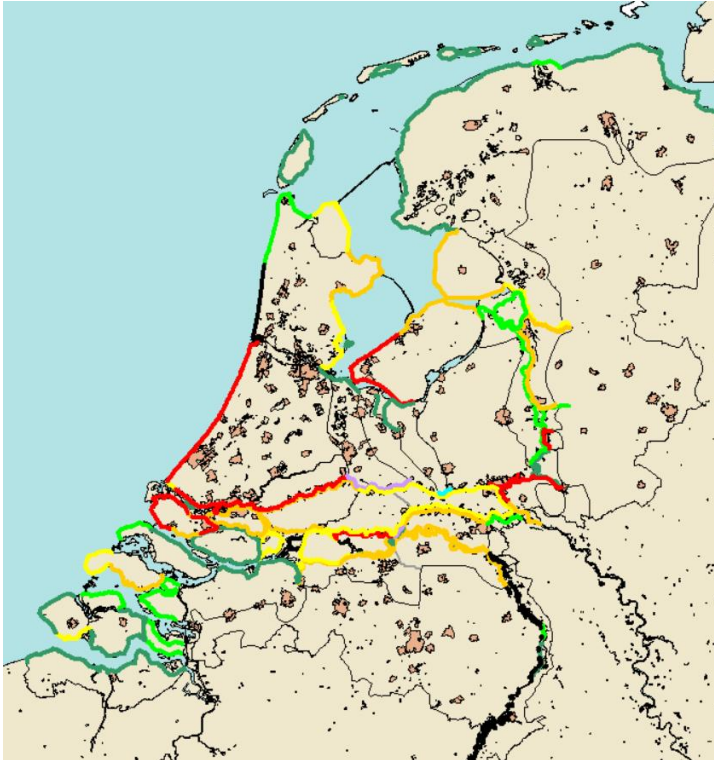
$S_{k+1}^+$  = *verwachte schade onmiddellijk na investeren op tijdstip  $t_{k+1}$*

$S_{k+1}^-$  = *verwachte schade onmiddellijk voor investeren op tijdstip  $t_{k+1}$*

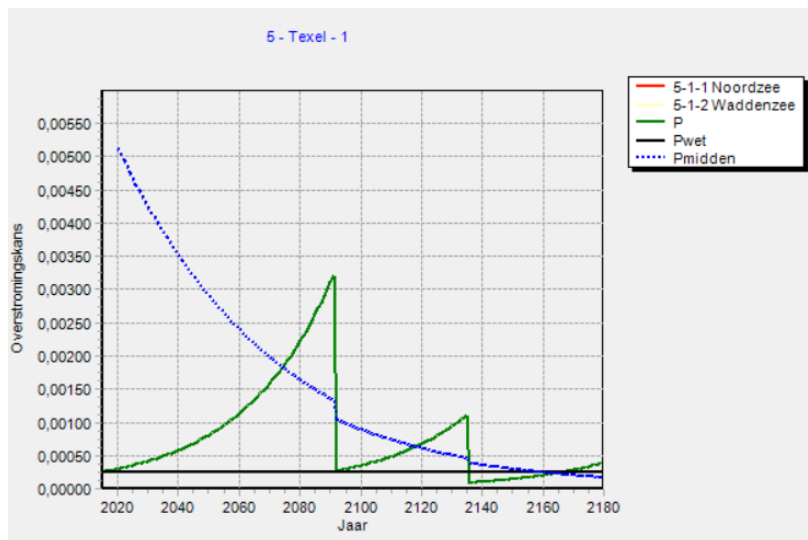
## Resultaten

Het model geeft als uitkomst een kaart van Nederland met daarop de dijkkringdelen waarvoor de economisch optimale overstromingskansen berekend zijn (zie figuur 3). In figuur 3 is donkergroen een hoge overstromingskans dus een minder strenge norm en donkerrood een lage overstromingskans dus een strengere norm. Zo is in figuur 3 te zien dat er in het kustgebied bij de randstad een strengere norm is dan bijvoorbeeld in het Waddengebied. Verder geeft OptimaliseRing een uitvoerbestand terug waar de gegevens in staan die gebruikt zijn voor de berekeningen. Ook staan in dit bestand de uitkomsten van de berekeningen. Dat wil zeggen, per investering het jaartal, de verhoging en wat dit gaat kosten, verder de optimale overstromingskansen in 2020 en 2050 voor het gehele dijkkringdeel en de totale investeringskosten. Naast het uitvoerbestand levert OptimaliseRing per dijkkringdeel een grafiek waarin het verloop van de investeringen te zien is (zie figuur 4). In figuur 4 is te zien dat er ongeveer geïnvesteerd wordt op het punt dat de middenkans en de overstromingskans elkaar snijden. Er kan ongeveer 20 jaar overheen gaan voordat de investering uitgevoerd is. De toppen van de overstromingskansen zijn de momenten vlak voordat de investering afgerond is. Op het moment dat de investering afgerond is, daalt de overstromingskans, omdat het veiligheidsniveau is toegenomen.





Figuur 3: Output van het model *OptimaliseRing*: kaart van Nederland met voor ieder dijkkringdeel de overstromingskans met een kleur aangeduid. Wanneer er op een dijkkringdeel wordt geklikt, dan verschijnt de overstromingskans van dat dijkkringdeel.



Figuur 4: Output van het model *OptimaliseRing*: voorbeeld van een grafiek met de investeringsstrategie voor dijkkring 5: Texel.

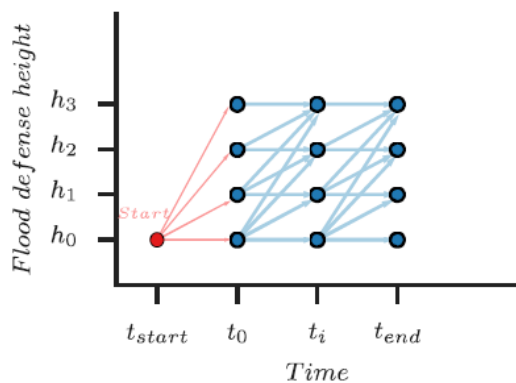
#### H2.4: *OptimaliseRing* Discreet

*OptimaliseRing* Discreet berekent ook de optimale investeringsstrategie (de optimale investeringsniveaus met bijbehorende jaartallen). In dit model werden echter nog geen economisch optimale overstromingskansen berekend. Dit is iets wat in deze stage geïmplementeerd wordt. Verder wordt er in dit model gebruik gemaakt van discrete wiskunde in plaats van continue wiskunde.

Het optimaliseren voor dijkversterking is complex, daarom is er in OptimaliseRing Discreet voor gekozen om het optimaliseren voor dijkversterking te beschrijven als een wiskundig optimalisatie probleem en het optimum te laten bepalen door CPLEX. CPLEX is een softwarepakket die geschikt is om relatief grootschalige problemen door te rekenen. Om achtergrondinformatie op te doen over OptimaliseRing Discreet is het document “Implementatie van Optimalisering-discreet en Dique-Opt” gebruikt, wat ontvangen is via persoonlijke communicatie. Naar dit document zal in het vervolg van het rapport verwezen worden als: (P. de Grave, persoonlijke communicatie, 14 maart 2022).

### Wiskundige optimalisatie

Het doel van de wiskundige optimalisatie is om per dijkringdeel te bepalen in welk jaar naar welk veiligheidsniveau verhoogd moet worden, zodat de som van de verwachte schade en investeringskosten geminimaliseerd wordt. De wiskundige optimalisatie bestaat uit een combinatie van de doelfunctie die geminimaliseerd wordt en een verzameling van voorwaarden waaraan de gekozen optimale oplossing moet voldoen. De doelfunctie is de som van de investeringskosten en verwachte schade en moet geminimaliseerd worden. Voor één dijkringdeel kan het optimalisatieprobleem weergegeven worden als een netwerkprobleem (zie figuur 5). De verzameling van tijdstippen en hoogtes is hierin discreet gemaakt. Als een tak horizontaal loopt, dan betekent dit dat er op dit tijdstip geen investeringskosten worden gemaakt, als de tak schuin omhoog loopt is dit wel het geval. Een tak kan niet schuin omlaag lopen, omdat de dijken niet zullen worden verzwakt. Het doel van het optimalisatieprobleem is om het pad te vinden met de laagste totale kosten. Dit probleem is daarmee een variant op het kortste pad probleem.



Figuur 5: Ophogen van dijkringdelen in de tijd in de vorm van een netwerkprobleem. Overgenomen uit (P. de Grave, persoonlijke communicatie, 14 maart 2022).

Het model bevat twee groepen beslisvariabelen, namelijk CY en DY. Deze beslisvariabelen zijn binair en kunnen dus de waarde 0 of 1 aannemen. Het doel van de optimalisatie is om de juiste combinatie van waardes voor de beslisvariabelen te vinden die voldoet aan de voorwaarden en leidt tot de minimale totale kosten.

Beslisvariabele CY geeft voor elk dijktraject ( $l$ ) voor elk tijdstip ( $t$ ) aan van welk startniveau ( $h_1$ ) het dijktraject wordt verhoogd naar niveau  $h_2$ , waarbij  $h_2 \geq h_1$ . Als een dijktraject wordt verhoogd, dan geldt  $CY(t, l, h_1, h_2) = 1$ , anders is dit 0. Het optimale pad wordt gevormd door alle combinaties van  $t$ ,  $l$ ,  $h_1$  en  $h_2$  waarvoor geldt  $CY(t, l, h_1, h_2) = 1$ , tenzij er geen investeringen plaatsvinden (P. de Grave, persoonlijke communicatie, 14 maart 2022).

Beslisvariabele DY geeft voor elk tijdstip (t) weer welk dijktraject de grootste kans heeft om te falen in het dijkkringdeel en wat het niveau is van dit dijktraject. Er is aangenomen als er een overstroming plaatsvindt in een dijkkringdeel, dat die dan plaatsvindt in het zwakste dijktraject in een dijkkringdeel, die grootste kans op falen heeft (P. de Grave, persoonlijke communicatie, 14 maart 2022). Het zwakste dijktraject is dus bepalend voor de verwachte overstromingsschade in een dijkkringdeel.

De keuze voor de CY waarden is direct bepalend voor de uitkomst van de DY waarden, omdat wanneer en tot welk niveau er wordt versterkt, invloed heeft op welk dijktraject het zwakste is. Hierdoor is DY in principe niet nodig, maar deze variabele maakt de wiskundige beschrijving van het wiskundig probleem wel eenvoudiger (P. de Grave, persoonlijke communicatie, 14 maart 2022). In de voorwaarden wordt rekening gehouden met de relatie tussen CY en DY.

Het wiskundig model bestaat uit invoer, een doelfunctie en voorwaarden. Het wiskundig model ziet er als volgt uit:

Het model heeft de volgende invoer (Zwaneveld, Verweij & Van Hoesel 2018, pp. 10):

- Investeringskosten als functie van tijdstap, dijktraject, startniveau en eindniveau, genaamd  $cost(t, l, h_1, h_2)$ .
- Overstromingskansen als functie van tijdstap, dijktraject en niveau, genaamd  $prob(t, l, h_2)$ .
- Verwachte schade als functie van tijdstap, dijktraject en niveau, genaamd  $damage(t, l, h_2)$ .
- Verwachte schade laatste tijdstap tot oneindige tijdshorizon als functie van dijktraject en hoogte.

Doelfunctie (Zwaneveld et al., 2018, p. 11):

$$\begin{aligned} \text{Min} \quad & \sum_{t \in T} \sum_{l \in L} \sum_{h_1 \in H^l} \sum_{h_2 \geq h_1 \in H^l} cost(t, l, h_1, h_2) \cdot CY(t, l, h_1, h_2) \\ & + \sum_{t \in T} \sum_{l \in L} \sum_{h_2 \in H^l} prob(t, l, h_2) \cdot damage(t, l, h_2) \cdot DY(t, l, h_2) \end{aligned}$$

De doelfunctie minimaliseert de totale kosten, met andere woorden de som van de investeringskosten en de verwachte schade. De eerste regel geeft de investeringskosten weer en de tweede regel de verwachte schade.

Voorwaarden (Zwaneveld et al., 2018, p. 11):

$$CY('0', l, '0', '0') = 1; CY('0', l, h_1, h_2) = 0 \quad \forall l \in L; h_1, h_2 \in H^l; h_2 \geq h_1 \wedge h_2 > '0' \quad (C.2)$$

$$\sum_{h_1 \leq h_2 \in H^l} CY(t-1, l, h_1, h_2) = \sum_{h_2 \geq h_1 \in H^l} CY(t, l, h_2, h_1) \quad \forall t \in T / \{ '0' \}, l \in L, h_2 \in H^l \quad (C.3)$$

$$\sum_{h_1 \in H^l} \sum_{\substack{h_2 \geq h_1 \in H^l: \\ prob(t, l, h_2) > prob(t, l^*, h_2^*)}} CY(t, l, h_1, h_2) + \sum_{l_h \in L} \sum_{\substack{h_2 \in H^{l_h}: \\ prob(t, l_h, h_2) \leq prob(t, l^*, h_2^*)}} DY(t, l_h, h_2) \leq 1 \quad (C.4)$$

$$\forall t \in T / \{ '0' \}, l \in L, l^* \in L, h_2^* \in H^{l^*}$$

$$\sum_{l \in L} \sum_{h_2 \in H^l} DY(t, l, h_2) = 1 \quad \forall t \in T / \{ '0' \} \quad (C.5)$$

$$CY(t, l, h_1, h_2) \in \{0, 1\} \quad \forall t \in T, l \in L, h_1 \in H^l, h_2 \geq h_1 \in H^l \quad (C.6)$$

$$DY(t, l, h_2) \in \{0, 1\} \quad \forall t \in T, l \in L, h_2 \in H^l \quad (C.7)$$

In deze voorwaarden is  $t$  de tijdstap,  $l$  het traject,  $h_1$  de beginhoogte en  $h_2$  de eindhoogte. Hieronder is een korte uitleg bij iedere voorwaarde gegeven:

C.2: Voor alle dijktrajecten is de startwaarde van het beschermingsniveau 0.

C.3: Continuïteitsvergelijking: als een dijktraject aan het eind van de periode  $t-1$  wordt opgehoogd tot beschermingsniveau  $h$ , dan heeft dit dijktraject aan het begin van periode  $t$  hetzelfde beschermingsniveau  $h$ .

C.4:  $DY$  representeert het zwakste dijktraject in het dijkkringdeel.

C.5: Per dijkkringdeel per tijdstip  $t$  kan er exact één zwakste dijktraject zijn. Dit is nodig om de overstromingsschade te bepalen.

C.6:  $CY$  is binair en kan dus alleen de waarde 0 of 1 aannemen.

C.7:  $DY$  is binair en kan dus alleen de waarde 0 of 1 aannemen.

De voorwaarden zijn clusters van voorwaarden. De totale hoeveelheid voorwaarden is afhankelijk van de dimensies van het probleem. Het totaal aantal voorwaarden kan met de volgende formules berekend worden:

$$Totaal\ aantal\ voorwaarden = C2 + C3 + C4 + C5$$

$$C2 = 2l$$

$$C3 = (t - 1)hl$$

$$C4 = thl^2$$

$$C5 = t$$

Bijvoorbeeld als er 39 tijdstappen ( $t$ ), 22 mogelijke beschermingsniveaus ( $h$ ) en 3 dijktrajecten ( $l$ ) zijn, dan is het aantal voorwaarden gelijk aan 10.275.

### Discretisatie

De discretisatie zit in de invoerparameters. Zo wordt in OptimaliseRing Discreet een tijdsperiode genomen van 2015 tot 2300 met tijdstappen van 5 jaar tot het jaar 2100

<sup>2</sup> In de oorspronkelijke documentatie is ervoor gekozen om de doelfunctie C.1 te noemen en de eerste voorwaarde C.2 te noemen, dit is overgenomen in het kader van consistentie.

en tijdstappen van 10 jaar daarna. Verder zijn er 22 veiligheidsniveaus. Deze veiligheidsniveaus zijn voor verhogingen tot 1 meter in stappen van 10 centimeter, tussen 1 meter en 2 meter in stappen van 20 centimeter en tussen 2 en 4 meter in stappen van 30 centimeter. Volgens (Zwaneveld et al., 2018, p. 14), die het model meerdere keren hebben opgelost met verschillende discretisatie schema's, hangt de optimale oplossing nauwelijks af van het discretisatie schema. Er kan dus ook voor een ander discretisatie schema worden gekozen.

## CPLEX

Het probleem is geformuleerd als een mixed-integer probleem in CPLEX voor Python. Een mixed-integer probleem is een probleem waarbij de doelfunctie en voorwaarden lineair zijn. Verder is een deel van de beslisvariabelen geheeltallig. In OptimaliseRing Discreet zijn alle beslisvariabelen (CY en DY) binair, dus geheeltallig. De algemene vorm van een mixed-integer probleem ziet er als volgt uit (P. de Grave, persoonlijke communicatie, 14 maart 2022):

Minimaliseer  $c^T x$   
 zodanig dat :  
 $Ax \leq b$   
 $lb \leq x \leq ub$   
 $x_j$  integer ,  $j \in J$

waarbij:

- x:  $n \times 1$  vector van beslisvariabelen
- c:  $n \times 1$  vector met coëfficiënten van de doelfunctie
- A:  $m \times n$  matrix met coëfficiënten van de constraints
- b:  $m \times 1$  vector met constanten van de constraints
- lb:  $n \times 1$  vector met ondergrenswaarden van x
- ub:  $n \times 1$  vector met bovengrenswaarden van x
- J: index die de deelverzameling van x-variabelen aangeeft die geheeltallig (integer) zijn.

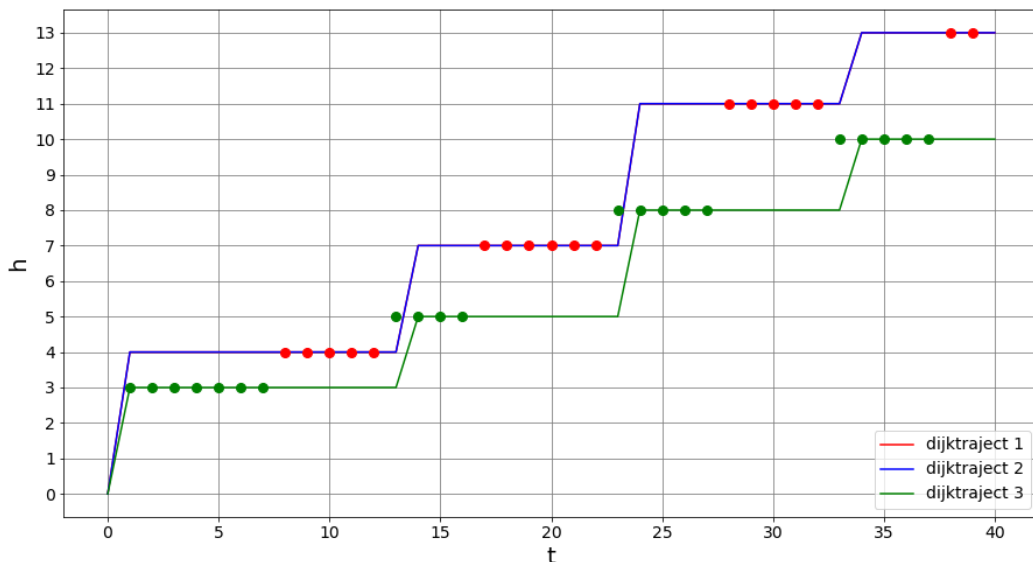
Het probleem van OptimaliseRing Discreet bevat “m” aantal voorwaarden en “n” aantal beslisvariabelen. De eerste stap van het formuleren van het mixed-integer probleem in CPLEX is het inlezen van de parameters. Als tweede stap wordt het CPLEX model gecreëerd en vervolgens wordt het model gevuld met de doelfunctie en voorwaarden. Verder worden de beslisvariabelen CY en DY opgeslagen als meerdimensionale lijst en als 1-dimensionale lijst en worden de beslisvariabelen toegevoegd aan het model (P. de Grave, persoonlijke communicatie, 14 maart 2022). Ook worden de coëfficiënten van de doelfunctie en boven- en ondergrenzen meegegeven aan het model. Als laatste stap wordt het model aangeroepen en wordt de berekening gestart.

## Resultaten

De resultaten van OptimaliseRing Discreet worden geploteerd zoals in figuur 6. In dit figuur liggen dijktraject 1 en 2 op elkaar. In figuur 6 is voor elk dijktraject voor elk tijdstip af te lezen wat de verhoging in veiligheidsniveau is. Zo worden dijktraject 1 en 2 op tijdstip 13 met 3 veiligheidsniveaus verhoogd. Dit is een verhoging van 7

veiligheidsniveaus ten opzichte van het veiligheidsniveau op tijdstip 0. De gekleurde stippen geven in figuur 6 per tijdstap het zwakste traject aan. Zo is op tijdstip 35 dijktraject 3 het zwakste traject. Het zwakste traject is het traject dat de grootste overstromingskans heeft op het bijbehorende tijdstip. Ten gevolge van de vaste kosten voor dijkversterking is het vaak economisch optimaal om dijktrajecten tegelijkertijd te verhogen. In figuur 6 is dat ook te zien, de dijktrajecten worden namelijk allemaal op hetzelfde tijdstip verhoogd.

dijkversterking voor alle dijktrajecten + zwakste vak



Figuur 6: Output *OptimaliseRing Discreet*: voorbeeld resultaat voor een dijkkringdeel met drie dijktrajecten. “h” is het veiligheidsniveau en “t” is de tijdstap. Traject 1 en 2 liggen op elkaar. Overgenomen uit (P. de Grave, persoonlijke communicatie, 14 maart 2022).

## H2.5: Statistische methoden

Er wordt onderzocht welke statistische methoden kunnen worden gebruikt om vast te stellen of er een statistisch significant verschil is tussen de overstromingskansen van *OptimaliseRing* en *OptimaliseRing Discreet*, dit is een eis. Er wordt op gelet in welke situaties deze methoden gebruikt kunnen worden; bijvoorbeeld of de eis dat de verschillen normaal verdeeld zijn wel of geen rol speelt. Verder wordt er onderzocht hoe de verschillen tussen de overstromingskansen van de modellen *OptimaliseRing* en *OptimaliseRing Discreet* verklaard kunnen worden met behulp van verklarende statistiek; dit is ook een eis. Ook hierbij wordt er gelet op de eisen van deze methoden en wordt er naar een alternatief gezocht voor als er niet aan deze eisen wordt voldaan.

Er is zelfstandig onderzoek gedaan naar de statistische methoden die kunnen worden toegepast om te testen of er een statistisch significant verschil is tussen de economisch optimale overstromingskansen in *OptimaliseRing* en *OptimaliseRing Discreet*. Ook is er onderzoek gedaan naar hoe de verschillen verklaard kunnen worden. Voor dit onderzoek is het boek “Statistiek om mee te werken” gebruikt. Ook zijn op internet een aantal andere artikelen en bronnen gevonden. De volgende zoektermen zijn hiervoor gebruikt: (binomial) sign test, wilcoxon signed rank test, linear regression, multiple regression, robust regression, test difference between two

groups, breusch pagan test en paired measurements test. Er is in verschillende databanken gezocht, waaronder Google Scholar en Science Direct.

Om erachter te komen of de verschillen in uitkomsten van OptimaliseRing en OptimaliseRing Discreet statistisch significant zijn, zal er een verschiltoets worden uitgevoerd op de uitkomsten van de modellen. Met een verschiltoets kunnen twee steekproeven worden vergeleken. De vraag is vervolgens of de waargenomen uitkomsten van een bepaalde variabele opvallend verschillen voor de twee steekproeven (Buijs, 2017, p. 339). In dit geval zijn de twee steekproeven de uitkomsten van de economisch optimale overstromingskansen van OptimaliseRing en OptimaliseRing Discreet. Om te weten wat voor soort verschiltoets er toegepast kan worden, is het nodig om te weten of het om gepaarde of ongepaarde waarnemingen gaat. Per dijkkringdeel zijn er twee uitkomsten, namelijk de economisch optimale overstromingskans van OptimaliseRing en OptimaliseRing discreet. Er kan per dijkkringdeel worden vastgesteld of het om een toename of afname gaat. Dit maakt dat het om gepaarde waarnemingen gaat. Als voor een object steeds twee meetresultaten bekend zijn, kunnen er toetsen voor gepaarde waarnemingen worden gebruikt, namelijk de tekentoets of de t-toets (Buijs, 2017, p. 348).

Bij het uitvoeren van statistische toetsen is er sprake van een nulhypothese en een alternatieve hypothese. De nulhypothese is een standpunt wat vooraf ingenomen wordt over de variabele die bestudeerd wordt (Buijs, 2017, p. 281). De nulhypothese wordt op juistheid onderzocht. Als op basis van de verzamelde gegevens wordt geconcludeerd dat de nulhypothese niet geloofwaardig is, dan wordt er overgestapt op een ander standpunt, namelijk de alternatieve hypothese (Buijs, 2017, p. 281).  $\alpha$  is de kans om ten onrechte de nulhypothese te verwerpen (Buijs, 2017, p. 287). In overleg met de opdrachtgever is er voor  $\alpha = 0,05$  gekozen in dit onderzoek. Als geldt  $p < \alpha$  dan wordt de nulhypothese verworpen (Buijs, 2017, p. 287). Als geldt  $p > \alpha$  dan wordt de nulhypothese behouden (Buijs, 2017, p. 287).

Bij de tekentoets wordt per tweetal waarnemingen vastgesteld of er vooruitgang of achteruitgang is. Voor getallenparen  $x$  en  $y$  wordt bij  $y > x$  een “+” genoteerd. Bij  $y < x$  wordt er een “-“ genoteerd. Het aantal “+” wordt geteld. De hypothesen zijn bij de tekentoets als volgt (Buijs, 2017, p. 349):

$H_0$ : er is geen effect

$H_1$ : er is wel een effect

Als de nulhypothese waar is, dan zijn er ongeveer evenveel “+” als “-“. De kans op een “+” is daarom  $\frac{1}{2}$ . Het gaat hier dus om de binomiale verdeling (Buijs, 2017, p. 349). Een voordeel van de tekentoets is dat het een verdelingsvrije toets is, omdat er geen uitspraak gedaan wordt over de kansverdeling van de waargenomen verschillen (Buijs, 2017, p. 349). Een nadeel van de tekentoets is dat er geen aandacht wordt besteed aan de grootte van de waargenomen verschillen (Buijs, 2017, p. 349).

Bij de gepaarde t-toets worden alle getallenparen  $x$  en  $y$  omgezet in een verschilvariabele  $v$ , waarbij  $v = y - x$  (Buijs, 2017, p. 350). Als de verschilvariabele groter is dan nul, dan is er sprake van vooruitgang. Het uitgangspunt is dat de verschillen worden beschouwd als trekkingen uit de normale verdeling (Buijs, 2017, p. 350).



De hypothesen luiden als volgt (Buijs, 2017, p. 350):

$H_0$ : er is geen effect ( $\mu_v = 0$ )

$H_1$ : er is wel een effect ( $\mu_v \neq 0$ )

Een voordeel van de t-toets is dat de grootte van de waargenomen verschillen wel wordt meegenomen (Buijs, 2017, p. 350). Het nadeel van de t-toets is dat het uitgangspunt is dat de waargenomen verschillen de normale verdeling volgen. Deze toets kan dus niet toegepast worden als de verschillen niet de normale verdeling volgen. Er zal dus van tevoren moeten worden getoetst op normaliteit.

Het toetsen op normaliteit wordt uitgevoerd in twee stappen. De eerste stap is het uitvoeren van de Shapiro-Wilk toets. De hypothesen van deze toets zijn (*Assessing the Assumption of Normality · UC Business Analytics R Programming Guide*, z.d.):

$H_0$ : de variabele is normaal verdeeld

$H_1$ : de variabele is niet normaal verdeeld

De toets is gevoelig voor uitbijters bij grote steekproeven (*Assessing the Assumption of Normality · UC Business Analytics R Programming Guide*, z.d.). Daarom wordt deze toets in combinatie gebruikt met een Q-Q plot. Op de y-as staan de waargenomen verschillen en op de x-as de z-waarden onder de aanname dat de verschillen normaal verdeeld zijn. Als de verschillen normaal verdeeld zijn, dan zal in de Q-Q plot een rechte lijn te zien zijn (*Assessing the Assumption of Normality · UC Business Analytics R Programming Guide*, z.d.).

Een alternatief voor de gepaarde t-toets is de Wilcoxon signed rank toets (Xia, 2020, par. 8.1.1.3). Het is een non-parametrische toets, dat wil zeggen dat er niet vanuit wordt gegaan dat de waarnemingen de normale verdeling volgen (Quené en Van den Bergh, 2022, Hoofdstuk 17). Bij de Wilcoxon signed rank toets wordt er gekeken naar de mediaan van de verschillen (King & Eckersley, 2019, par. 6.3). De mediaan is het middelste getal als de verschillen gerangschikt zijn van laag naar hoog. De hypothesen van de Wilcoxon signed rank toets zijn (King & Eckersley, 2019, par. 6.3):

$H_0$ : de mediaan van de verschillen is 0

$H_1$ : de mediaan van de verschillen is niet 0

Het voordeel van de Wilcoxon signed rank toets is dat deze krachtiger is dan de tekentoets (Scheff, 2016). Dit komt doordat er rekening wordt gehouden met de grootte van de verschillen (Scheff, 2016).

Een van de doelen van regressie is het achterhalen welke van de verklarende variabelen een statistisch significante invloed hebben op de te verklaren variabele en hoe groot deze invloed is (Buijs, 2017). Een lineair regressiemodel ziet er als volgt uit (Buijs, 2017, p. 427):

$X_1, \dots, X_m$  zijn de verklarende variabelen waarmee  $Y$  verklaard wordt.

$$Y = \alpha + \beta_1 X_1 + \dots + \beta_m X_m + \varepsilon$$

Bij enkelvoudige regressie wordt de te verklaren variabele verklaard met één verklarende variabelen. Bij meervoudige regressie gebeurt dit met meerdere verklarende variabelen. Een valide regressiemodel heeft vier voorwaarden, namelijk (Bevans, 2022):

1. De residuen zijn bij benadering normaal verdeeld.
2. De residuen zijn bij benadering gemiddeld nul.
3. De residuen zijn homoscedastisch.
4. De verklarende variabelen vertonen nauwelijks samenhang.



De eerste voorwaarde kan worden gecontroleerd door eerst een Shapiro-Wilk toets uit te voeren en vervolgens een Q-Q plot te maken (zie uitleg hierboven over het toetsen op normaliteit). De tweede voorwaarde kan worden getoetst met de t-toets. De hypothesen zijn dan als volgt (*Understanding the T-Test in Linear Regression*, 2021):

$H_0$ : het gemiddelde van de residuen is gelijk aan nul

$H_1$ : het gemiddelde van de residuen is niet gelijk aan nul

De derde voorwaarde wordt in twee stappen gecontroleerd. De eerste stap is het plotten van de residuen op de y-as en de gefitte waardes op de x-as (*The Breusch-Pagan Test: Definition & Example*, 2022). Als de punten steeds verder of juist steeds dichterbij de horizontale as komen te liggen, kan dat een teken zijn van heteroscedasticiteit (*The Breusch-Pagan Test: Definition & Example*, 2022). Als tweede stap wordt de Breusch-Pagan toets uitgevoerd, mits de residuen normaal verdeeld zijn. De hypothesen voor de Breusch-Pagan toets zijn als volgt (*The Breusch-Pagan Test: Definition & Example*, 2022):

$H_0$ : de residuen zijn homoscedastisch

$H_1$ : de residuen zijn niet homoscedastisch

De vierde voorwaarde geldt niet in het geval van enkelvoudige lineaire regressie, maar wel in het geval van meervoudige lineaire regressie. De voorwaarde kan getest worden met behulp van de VIF-waarde<sup>3</sup> (Glen, 2020). De VIF-waarde wordt berekend voor alle verklarende variabelen. Een VIF-waarde van boven de 10 of meerdere VIF-waarden van boven de 5 wijzen op samenhang tussen de variabelen (Glen, 2020).

Mocht het regressiemodel niet aan alle hierboven genoemde voorwaarden voldoen, dan is het niet valide en kan het niet gebruikt worden om de verschillen te verklaren. In dat geval kan er gebruik worden gemaakt van robuuste regressie. Robuuste regressie is een alternatief voor lineaire regressie wanneer de gegevens vervuild zijn met uitbijters of invloedrijke waarnemingen (*Robust Regression | R Data Analysis Examples*, z.d.). Ook hoeft een robuust regressiemodel niet gevalideerd te worden.

Alle hierboven genoemde methoden voldoen aan de eisen, namelijk dat ze kunnen vaststellen dat er een significant verschil is of dat ze de verschillen kunnen verklaren. Echter hebben de methoden ook voor- en nadelen. Zo is een voordeel van de tekentoets dat het een verdelingsvrije toets is en een nadeel van de tekentoets dat er geen aandacht wordt besteed aan de grootte van de waargenomen verschillen. Een voordeel van de t-toets is dat de grootte van de waargenomen verschillen wel wordt meegenomen en een nadeel van de t-toets is dat de toets alleen kan worden gebruikt als de verschillen de normale verdeling volgen. De t-toets kan dus wellicht niet toegepast worden, als de verschillen niet de normale verdeling volgen. Een voordeel van de Wilcoxon signed rank toets is dat er rekening wordt gehouden met de grootte van de verschillen en dat er geen eis van normaliteit is bij deze toets. Een nadeel van lineaire regressie is dat het model aan vier eisen moet voldoen om valide te zijn en anders niet gebruikt kan worden voor het verklaren van de verschillen. Om deze reden is er naar een alternatief gezocht en dit alternatief is robuuste regressie. Het voordeel van robuuste regressie is dat dit model niet gevalideerd hoeft te worden.

---

<sup>3</sup> Variance inflation factor

### H3: Testen aan kwaliteitscriteria

Het doel van de stage is om de modellen OptimaliseRing en OptimaliseRing Discreet te vergelijken. Dit gebeurt aan de hand van kwaliteitscriteria die in dit hoofdstuk worden opgesteld. Op basis van interviews met een aantal deskundigen op het gebied van overstromingskansen en mensen die verstand hebben van de modellen OptimaliseRing en OptimaliseRing Discreet (André Wooning, Robert Slomp, Jarl Kind & Peter de Grave, persoonlijke communicatie, 16-18 maart 2022) kwamen de kwaliteitscriteria naar voren. De kwaliteitscriteria zijn gecategoriseerd in de volgende categorieën: juistheid, verklaarbaarheid, stabiliteit, aanpasbaarheid, gebruiksvriendelijkheid, reproduceerbaarheid en rekentijd. Per kwaliteitscriterium categorie zullen de criteria uitgelegd worden en vervolgens wordt een korte test of omschrijving gegeven. De criteria gelden voor zowel OptimaliseRing als OptimaliseRing Discreet, met uitzondering van het tweede criterium van de categorie rekentijd, deze geldt alleen voor OptimaliseRing Discreet. De resultaten van de testen worden in hoofdstuk 6 besproken.

#### H3.1: Juistheid

Juistheid kan worden onderverdeeld in twee criteria. Het eerste criterium gaat over de vraag in hoeverre de economisch optimale overstromingskansen in het jaar 2050 van OptimaliseRing Discreet overeenkomen met de berekende overstromingskansen uit OptimaliseRing? Voor alle dijkkringdelen wordt de economisch optimale overstromingskans in het jaar 2050 berekend met OptimaliseRing Discreet. Per dijkkringdeel wordt het procentuele verschil berekend tussen de uitkomsten van OptimaliseRing en OptimaliseRing Discreet. Met behulp van beschrijvende statistiek wordt het gemiddelde procentuele verschil berekend. (Voorbeeld: de overstromingskansen van OptimaliseRing Discreet liggen gemiddeld 3,2% hoger dan van OptimaliseRing.) De overstromingskans met bijbehorende totale kosten kunnen worden geplot in één grafiek met per dijkkringdeel een stip en voor elk model een andere kleur stippen.

Het tweede criteria dat aan juistheid gekoppeld kan worden is in hoeverre de minimale totale kosten per dijkkringdeel van OptimaliseRing Discreet overeenkomen met de berekende minimale totale kosten uit OptimaliseRing. Voor alle dijkkringdelen worden de minimale totale kosten berekend met OptimaliseRing Discreet. Per dijkkringdeel wordt het procentuele verschil berekend tussen de uitkomsten van OptimaliseRing en OptimaliseRing Discreet. Met behulp van beschrijvende statistiek wordt het gemiddelde procentuele verschil berekend. (Voorbeeld: de minimale totale kosten van OptimaliseRing Discreet liggen gemiddeld 4,5% hoger dan van OptimaliseRing.) De overstromingskans met bijbehorende totale kosten kunnen worden geplot in één grafiek met per dijkkringdeel een stip en voor elk model een andere kleur stippen.

#### H3.2: Verklaarbaarheid

Bij het criterium verklaarbaarheid wordt nagegaan of de uitkomsten van de modellen logisch en uitlegbaar zijn. Met behulp van verklarende statistiek zal er vastgesteld worden of er een significant verschil in overstromingskansen is tussen de modellen. Als er een statistisch significant verschil is, dan zal dit verschil worden verklaard met behulp van regressie.

### H3.3: Stabiliteit

Dit criterium gaat over hoe stabiel het programma is, dat wil zeggen als één parameter relatief weinig verandert, hoe groot is dan de verandering in de uitkomst? Dit criterium zal op de volgende manier worden getest: de volgende parameters zullen één voor één worden veranderd: discontovoet schade, discontovoet investeringen, economisch scenario, bedrag per dodelijk slachtoffer en bedrag per getroffen. De uitkomsten van de minimale totale kosten en overstromingskans per dijkkringdeel worden vergeleken met de uitkomsten van OptimaliseRing en OptimaliseRing Discreet. Per dijkkringdeel per parameter wordt het procentuele verschil in uitkomsten berekend.

### H3.4: Aanpasbaarheid

Er zijn drie criteria te verbinden aan aanpasbaarheid. Met het eerste criterium wordt bedoeld dat de invoerdata makkelijk aanpasbaar is. Er wordt beschreven hoe de invoerdata eruit ziet en of er nieuwe data kan worden ingevoerd. Ook wordt omschreven wat de stappen zijn die gemaakt moeten worden om nieuwe data in te kunnen voeren.

Het tweede criterium gaat over de kwaliteit van de code. In hoeverre is de code voorzien van duidelijk commentaar, duidelijke variabelenamen, geen ongebruikte code, consistentie (spaties, insprong, hoofdletters) en een kleine scope<sup>4</sup>. De kwaliteit van de code is deels bepalend voor hoe makkelijk/moeilijk het is om aanpassingen of toevoegingen te doen in het programma en om het programma te kunnen onderhouden. De code achter OptimaliseRing en OptimaliseRing Discreet wordt geanalyseerd. Per categorie (commentaar, variabelenamen, dode code, consistentie, scope) zal er worden samengevat wat er is waargenomen in de code.

Het derde criterium geeft weer in welke taal en met gebruik van welke pakketten het programma is opgebouwd en hoeveel mensen er zijn met kennis hierover. Er wordt omschreven in welke taal en met welke pakketten de programma's OptimaliseRing en OptimaliseRing Discreet zijn opgebouwd. Er wordt gezocht naar het geschatte aantal mensen met kennis hierover.

### H3.5: Gebruiksvriendelijkheid

Gebruiksvriendelijkheid kan worden onderverdeeld in drie criteria. Het eerste criterium dat aan gebruiksvriendelijkheid gekoppeld kan worden is dat het model als uitkomst per dijkkringdeel de overstromingskans en een grafiek en tabel geeft, waarin af te lezen is wat de optimale tijdstippen zijn om te verhogen en naar welk veiligheidsniveau er verhoogd wordt. Er zal worden beschreven hoe de output van de modellen OptimaliseRing en OptimaliseRing Discreet eruit ziet.

Het tweede criterium bij gebruiksvriendelijkheid is dat het programma te gebruiken is door mensen die niet beschikken over diepgaande programmeerkennis. Er wordt een zeer korte samenvatting gegeven van hoe OptimaliseRing en OptimaliseRing Discreet kunnen worden gebruikt en welke kennis daarvoor nodig is.

Het derde criterium gaat over dat de user interface van het programma duidelijk is en dat er documentatie beschikbaar is over het model. Er wordt omschreven wat er qua

---

<sup>4</sup> Het deel van de broncode waarin de variabele gedefinieerd is.

documentatie beschikbaar is over OptimaliseRing en OptimaliseRing Discreet. Er wordt in enkele zinnen uitgelegd wat er in deze documentatie te vinden is.

### H3.6: Reproduceerbaarheid

Het criterium dat aan reproduceerbaarheid gekoppeld kan worden is dat het model dezelfde uitkomsten geeft, ook als dit meerdere keren gedraaid wordt.

OptimaliseRing en OptimaliseRing Discreet worden twee keer gedraaid. De uitkomsten worden opgeslagen en naderhand met elkaar vergeleken.

### H3.7: Rekentijd

Er zijn twee criteria die gekoppeld kunnen worden aan rekestijd. Het eerste criterium gaat over hoelang het duurt om voor alle dijkkringdelen de economisch optimale overstromingskans te berekenen en voor elk dijkkringdeel een grafiek met optimale tijdstippen en beschermingsniveaus te genereren. OptimaliseRing en OptimaliseRing Discreet worden twee keer gedraaid en de tijd wordt bijgehouden. Hiervan wordt per programma een gemiddelde tijd berekend en deze gemiddelden worden met elkaar vergeleken.

Het tweede criterium dat bij rekestijd hoort, gaat over hoeveel de uitkomst verschilt als OptimaliseRing Discreet afgekapt wordt na een bepaald aantal minuten ten opzichte van OptimaliseRing Discreet laten rennen tot het klaar is en ten opzichte van OptimaliseRing? OptimaliseRing Discreet wordt na een bepaald aantal minuten afgekapt en de resultaten worden opgeslagen. De resultaten worden vergeleken met de resultaten van OptimaliseRing Discreet. De overstromingskans met bijbehorende totale kosten kunnen worden geplot in één grafiek per dijkkringdeel een stip, voor elk model een andere kleur stippen.

### H3.8: Kosten

De kosten waren in eerste instantie meegenomen als kwaliteitscriterium, echter bleek later dat deze niet relevant zijn voor Rijkswaterstaat. Dit komt doordat de software en licenties die nodig zijn voor het gebruik van OptimaliseRing en OptimaliseRing Discreet al zijn aangeschaft in het verleden door Rijkswaterstaat en onbeperkt bruikbaar zijn. Waarschijnlijk zijn de kosten wel relevant voor andere bedrijven die de programma's OptimaliseRing en OptimaliseRing Discreet zouden willen gebruiken en nog niet beschikken over de benodigde software en licenties.

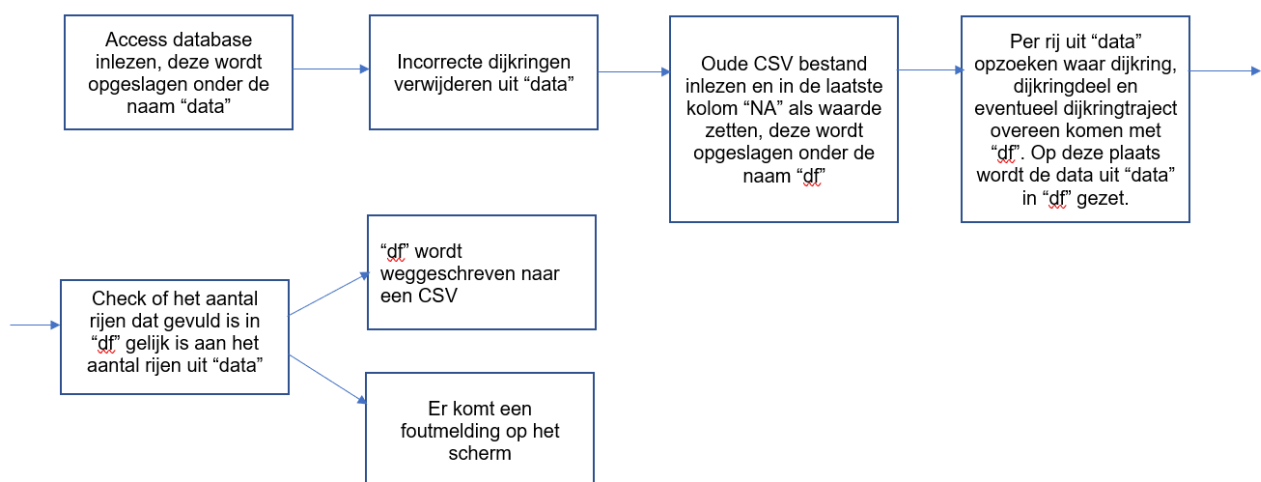
## H4: Bepalen economisch optimale overstromingskansen

In dit hoofdstuk wordt omschreven hoe het model OptimaliseRing Discreet is aangevuld, zodat hiermee input data gegenereerd kan worden en voor elk dijkkringdeel voor elk jaar de economisch optimale overstromingskans kan worden bepaald. De stappen die uitgevoerd zijn om de input data te genereren zijn omschreven in hoofdstuk 4.1. In hoofdstuk 4.2 wordt toegelicht hoe de berekening is geïmplementeerd. Tot slot worden in hoofdstuk 4.3 de resultaten van het model OptimaliseRing Discreet besproken.

### H4.1: Input data genereren

Voordat de economisch optimale overstromingskansen kunnen worden berekend, zijn er eerst een aantal stappen uitgevoerd om de juiste input data te genereren.

Als eerste stap is er nieuwe input data gegenereerd voor het programma “aanmaken\_variabelen.py” met behulp van vier zelf geschreven scripts. De input data betreft acht CSV bestanden die gemaakt zijn met de informatie uit de aangeleverde Access database. De vier scripts lijken op elkaar en in Bijlage 1 wordt stapsgewijs uitgelegd welke stappen er zijn genomen om één CSV bestand te maken. In figuur 7 is een schematische weergave van één van deze scripts.



Figuur 7: Schematische weergave van input data genereren voor het programma “aanmaken\_variabelen.py”.

Ten tweede wordt de gegenereerde input data ingevoerd in het programma “aanmaken\_variabelen.py” en wordt dit programma gedraaid. Hieruit volgen voor ieder dijkkringdeel vier bestanden.

Tot slot worden de bestanden die gegenereerd zijn in stap 2 meegegeven als input voor het programma “OptimaliseringMIP.py” en wordt dit programma gedraaid. Hieruit komt per dijkkringdeel de optimale investeringsstrategie als afbeelding en als CSV bestand. Hierin staat per dijkkring per veiligheidsniveau in welk jaar er een verhoging plaatsvindt.

#### H4.2: Berekening implementeren

De berekening van de economisch optimale overstromingskansen is in deze stage geïmplementeerd in het Python programma Berekening economisch optimale overstromingskansen. Het programma wordt gedraaid met als input data de CSV bestanden die in de laatste stap in hoofdstuk 4.1 gegenereerd zijn. De gegevens uit de Access database staan in Python opgeslagen in twee dictionary's. De uitkomsten van de berekeningen worden opgeslagen in een aparte dictionary. In Bijlage 2 in tabel 1, 2 en 3 staat omschreven wat er in deze dictionary's zit. Er is gekozen om deze gegevens op te slaan in het datatype dictionary, omdat het zoeken in een dictionary een complexiteit heeft van  $O(1)$ . Dit is sneller dan bijvoorbeeld het zoeken in een lijst, omdat dit een complexiteit heeft van  $O(n)$ .

De economisch optimale overstromingskansen worden per dijkkringdeel berekend. Hieronder is de berekening van de economisch optimale overstromingskansen beschreven. De formules en achtergrondinformatie hierover is te vinden in hoofdstuk 2.3. Een schematische weergave van de Python code is weergegeven in Bijlage 1 figuur 1.

Als eerste wordt de overstromingsschade op tijdstip 0, dus zonder dijkverhoging, berekend. Hiervoor wordt formule 8 in hoofdstuk 2.3 gebruikt.

Als tweede wordt de overstromingskans per dijktraject voor elke tijdstap berekend. Een schematische weergave van de deze stap in de code is terug te vinden in Bijlage 1 figuur 2. Om de overstromingskans per dijktraject te berekenen wordt formule 5 in hoofdstuk 2.3 gebruikt. De overstromingskans per dijkkringdeel is het maximum van de overstromingskansen per dijktraject.

Als derde wordt de overstromingsschade en de verwachte schade voor elk dijkkringdeel voor iedere tijdstap berekend. De overstromingsschade is berekend met formule 7 in hoofdstuk 2.3 en de verwachte schade is berekend met formule 4 in hoofdstuk 2.3.

Vervolgens wordt per dijkkringdeel op elk investeringsmoment het logaritmisch gemiddelde van de verwachte schade berekend.  $S_{k+1}^{midden}$  blijft de jaren voor de investering gelijk, tot het investeringsmoment. Bijvoorbeeld als het eerste investeringsmoment in 2071 plaatsvindt, dan is  $S_{k+1}^-$  de verwachte schade in 2071 en  $S_{k+1}^+$  de verwachte schade in 2076. Tot het jaar 2071 geldt hetzelfde logaritmische gemiddelde van de verwachte schade. Om het logaritmische gemiddelde van de verwachte schade te berekenen is formule 10 in hoofdstuk 2.3 gebruikt.

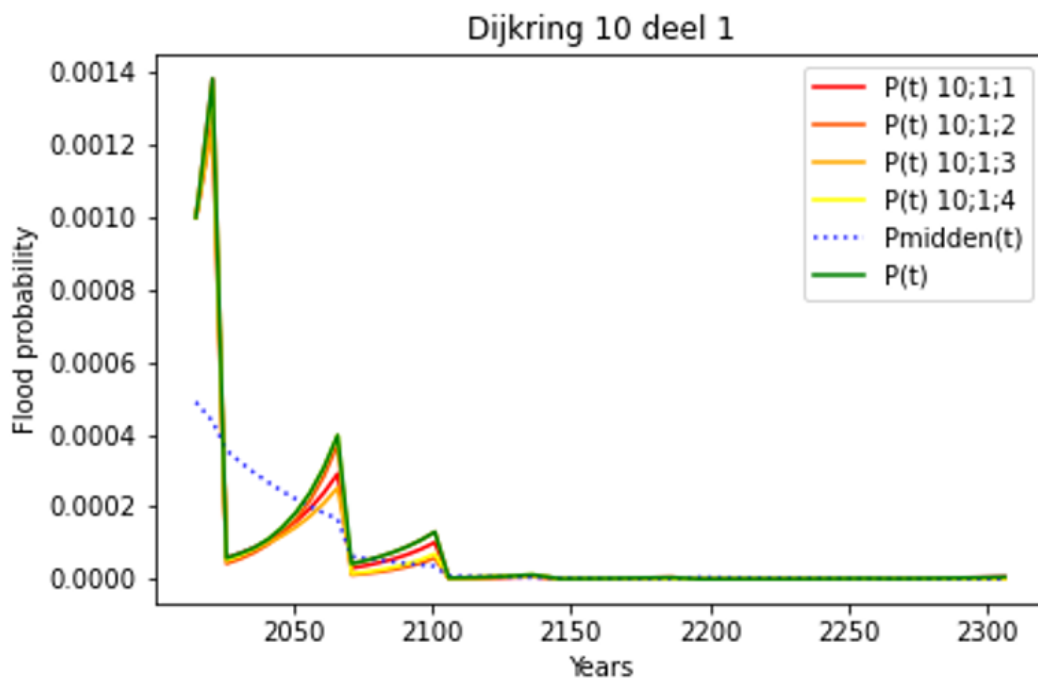
Als laatste wordt de economisch optimale overstromingskans voor elk dijkkringdeel voor iedere tijdstap berekend. Hiervoor is formule 9 in hoofdstuk 2.3 gebruikt.

#### H4.3: Resultaten OptimaliseRing Discreet

OptimaliseRing Discreet bepaalde al de economisch optimale investeringsstrategie, weergegeven in een afbeelding zoals in figuur 6 in hoofdstuk 2.4 te zien is. Hieraan is toegevoegd dat het programma ook voor ieder dijkkringdeel de investeringsstrategie in een CSV bestand opslaat. Hierin staat per traject per veiligheidsniveau wanneer deze investering plaatsvindt. Verder is het programma geautomatiseerd, dat wil

zeggen, het programma kan nu door het één keer te draaien de economisch optimale investeringsstrategie berekenen van alle dijkringdelen.

Tijdens de stage is de berekening van de economisch optimale overstromingskansen geïmplementeerd. OptimaliseRing Discreet berekent in het programma Berekening economisch optimale overstromingskansen, voor alle dijkringdelen per jaar de economisch optimale overstromingskans. In Bijlage 3 zijn de resultaten uit het programma “Berekening\_economisch\_optimale\_overstromingskansen.py” te vinden. In deze tabel staat per dijkringdeel de overstromingskans en de economisch optimale overstromingskans voor het jaar 2021, 2046 en 2051. Verder wordt voor elk dijkringdeel een afbeelding gegenereerd, waarin  $P(t)$  en  $P_{midden}(t)$  voor elk traject in het dijkringdeel is weergegeven (figuur 8). In figuur 8 is af te lezen dat er per traject op dezelfde momenten geïnvesteerd wordt, doordat de lijnen op hetzelfde punt naar beneden gaan. Op dit punt zijn de investeringen uitgevoerd en gaat de overstromingskans omlaag. De economisch optimale overstromingskans  $P_{midden}(t)$  daalt als gevolg van economische groei.



Figuur 8: Output van het programma Berekening economisch optimale overstromingskansen: resultaat van dijkkring 10 deel 1.



## H5: Statistische analyse

Dit hoofdstuk geeft antwoord op de test van het kwaliteitscriterium verklaarbaarheid. Als eerste wordt er een voorbereiding van de data gedaan, voordat de statistische analyse uitgevoerd kan worden. Dit is toegelicht in hoofdstuk 5.1. In hoofdstuk 5.2 wordt het gemiddelde procentuele verschil berekend. Vervolgens worden in hoofdstuk 5.3 verschiltoetsen uitgevoerd. Tot slot wordt in hoofdstuk 5.4 regressie toegepast om een verklaring van de verschillen in uitkomsten tussen de modellen gegeven.

### H5.1: Voorbewerking data

Het is op dit punt nog onjuist om de economisch optimale overstromingskansen van OptimaliseRing en OptimaliseRing Discreet te vergelijken, omdat de kansen in OptimaliseRing bekend zijn voor de jaren 2020 en 2050 en in OptimaliseRing Discreet voor de jaren 2021 en 2051. Daarom worden als eerst de economisch optimale overstromingskansen voor 2020 en 2050 benaderd, met behulp van lineaire interpolatie. De economisch optimale overstromingskans in 2020 is benaderd met de kansen in 2016 en 2021. Voor 2050 is de kans benaderd met de kansen in 2046 en 2051. Er is gekozen om lineaire interpolatie toe te passen, omdat kijkende naar de grafieken, een lineair verband in een kleine tijdstap aannemelijker is dan bijvoorbeeld een kwadratisch verband. Verder is er gekozen voor de jaren 2020 en 2050, omdat er geen andere data dan voor de jaren 2020 en 2050 bekend is van het model OptimaliseRing.

Bij lineaire interpolatie wordt er van twee punten uitgegaan:  $(x_0, y_0)$  en  $(x_1, y_1)$ . Hieruit kan een eerstegraads functie opgesteld worden. Voor de richtingshoek  $\alpha$  van deze lijn geldt (De Joode, 2016, p. 17):

$$\tan \alpha = \frac{y_1 - y_0}{x_1 - x_0}$$

Hieruit kan de interpolatiepolynoom worden afgeleid, namelijk (De Joode, 2016, p. 18):

$$p(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0)$$

Als voorbeeld wordt de economisch optimale overstromingskans in 2050 benadert voor dijkkring 1 deel 1. De economisch optimale overstromingskans in 2046 is hier 0,002091119 en in 2051 is de kans 0,001901607. Het invullen van de formule van de interpolatiepolynoom geeft het volgende:

$$0,002091119 + \frac{0,001901607 - 0,002091119}{2051 - 2046}(2050 - 2046) \approx 0,001939509$$

Hiermee is 0,001939509 de benaderde economisch optimale overstromingskans in 2050. Voor de overige dijkkringdelen is dezelfde werkwijze toegepast om de economisch optimale overstromingskansen in 2050 te benaderen.

### H5.2: Gemiddelde procentuele verschil

Als eerst wordt de verschilvariabele  $V$  aangemaakt.  $X$  is de economisch optimale overstromingskans in OptimaliseRing,  $Y$  de economisch optimale overstromingskans in OptimaliseRing Discreet en  $i$  het dijkkringdeel, dan wordt  $V$  als volgt berekend:

$V_i = Y_i - X_i$ . Het procentuele verschil is per dijkkringdeel berekend op de volgende wijze: *procentuele verschil* =  $\frac{V_i}{X_i} \cdot 100$ . De uitkomsten van de economisch optimale



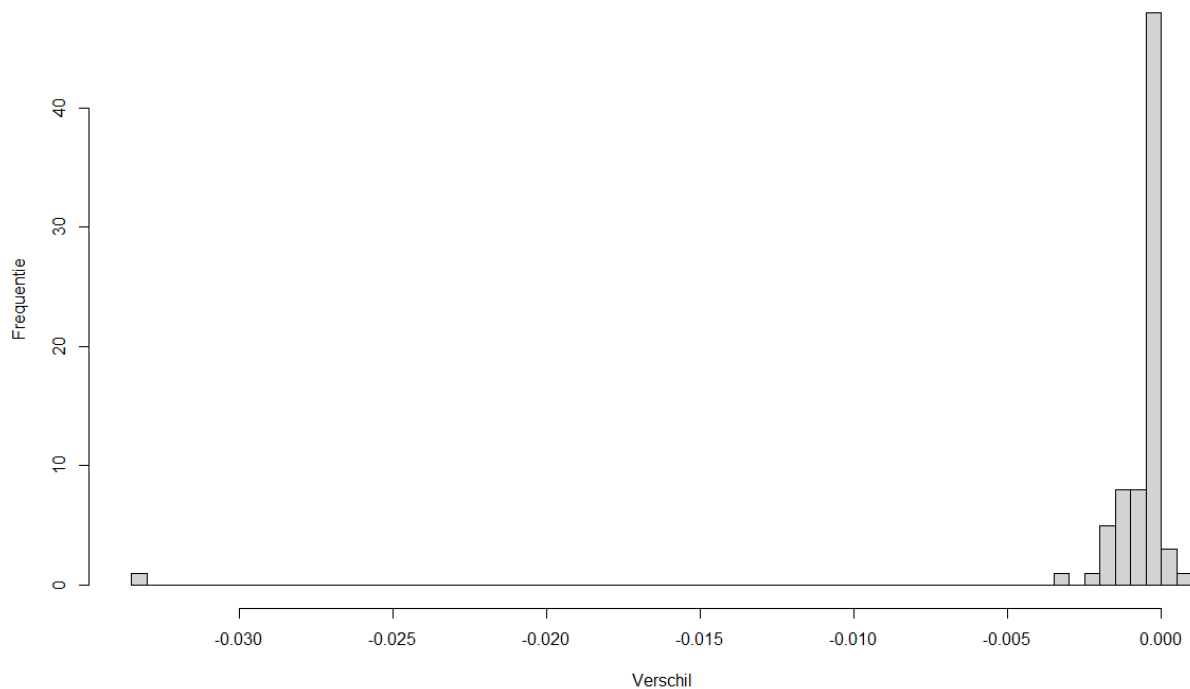
overstromingskansen in 2050 liggen gemiddeld genomen in OptimaliseRing Discreet 50% lager dan in OptimaliseRing. De totale kosten liggen gemiddeld genomen in OptimaliseRing Discreet 19% lager dan in OptimaliseRing.

### H5.3: Verschiltoetsen

Om te toetsen of de verschillen in uitkomsten van de economisch optimale overstromingskansen tussen OptimaliseRing en OptimaliseRing Discreet statistisch significant zijn, worden er verschiltoetsen uitgevoerd. In hoofdstuk 2.5 is beschreven dat het gaat om gepaarde waarnemingen en welke keuzes er voor verschiltoetsen in dat geval zijn. Voordat er een keuze voor een verschiltoets kan worden gemaakt, wordt de verschilvariabele  $V$  eerst getoetst op normaliteit. Verschilvariabele  $V$  is in hoofdstuk 5.2 gedefinieerd.

Als eerste worden de verschillen geplott in een histogram (zie figuur 9). Op basis van deze plot wordt het vermoeden geuit dat de verschillen niet normaal verdeeld zijn. Op basis van deze plot wordt er een uitbijter, namelijk dijkkring 46 deel 1 uit de data verwijderd, omdat hier het absolute verschil erg groot is. Dit is gedaan in overleg met de opdrachtgever. In het vervolg van de statistische analyse zal dijkkring 46 deel 1 niet worden meegenomen in de data.

Histogram verschil in overstromingskans tussen OptimaliseRing en OptimaliseRing Discreet



Figuur 9: Histogram verschil in overstromingskans tussen OptimaliseRing en OptimaliseRing Discreet

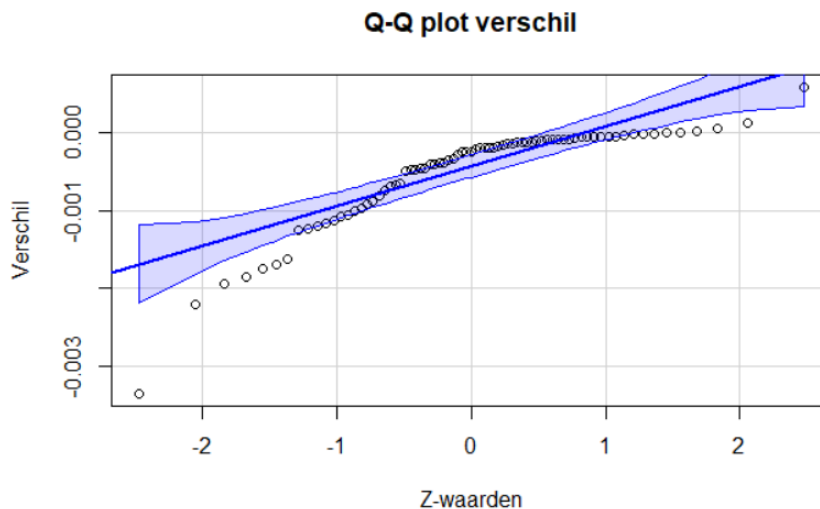
Het toetsen op normaliteit gebeurt in twee stappen. De eerste stap is het uitvoeren van de Shapiro-Wilk toets. De hypothesen van deze toets zijn:

$H_0$ :  $V$  volgt als achterliggende kansverdeling de normale verdeling

$H_1$ :  $V$  volgt niet als achterliggende kansverdeling de normale verdeling

Het resultaat van de Shapiro-Wilk toets is  $p < 0,001$ . Omdat de  $p$ -waarde kleiner is dan  $\alpha$ , wordt de nulhypothese verworpen. Dit wil zeggen dat het niet statistisch significant is om aan te nemen dat de verschilvariabele de normale verdeling als achterliggende kansverdeling heeft,  $W=0,809$ ,  $p < 0,001$ .

Als tweede stap is er een Q-Q plot gemaakt (zie figuur 10). In figuur 10 is te zien dat veel waarden buiten het acceptatiegebied liggen en er enkele uitschieters zijn. De Q-Q plot geeft aanleiding om aan te nemen dat de verschilvariabele niet de normale verdeling als achterliggende kansverdeling heeft.



Figuur 10: Q-Q plot van de verschillen met bijbehorende Z-waarden

Uit de twee stappen is geconcludeerd dat er geen aanleiding is om te concluderen dat de achterliggende kansverdeling van de verschilvariabele de normale verdeling is. Het is dus onjuist om een gepaarde t-toets toe te passen. In deze situatie wordt de binomiale tekentoets en de Wilcoxon signed rank toets toegepast, omdat deze toetsen niet als eis hebben dat de verschilvariabele de normale verdeling volgt.

Er is een binomiale tekentoets uitgevoerd. Hiervoor wordt er een “+” toegekend aan de verschillen groter dan 0 en een “-“ aan de verschillen kleiner dan 0. De variabele “x” is het aantal “-” opgeteld. De binomiale tekentoets is uitgevoerd met de functie “binom.test”. “x” is het aantal successen en het totaal aantal verschillen is het aantal pogingen tot succes. Er is aangegeven dat het om een rechtszijdige hypothesetoets gaat, omdat het vermoeden is dat het verschil groter dan nul is. De hypothesen voor de binomiale tekentoets zijn:

$H_0$ : de kans op – is gelijk aan 0,5

$H_1$ : de kans op – is groter dan 0,5

Uit de binomiale tekentoets volgt  $p < 0,001$ . De p-waarde is kleiner dan alfa, waardoor de nulhypothese wordt verworpen. Er is geen aanleiding om aan te nemen dat de kans op “-” gelijk is aan 0,5. Er is dus geen aanleiding om aan te nemen dat er geen verschil is tussen de economisch optimale overstromingskansen van OptimaliseRing en OptimaliseRing Discreet. Er is geconcludeerd dat er een statistisch significant verschil is.

Als alternatief voor de gepaarde t-toets, wordt de Wilcoxon signed rank toets uitgevoerd. Deze toets is krachtiger dan de binomiale tekentoets. De hypothesen voor de Wilcoxon signed rank toets zijn:

$H_0$ : de mediaan van de verschillen is nul

$H_1$ : de mediaan van de verschillen is niet nul

Uit de Wilcoxon signed rank test volgt  $p < 0,001$ . De p-waarde is kleiner dan alfa, waardoor de nulhypothese wordt verworpen. De kansverdelingen zijn dus niet significant hetzelfde.

## H5.4: Regressie

Het verschil in uitkomsten van de economisch optimale overstromingskansen tussen OptimaliseRing en OptimaliseRing Discreet is mogelijk te verklaren aan de hand van de eigenschappen van de dijkringdelen. De volgende eigenschappen zijn meegenomen: schade, slachtoffers, getroffen, psi en zeta. In hoofdstuk 2.3 staat uitgelegd wat deze eigenschappen betekenen. Deze eigenschappen zijn namelijk voor beide modellen gelijk. De kosten parameters zijn niet meegenomen, omdat deze niet beschikbaar zijn op het niveau van dijkringdeel. Er is gekozen om de verschillen te verklaren aan de hand van een meervoudig lineair regressiemodel, omdat er meerdere verklarende variabelen zijn die invloed hebben op de te verklaren variabele het verschil.

Er is een meervoudig lineair regressiemodel opgesteld om te achterhalen welke van de verklarende variabelen een statistisch significante invloed hebben op de te verklaren variabele, het verschil. In het meervoudige regressiemodel worden alle eigenschappen meegenomen. Het meervoudige regressiemodel met bijbehorende waarde van de F-toets en de waarde van adjusted  $R^2$  is weergegeven in tabel 1. Adjusted  $R^2$  geeft aan in welke mate de variantie van de te verklaren variabele verklaard is door de verklarende variabelen. Een waarde van 1 wil zeggen dat de variantie van de te verklaren variabelen volledig verklaard kan worden met het regressiemodel. De F-toets heeft de volgende nulhypothese:

$H_0$ : het regressiemodel voegt niets toe ten opzichte van een regressiemodel zonder onafhankelijke variabelen

Meervoudig lineair regressiemodel	Uitkomst F-toets	Adjusted $R^2$
$-8,179 \cdot 10^{-4} + 1,933 \cdot 10^{-8} \cdot \text{schade} + 1,055 \cdot 10^{-7}$ $\cdot \text{slachtoffers} + 1,505 \cdot 10^{-9}$ $\cdot \text{getroffenen} + 9,120 \cdot 10^{-2} \cdot \text{psi} + 1,189$ $\cdot 10^{-3} \cdot \text{zeta}$	$p \approx 0,003$	0,165

Tabel 1: Meervoudige lineair regressiemodel met bijbehorende uitkomst van de F-toets en adjusted  $R^2$  waarde.

Bij meervoudige regressie kan er sprake zijn van multicollineariteit. Dit houdt in dat er een sterke samenhang tussen verschillende onafhankelijke variabelen is. Het regressiemodel wordt eerst op multicollineariteit getoetst. Met behulp van de VIF-waarde van elke onafhankelijke variabelen wordt vastgesteld of er sprake is van multicollineariteit. In tabel 2 is de VIF-waarde per onafhankelijke variabele te vinden. Een VIF-waarde tussen de 5 en 10 kan op multicollineariteit duiden als de VIF-waarde in verhouding veel groter is dan de andere VIF-waarden. In dit geval zijn er twee VIF-waarde die ongeveer tussen de 5 en 10 zijn en dit duidt dus op multicollineariteit.

Onafhankelijke variabele	VIF-waarde
Schade	5,706
Slachtoffers	1,475
Getroffenen	4,975
Psi	1,097
Zeta	1,171

Tabel 2: Onafhankelijke variabelen met bijbehorende VIF-waarde

Er zijn correlatie toetsen uitgevoerd om te kijken welke van de onafhankelijke variabelen een sterke samenhang hebben. De uitkomsten van de correlatietoetsen zijn weergegeven in tabel 3.

Variabelen	Uitkomst correlatietoets
Schade, getroffen	0,893
Schade, slachtoffers	0,484
Getroffenen, slachtoffers	0,433
Psi, zeta	-0,019
Psi, schade	-0,168
Zeta, schade	-0,265

Tabel 3: Sets van onafhankelijke variabelen met bijbehorende uitkomst van de correlatietoets

Eén van de groepen correleert sterk, namelijk de groep van schade, slachtoffers en getroffen. Er is gekozen om één van deze variabelen op te nemen in het nieuwe model, namelijk schade. Het nieuwe model met bijbehorende waarde van de F-toets en de waarde van adjusted  $R^2$  is weergegeven in tabel 4. De VIF-waarden van het nieuwe model zijn voor alle variabelen rond 1. Dit betekent dat er geen sprake meer is van multicollineariteit. Voor het meervoudige lineaire regressiemodel geldt dat de p-waarde kleiner is dan alfa en dus wordt de nulhypothese van de F-toets verworpen. Het regressiemodel is dus beter in staat het verschil in overstromingskans te verklaren dan een regressiemodel zonder onafhankelijke variabelen,  $F(3,70)=6,218$ ,  $p<0,001$ .

Meervoudig lineair regressiemodel	Uitkomst F-toets	Adjusted $R^2$
$-8,309 \cdot 10^{-4} + 4,610 \cdot 10^{-8} \cdot schade + 9,914 \cdot 10^{-2} \cdot psi + 7,812 \cdot 10^{-3} \cdot zeta$	$p \approx 0,001$	0,175

Tabel 4: Meervoudige lineair regressiemodel met bijbehorende uitkomst van de F-toets en adjusted  $R^2$  waarde.

Voordat het meervoudige lineaire regressiemodel gebruikt mag worden, zal deze eerst moeten worden gevalideerd. Het valideren gebeurt in drie stappen op dezelfde wijze als bij enkelvoudige lineaire regressie.

De eerste stap is het testen of er aan de voorwaarde dat de residuen normaal verdeeld zijn wordt voldaan. Dit gebeurt in twee stappen. Als eerst wordt de Shapiro-Wilk toets uitgevoerd. De hypothesen van deze toets zijn:

$H_0$ : de residuen volgen als achterliggende kansverdeling de normale verdeling

$H_1$ : de residuen volgen niet als achterliggende kansverdeling de normale verdeling

De p-waarde is kleiner dan alfa, dus de nulhypothese wordt verworpen. Er is dus geen statistisch significante reden om aan te nemen dat de residuen normaal verdeeld zijn,  $p<0,001$ . De tweede stap is het maken van een Q-Q plot. In Bijlage 4 figuur 1 is de Q-Q plot te vinden. Een groot deel van de uitkomsten ligt buiten het acceptatiegebied en de uitkomsten liggen niet in een rechte lijn. Samen met de conclusie van de Shapiro-Wilk toets lijkt het redelijk om aan te nemen dat de residuen van het model niet normaal verdeeld zijn. Er wordt dus niet aan de voorwaarde van normaliteit voldaan.

Als tweede stap wordt er getest of de residuen bij benadering nul zijn met behulp van een t-toets. De hypothesen voor deze toets zijn:

$H_0$ : het gemiddelde van de residuen is gelijk aan 0

$H_1$ : het gemiddelde van de residuen is niet gelijk aan 0

De p-waarde is groter dan alfa, dus de nulhypothese wordt behouden. Er is genoeg aanleiding om aan te nemen dat de residuen gemiddeld nul zijn,  $p=1$ . Er wordt dus voldaan aan de eis dat de residuen bij benadering gemiddeld nul zijn.

Als laatste stap wordt er getest of de residuen homoscedastisch zijn. Hiervoor worden de residuen geplot. De plot is te vinden in Bijlage 4 figuur 2. Er lijkt geen duidelijk patroon te zien, dit wijst op homoscedasticiteit. Om deze uitkomst te versterken had er een Breusch-Pagan toets toegepast kunnen worden. Dit kan in dit geval niet, omdat de residuen niet de normale verdeling volgen en dit een eis is om de Breusch-Pagan toets toe te mogen passen.

Aan de voorwaarden dat de residuen bij benadering nul zijn en dat de residuen homoscedastisch zijn is voldaan. Ook is aan de voorwaarde voldaan dat er geen sprake is van multicollineariteit. Echter is niet aan de voorwaarde van normaliteit voldaan. Omdat het regressiemodel niet aan alle eisen voldoet is geconcludeerd dat het model niet valide is voor het verklaren van het verschil in overstromingskans.

Om toch de verschillen in overstromingskans te kunnen verklaren, wordt er robuuste meervoudige regressie toegepast. Het robuuste meervoudige regressiemodel is weergegeven in tabel 5. Schade, slachtoffers en getroffen en zijn in dit model wel meegenomen, alleen zij gaven de waarde 0, waardoor deze niet in de tabel zijn opgenomen. Met behulp van meervoudige robuuste regressie is er een relatie gevonden tussen het verschil in overstromingskans en psi en tussen het verschil in overstromingskans en zeta. Er is een verhoging van 0,0796% van de overstromingskans voor elke 1% verhoging van psi. Er is een verlaging van 0,0279% van de overstromingskans voor elke 1% verhoging van zeta.

Meervoudig robuust regressiemodel
$-0,0006 + 0,0796 \cdot psi - 0,0279 \cdot zeta$

Tabel 5: Meervoudig robuust regressiemodel

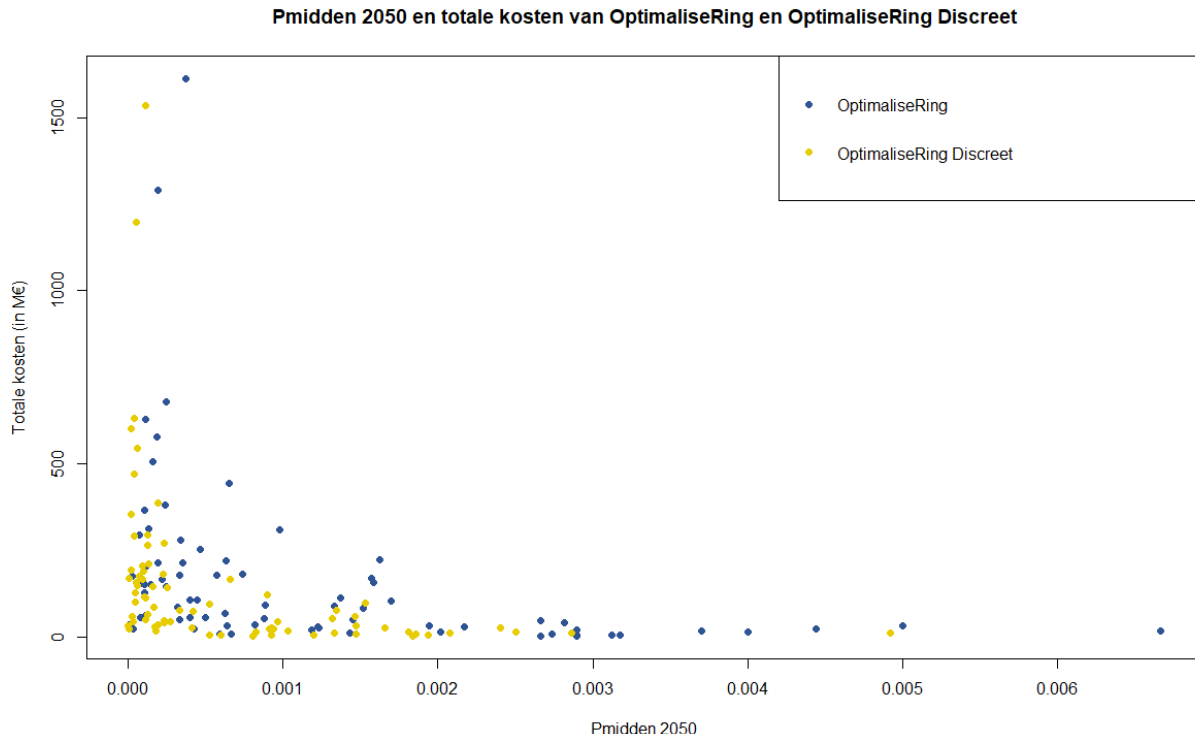
## H6: Resultaten testen aan kwaliteitscriteria

In hoofdstuk 3 zijn de kwaliteitscriteria opgesteld en is omschreven hoe deze getest kunnen worden. In dit hoofdstuk worden per kwaliteitscriterium categorie de uitkomsten van de testen gegeven.

### H6.1: Juistheid

Bij de categorie juistheid is gekeken naar twee criteria. Voor alle dijkkringdelen is de economisch optimale overstromingskans berekend met OptimaliseRing Discreet. Vervolgens is per dijkkringdeel het procentuele verschil berekend, zoals omschreven in hoofdstuk 5.2, waaruit het gemiddelde procentuele verschil berekend is. De aanname is dat de uitkomsten van OptimaliseRing correct zijn. De uitkomsten van de economisch optimale overstromingskans in 2050 liggen gemiddeld genomen in OptimaliseRing Discreet ongeveer 50% lager dan in OptimaliseRing.

Voor alle dijkkringdelen zijn ook de totale kosten bepaald met OptimaliseRing Discreet. Per dijkkringdeel is het procentuele verschil berekend, waaruit het gemiddelde procentuele verschil is berekend. De totale kosten liggen gemiddeld genomen in OptimaliseRing Discreet 19% lager dan in OptimaliseRing. Verder is er een spreidingsdiagram (zie figuur 11) gemaakt waarin de uitkomsten van OptimaliseRing en OptimaliseRing Discreet geplott zijn, met op de x-as de overstromingskans en op de y-as de totale kosten. In figuur 11 is te zien dat OptimaliseRing Discreet gemiddeld genomen lagere kansen en lagere kosten heeft dan OptimaliseRing.



Figuur 11: Spreidingsdiagram Pmidden 2050 en totale kosten (in M€) van OptimaliseRing en OptimaliseRing Discreet

## H6.2: Verklaarbaarheid

Met behulp van de binomiale tekentoets en de Wilcoxon signed rank toets is in hoofdstuk 5.3 geconcludeerd dat er een statistisch significant verschil is tussen de economisch optimale overstromingskansen van OptimaliseRing en OptimaliseRing Discreet,  $p < 0,001$ . Er is een meervoudig regressiemodel opgesteld om de verschillen te verklaren. Dit model voldeed niet aan de eis van normaliteit en is daardoor niet valide. Dit model kon dus niet worden gebruikt voor het verklaren van de verschillen, daarom is er als alternatief een robuust meervoudig regressiemodel opgesteld. Met behulp van meervoudige robuuste regressie is er een relatie gevonden tussen het verschil in overstromingskans en psi en tussen het verschil in overstromingskans en zeta. Er is een verhoging van 0,0796% van de overstromingskans voor elke 1% verhoging van psi. Er is een verlaging van 0,0279% van de overstromingskans voor elke 1% verhoging van zeta.

## H6.3: Stabiliteit

Om te testen hoe stabiel het programma is kunnen de parameters van het model één voor één veranderd worden. Omwille van opslagcapaciteit en tijd is deze test niet meer uitgevoerd en is er dus geen resultaat voor dit kwaliteitscriterium. Deze test zou in een vervolgonderzoek wel meegenomen kunnen worden.

## H6.4: Aanpasbaarheid

Aanpasbaarheid is onderverdeeld in drie criteria. Het eerste criteria gaat over hoe de input data eruit ziet. De data die ingevoerd wordt in OptimaliseRing zit in een Access database. Uit deze database worden 14 tabellen gebruikt. In het document technische documentatie OptimaliseRing (Duits, 2011b, pp. 19-37) wordt de database uitgebreid omschreven en op basis van deze informatie kan de database worden aangepast. De data van OptimaliseRing Discreet bestaat uit 11 CSV en 2 INC bestanden, die ingevoerd worden in het Python script “aanmaken\_variabelen”. In dit script worden er vanuit de CSV data vier variabelen gemaakt. Deze variabelen worden vervolgens als input gegeven voor OptimaliseRing Discreet. Als men de data aan wil passen, dan kan dit door de CSV bestanden aan te passen, wel moet dit consistent gebeuren. Zo moeten er in alle bestanden evenveel dijkringdelen staan. Vervolgens kan met de nieuwe CSV bestanden het programma “aanmaken\_variabelen” opnieuw worden uitgevoerd en kunnen de nieuwe variabelen worden meegegeven aan OptimaliseRing Discreet. Volgens het rapport documentatie nabootsen optimalisering opgestuurd door stagebegeleider (A. Wooning, persoonlijke communicatie, 7 februari 2022) is een voordeel van OptimaliseRing Discreet dat er bijna volledige flexibiliteit met betrekking tot de input data is. Bij beide modellen kan de input data aangepast worden. Bij OptimaliseRing Discreet is er bijna volledige flexibiliteit met betrekking tot de input data, echter kan de documentatie over de input data nog worden uitgebreid.

Het tweede criterium gaat over de kwaliteit van de code. OptimaliseRing is niet voorzien van commentaar, echter zijn de variabelen namen erg duidelijk, waardoor de code te lezen is. Er zit weinig ongebruikte code in OptimaliseRing. Consistentie in de code is zichtbaar, door het consistent gebruiken van hoofdletters en spaties. Verder wordt de scope van de variabelen klein gehouden. OptimaliseRing discreet is voorzien van een klein beetje commentaar. Echter zijn dit korte zinnen die niet altijd erg duidelijk zijn. Ook staan er soms wat ingewikkeldere stukken code zonder enig commentaar in. De variabelen namen zijn soms duidelijk, bijvoorbeeld “cost” en “mdl”.



Echter is een groot deel van de variabelen namen onduidelijk, omdat ze uit één of twee letters bestaan. Er zitten ongeveer 100 regels ongebruikte code in OptimaliseRing Discreet. Consistentie is in de code zichtbaar door het gebruik van enters en doordat het commentaar bijvoorbeeld met een kleine letter begint. Verder wordt de scope van variabelen in functies klein gehouden. In OptimaliseRing discreet kunnen de aspecten duidelijke variabelen namen en commentaar nog verbeterd worden. In OptimaliseRing kan er commentaar toegevoegd worden.

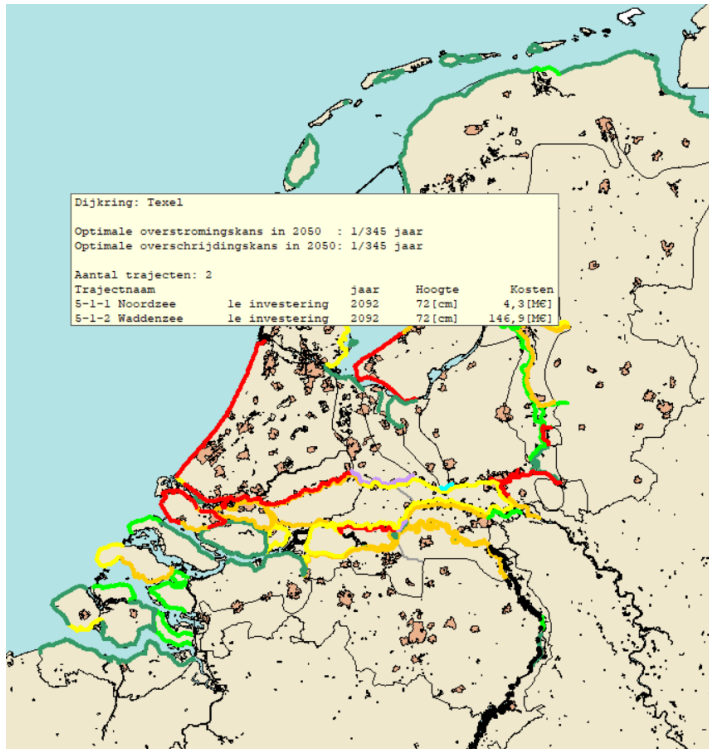
Het derde criterium bij aanpasbaarheid gaat over in welke taal en met welke pakketten de programma's opgebouwd zijn. OptimaliseRing is gebouwd in AIMMS en heeft als beperking dat het programma op één computer niet meer dan één keer tegelijk gestart kan worden. Om OptimaliseRing te kunnen draaien is het programma AIMMS nodig van versie 3.8.5.11246 met bijbehorende licentie en Microsoft .NET Framework 2.0. Verder zijn de volgende packages nodig: dotnetfx en vcredist\_x86 (Duits, 2011a, p. 5). Volgens het rapport kennisoverdracht plan FD opgestuurd door stagebegeleider (A. Wooning, persoonlijke communicatie, 7 februari 2022) is de taal niet erg handig in de voor- en nabewerking van de data. Volgens datzelfde rapport heeft AIMMS een heel veel kleinere gebruikersgroep dan Python. OptimaliseRing Discreet is gebouwd in Python. Door de code door te nemen, is gevonden welke packages er in OptimaliseRing Discreet worden gebruikt. De volgende packages worden gebruikt: cplex, numpy, matplotlib, scipy, pyodbc en csv. Dit zijn standaard packages waar op internet aardig wat documentatie over te vinden is. Het voordeel van Python is dat het open source is. De voordelen van open source zijn dat het flexibel is, dat het openbaar is en gratis is. Volgens onderzoek (Statista, 2022) gebruikt 48,24% van de ontwikkelaars wereldwijd Python in het jaar 2021. Cijfers over welk percentage of hoeveel mensen kennis bezitten over AIMMS zijn niet gevonden. OptimaliseRing Discreet heeft meer voordelen in dit criterium.

#### H6.5: Gebruiksvriendelijkheid

Er zijn drie criteria die horen bij de categorie gebruiksvriendelijkheid. Het eerste criterium gaat over de output van de modellen. OptimaliseRing geeft als uitkomst een kaart van Nederland met daarop de dijkkringdelen waarvoor de economisch optimale overstromingskansen berekend zijn (zie figuur 12). In zie figuur 12 staat de muis op Texel, er verschijnt dan een informatievenster die de economisch optimale overstromingskansen weergeeft. Ook geeft dit venster per dijktraject in het dijkkringdeel aan wanneer er voor het eerst geïnvesteerd wordt en met hoeveel het traject verhoogd gaat worden en wat dit gaat kosten. Wat OptimaliseRing nog meer teruggeeft is een uitvoerbestand waar de gegevens in staan die gebruikt zijn voor de berekeningen. Ook staan in dit bestand de uitkomsten van de berekeningen. Dat wil zeggen, per investering het jaartal, de verhoging en wat dit gaat kosten, verder de optimale overstromingskans in 2020 en 2050 voor het gehele dijkkringdeel en de totale investeringskosten. Naast het uitvoerbestand levert OptimaliseRing per dijkkringdeel een grafiek waarin het verloop van de investeringen te zien is (zie figuur 4 in hoofdstuk 2.3). OptimaliseRing Discreet berekende eerst alleen de optimale investeringsstrategie en nog niet de economisch optimale overstromingskansen, de berekening voor dit laatste is pas in deze stage geïmplementeerd. Als output worden de totale investeringskosten per dijkkringdeel gegeven en een grafiek per dijkkringdeel met de investeringsstrategie (zie figuur 6 in hoofdstuk 2.4). Hierin is te zien welk dijktraject het zwakste is gedurende de tijd, namelijk de gekleurde stippen. Ook zijn de tijdstippen waarop geïnvesteerd wordt af te lezen en naar welk



beschermingsniveau er wordt gegaan. Ook wordt er nu per dijkkringdeel een grafiek gegeven van het verloop van de overstromingskans en economisch optimale overstromingskans (zie figuur 8 in hoofdstuk 4.3). Beide modellen geven als uitkomst de overstromingskans en een grafiek en tabel waarin af te lezen is wat de optimale tijdstippen zijn om te verhogen en naar welk niveau er wordt verhoogd. Ook geven beide modellen de investeringsstrategie weer in een grafiek.



Figuur 12: Output OptimaliseRing: kaart van Nederland met de economisch optimale overstromingskansen voor alle dijkkringdelen. De muis staat op het dijkkringdeel Texel, om te laten zien wat voor informatie hierover gegeven wordt.

Het tweede criterium gaat over hoe toegankelijk de programma's zijn voor mensen zonder programmeerkennis. Het installeren van beide programma's vraagt algemene computerkennis. Hoe de installatie van OptimaliseRing verloopt staat duidelijk beschreven in de gebruikershandleiding (Duits, 2011a). Door de gebruikersschil die gebouwd is om OptimaliseRing heen, is het programma te gebruiken door mensen zonder programmeerkennis. Door achtereenvolgens op de knoppen "Rekenen" en "Berekeningen" te klikken, komt de gebruiker in een scherm waarin de gewenste dijkringen geselecteerd kunnen worden. Ook kunnen de berekeningsinstellingen hier worden aangepast. Wanneer de gebruiker dijkringdelen heeft geselecteerd, worden de berekeningen uitgevoerd en verschijnt er een kaart van Nederland met daarop de dijkringen met overstromingskansen. Ook zijn via "Rekenen" en vervolgens "Resultaten" de resultaten van eerdere berekeningen te zien en per dijkkring een grafiek van de optimale investeringsstrategie. OptimaliseRing Discreet heeft geen gebruikershandleiding en er is ook geen documentatie over de installatie beschikbaar, waardoor het voor de gebruiker een stuk lastiger is om het programma te installeren. Ook moet er een Integrated Development Environment (IDE) geïnstalleerd worden waarin de Python code van OptimaliseRing Discreet gedraaid kan worden, er is dan ook enige kennis over de gebruikte IDE vereist. Het programma heeft geen gebruikersschil en de gebruiker ziet dus direct de code van OptimaliseRing Discreet. Ook staat er niet in de documentatie waar de resultaten

worden opgeslagen, dit is wel uit de code af te lezen. Kortom, OptimaliseRing is toegankelijker voor gebruikers zonder programmeerkennis.

Het derde criterium gaat over de user interface en documentatie van de programma's. OptimaliseRing is zoals gezegd een programma met een gebruikersschil erom heen. Er is documentatie over OptimaliseRing beschikbaar, zoals het rapport MKBA WV21 (Kind, 2011), een technische documentatie (Duits, 2011b) en een gebruikershandleiding (Duits, 2011a). Al deze documentatie is erg uitgebreid. In de gebruikershandleiding wordt met behulp van afbeeldingen stap voor stap uitgelegd hoe de installatie eraan toegaat en hoe men het programma kan gebruiken om berekeningen te maken. Ook worden de instellingen van de berekening besproken en uitgelegd hoe deze kunnen worden aangepast. Verder wordt er uitgelegd welke resultaten het programma produceert en waar deze resultaten te vinden zijn. Over het model OptimaliseRing Discreet is minder informatie beschikbaar. Ook zit hier geen gebruikersschil omheen. In het rapport Implementatie van Optimalisering-discreet en Dique-Opt (P. de Grave, persoonlijke communicatie, 14 maart 2022) wordt het optimaliseringsmodel uitgelegd. Verder is het rapport documentatie nabootsing optimalisering en het rapport testverslag optimalisering discreet toegestuurd door stagebegeleider (A. Wooning, persoonlijke communicatie, 7 februari 2022). Hierin is het optimaliseringsmodel opnieuw kort uitgelegd en wordt ook beschreven hoe de resultaten eruit zien. Wat ontbreekt is een uitleg hoe het programma OptimaliseRing Discreet te installeren en te gebruiken. Concluderend is OptimaliseRing gebruiksvriendelijker in de user interface en documentatie.

#### H6.6: Reproduceerbaarheid

Het criteria dat hoort bij reproduceerbaarheid is of de programma's dezelfde uitkomsten geven als deze meerdere keren worden gedraaid. De programma's OptimaliseRing en OptimaliseRing Discreet zijn beide twee keer gedraaid en gaven de eerste keer dezelfde uitkomsten als de tweede keer.

#### H6.7: Rekeningtijd

Er zijn twee criteria toegekend aan rekeningtijd. Het eerste criterium gaat over hoelang het duurt om voor alle dijkkringdelen de economisch optimale overstromingskans te bepalen en bijbehorende output te genereren. De tijd dat OptimaliseRing hierover heeft gedaan is ongeveer 2 uur en 4 minuten. De tijd dat OptimaliseRing Discreet heeft gedaan over het aanmaken van de input variabelen die nodig zijn om de economisch optimale overstromingskansen te berekenen is ongeveer 6 uur en 26 minuten. Het daadwerkelijk berekenen van de economisch optimale overstromingskansen en bijbehorende output duurt ongeveer 15 minuten. Dit maakt dat de totale tijd die nodig is voor de berekening van de economisch optimale overstromingskansen bij OptimaliseRing minder is.

Het tweede kwaliteitscriterium kan getest worden door het programma OptimaliseRing Discreet af te laten kappen na een bepaald aantal minuten. Omwille van tijd is er niet meer onderzocht of dit mogelijk is en zo ja hoe dit uitgevoerd kan worden. Deze test zou in een vervolgonderzoek wel meegenomen kunnen worden.

## H6.8: Kosten

Kosten is niet meegenomen als kwaliteitscriterium, maar de bevindingen hierover worden voor de volledigheid toch omschreven. Dit is gedaan voor het geval dat een andere organisatie de programma's OptimaliseRing of OptimaliseRing Discreet zou willen gebruiken. OptimaliseRing is gebouwd in AIMMS. Om OptimaliseRing te kunnen draaien is het programma AIMMS nodig van versie 3.8.5.11246 met bijbehorende licentie (hier zitten kosten aan verbonden) en Microsoft .NET Framework 2.0. Verder zijn de volgende packages nodig: dotnetfx en vcredist\_x86. Het voordeel van AIMMS is dat het komt met de CPLEX optimalisering solver van IMB. Deze solver is vele malen sneller dan de gratis alternatieven en wordt daarom ook gebruikt in OptimaliseRing Discreet (A. Wooning, persoonlijke communicatie, 7 februari 2022). Om de solver in OptimaliseRing Discreet te kunnen gebruiken is een Python versie van de solver gekocht. OptimaliseRing Discreet is gebouwd in Python. In OptimaliseRing Discreet worden de volgende packages gebruikt: cplex, numpy, matplotlib, scipy, pyodbc en csv. Op cplex na zijn deze packages gratis te downloaden. Python is een open source taal en kan ook gratis worden gebruikt.

## H7: Conclusie, aanbevelingen en discussie

In hoofdstuk 7.1 wordt er een antwoord gegeven op de hoofdvraag en worden de voor- en nadelen van OptimaliseRing en OptimaliseRing Discreet beschreven. In hoofdstuk 7.2 worden aanbevelingen gegeven. In hoofdstuk 7.3 staat een discussie over welke aspecten mogelijk veranderd of verbeterd hadden kunnen worden in het onderzoek.

### H7.1: Conclusie

Er is onderzoek gedaan naar de volgende hoofdvraag: *Wat zijn de voor- en nadelen van het gebruik van OptimaliseRing en OptimaliseRing Discreet voor de toekomstige berekening van de economisch optimale overstromingskansen?* Om deze hoofdvraag te kunnen beantwoorden zijn er deelvragen opgesteld. Als eerst is onderzocht hoe de modellen OptimaliseRing en OptimaliseRing Discreet in elkaar zitten. Daarna is het model OptimaliseRing Discreet aangevuld, zodat het de economisch optimale overstromingskansen berekend. Vervolgens zijn er kwaliteitscriteria opgesteld met bijbehorende testen en zijn de testen uitgevoerd. Een onderdeel hiervan was een statistische analyse uitvoeren. Uit de statistische analyse volgt dat er een statistisch significant verschil in economisch optimale overstromingskansen is tussen OptimaliseRing en OptimaliseRing Discreet. Met behulp van meervoudige robuuste regressie is er een relatie gevonden tussen het verschil in overstromingskans en  $\psi$  en tussen het verschil in overstromingskans en  $\zeta$ . Dit zijn twee belangrijke parameters ten aanzien van klimaatverandering en potentiële schade. Twee testen zijn niet uitgevoerd door gebrek aan opslagcapaciteit en tijd. Uiteindelijk zijn de volgende voor- en nadelen uit de resultaten van de testen gekomen:

De voordelen van OptimaliseRing zijn:

- De uitkomsten van het model zijn correct.
- Er is uitgebreide informatie beschikbaar zoals een gebruikershandleiding en technische documentatie.
- In de code achter OptimaliseRing worden duidelijke variabelen namen gebruikt, is weinig ongebruikte code te vinden, wordt er consistent gebruik gemaakt van hoofdletters, spaties en enters en wordt de scope van de variabelen klein gehouden.
- Het programma heeft een gebruikersschil. Dit maakt dat het programma ook toegankelijk is voor mensen die niet over programmeerkennis beschikken.
- Het programma heeft een kortere rekentijd dan OptimaliseRing Discreet.

De nadelen van OptimaliseRing zijn:

- Er staat vrijwel geen commentaar in de code achter OptimaliseRing.
- OptimaliseRing is gebouwd in AIMMS, waar een licentie voor nodig is.

De voordelen van OptimaliseRing Discreet zijn:

- Het model biedt bijna volledige flexibiliteit met betrekking tot de invoer data.
- In de code achter OptimaliseRing Discreet wordt consistent gebruik gemaakt van hoofdletters, spaties en enters en wordt de scope van de variabelen klein gehouden.

- OptimaliseRing Discreet is gebouwd in Python, wat open source is. De voordelen van open source zijn dat het flexibel is, dat het openbaar is en gratis is.

De nadelen van OptimaliseRing Discreet zijn:

- De uitkomsten van het model wijken sterk af van de uitkomsten van OptimaliseRing. Zo is de economisch optimale overstromingskans gemiddeld 50% lager en de totale kosten 19% lager dan in OptimaliseRing. Met verschiltoetsen is aangetoond dat er een statistisch significant verschil is tussen de economisch optimale overstromingskansen in OptimaliseRing en OptimaliseRing Discreet. Er kan worden geconcludeerd dat er een systematische fout in het model zit, waardoor de uitkomsten onjuist zijn.
- Er is weinig documentatie beschikbaar over hoe het model te gebruiken en de invoer data. Dit maakt het programma minder toegankelijk.
- In de code achter OptimaliseRing Discreet worden soms onduidelijke variabelen namen gebruikt. Ook kan het commentaar uitgebreider en kan ongebruikte code worden weggehaald.
- Het programma heeft geen gebruikersschil erom heen. Dit maakt het programma minder gebruiksvriendelijk.
- Het programma heeft een langere rekentijd dan OptimaliseRing.

Kortom, OptimaliseRing heeft meer voordelen ten opzichte van OptimaliseRing Discreet. OptimaliseRing Discreet heeft meer nadelen dan OptimaliseRing. Ook weegt de juistheid van de uitkomsten zwaar en zou OptimaliseRing Discreet eerst verbeterd moeten worden om dit programma in de toekomst te kunnen gebruiken voor de berekening van de economisch optimale overstromingskansen.

## H7.2: Aanbevelingen

Uit de conclusie volgen een aantal aanbevelingen, namelijk:

Twee testen zijn niet uitgevoerd door gebrek aan opslagcapaciteit en tijd. Er wordt aanbevolen om de twee testen uit te voeren, om een nog completer beeld te krijgen van de voor- en nadelen van de programma's. Het gaat om de test hoe stabiel de programma's zijn en de test of OptimaliseRing discreet afgekapt kan worden na een bepaalde tijd en wat voor effect dit heeft.

Het programma OptimaliseRing Discreet geeft uitkomsten die te veel afwijken van de uitkomsten van OptimaliseRing. Het programma zou nog een keer nagekeken kunnen worden, om de fout in het programma op te sporen en op te lossen. Vervolgens zou de statistische analyse opnieuw uitgevoerd kunnen worden, om te kijken of het verschil tussen de modellen nog significant is en zo ja welke factoren er invloed hebben op dit verschil.

Het programma OptimaliseRing Discreet zou verbeterd kunnen worden, door meer commentaar toe te voegen en duidelijke variabelen namen te gebruiken. Verder zou een gebruikersschil om dit programma heen, het programma toegankelijker maken voor mensen die niet beschikken over programmeerkennis. Wat het programma ook toegankelijker zou maken is documentatie over hoe het programma te gebruiken. Als laatste kan er gekeken worden of het aanmaken variabelen op een efficiëntere manier kan gebeuren, om de rekentijd in te perken.

### H7.3: Discussie

In deze paragraaf wordt besproken welke aspecten in het onderzoek veranderd of verbeterd hadden kunnen worden.

Ten eerste was de uitkomst van de statistische analyse deels voorspelbaar, omdat de uitkomsten van OptimaliseRing en OptimaliseRing Discreet sterk afweken. Het was dus geen verrassing toen uit de statistische analyse kwam dat er een significant verschil was tussen de overstromingskansen van de modellen. Hierdoor leek dit deel van de analyse wellicht minder zinvol. Echter is aanbevolen om de statistische analyse nogmaals uit te voeren, als het model OptimaliseRing Discreet uitkomsten geeft die dichterbij de uitkomsten van OptimaliseRing liggen. Rijkswaterstaat kan dan deze analyse als voorbeeld gebruiken en grotendeels dezelfde stappen en dezelfde code gebruiken om de analyse opnieuw uit te voeren.

Ten tweede zijn in het regressiemodel geen kostenparameters meegenomen, omdat deze niet beschikbaar waren op het niveau van dijkkringdeel. Dit maakt het regressiemodel minder volledig, omdat deze parameters wel een invloed kunnen hebben op de economisch optimale overstromingskansen.

Als de discontovoet wordt verlaagd, duren de berekeningen van de economisch optimale overstromingskansen in ieder geval in OptimaliseRing een stuk langer. Naarmate de discontovoet lager is, dragen investeringen en schades in de verre toekomst verhoudingsgewijs meer bij. Bij een hogere discontovoet zou het optimalisatie algoritme de oplossing eerder accepteren dan bij een lagere discontovoet. De discontovoet heeft dus invloed op de rekentijd van OptimaliseRing. Met een andere discontovoet had OptimaliseRing wellicht een langere rekentijd gehad. Het verband tussen de discontovoet en de rekentijd van de modellen zou onderzocht kunnen worden, om een completer beeld te geven van de rekentijd.

In de aanbevelingen is genoemd om de test over het afkappen van OptimaliseRing Discreet nog uit te voeren. Aangezien het berekenen van de economisch optimale overstromingskansen maar ongeveer 15 minuten duurt, kan erover gediscussieerd worden of dit wel zinvol is. Als door de discontovoet te verlagen de berekeningen van OptimaliseRing Discreet ook langer duren, is het wel weer zinvol om naar dit criterium te kijken.

Als er werd gekeken of bepaalde groepen normaal verdeeld waren in de statistische analyse is er eerst een Q-Q plot gemaakt en daarna de Shapiro-Wilk toets uitgevoerd. Hierna is geen Mauchly's toets uitgevoerd. Dit heeft waarschijnlijk geen invloed gehad op de uitkomst, want de Q-Q plot en Shapiro-Wilk toets waren overtuigend. De Mauchly's toets zou de uitkomsten wel betrouwbaarder hebben gemaakt.



## Literatuurlijst

*Assessing the Assumption of Normality · UC Business Analytics R Programming Guide*. (z.d.). Github. Geraadpleegd op 23 juni 2022, van [https://uc-r.github.io/assumptions\\_normality](https://uc-r.github.io/assumptions_normality)

Bevans, R. (2022, 1 juni). *Multiple Linear Regression | A Quick Guide (Examples)*. Scribbr. Geraadpleegd op 18 juli 2022, van <https://www.scribbr.com/statistics/multiple-linear-regression/>

Buijs, A. (2017). *Statistiek om mee te werken* (10de editie). Noordhoff Uitgevers.

De Joode, C. (2016). *Dictaat numerieke wiskunde* (Versie 6.3 ed.) [dictaat]. Blackboard. Geraadpleegd op 7 juni 2022

Duits, M. T., Kind, J., & Bak, C. (2011, maart). *Maatschappelijke kosten-batenanalyse Waterveiligheid 21e eeuw Bijlage C: Wiskundige achtergrond MKBA model*. Deltares. [https://puc.overheid.nl/rijkswaterstaat/doc/PUC\\_139011\\_31/](https://puc.overheid.nl/rijkswaterstaat/doc/PUC_139011_31/)

Duits, M. T. (2011a, april). *OptimaliseRing Gebruikershandleiding van een numeriek rekenmodel voor de economische optimalisatie van veiligheidsniveaus van dijkringen*. HKV.

Duits, M. T. (2011b, april). *OptimaliseRing Technische documentatie van een numeriek rekenmodel voor de economische optimalisatie van veiligheidsniveaus van dijkringen*. HKV.

Glen, S. (2020, 16 december). *Variance Inflation Factor*. Statistics How To. Geraadpleegd op 18 juli 2022, van <https://www.statisticshowto.com/variance-inflation-factor/>

Kind, J. (2011, maart). *Maatschappelijke kosten-batenanalyse Waterveiligheid 21e eeuw*. Deltares. [https://puc.overheid.nl/rijkswaterstaat/doc/PUC\\_139011\\_31/](https://puc.overheid.nl/rijkswaterstaat/doc/PUC_139011_31/)

King, A. P., & Eckersley, R. J. (2019). Inferential Statistics III: Nonparametric Hypothesis Testing. *Statistics for Biomedical Engineers and Scientists*, 119–145. <https://doi.org/10.1016/b978-0-08-102939-8.00015-3>

Ministerie van Infrastructuur en Waterstaat. (z.d.). *De Deltawerken*. Rijkswaterstaat. Geraadpleegd op 26 april 2022, van <https://www.rijkswaterstaat.nl/water/waterbeheer/bescherming-tegen-het-water/waterkeringen/deltawerken>

Ministerie van Infrastructuur en Waterstaat. (2022, 22 februari). *Discontovoet*. RWSeconomie.nl. Geraadpleegd op 1 juni 2022, van <https://www.rwseconomie.nl/discontovoet#:~:text=De%20discontovoet%20is%20een%20percentage,het%20basisjaar%20van%20het%20project.>

Quené, H. & Van den Bergh, H. (2022). *Kwantitatieve Methoden en Statistiek*. Geraadpleegd op 23 juni 2022, van <https://hugoquene.github.io/KMS-NL/> .

*Robust Regression | R Data Analysis Examples.* (z.d.). UCLA Advanced Research Computing Statistical Methods and Data Analytics. Geraadpleegd op 18 juli 2022, van <https://stats.oarc.ucla.edu/r/dae/robust-regression/>

Scheff, S. W. (2016). Nonparametric Statistics. *Fundamental Statistical Principles for the Neurobiologist*, 157–182. <https://doi.org/10.1016/b978-0-12-804753-8.00008-7>  
<https://www.sciencedirect.com/topics/medicine-and-dentistry/wilcoxon-signed-ranks-test>

Statista. (2022, 23 februari). Most widely utilized programming languages among developers worldwide 2021. Geraadpleegd op 13 april 2022, van <https://www.statista.com/statistics/793628/worldwide-developer-survey-most-used-languages/>

*The Breusch-Pagan Test: Definition & Example.* (2022, 25 januari). Statology. Geraadpleegd op 18 juli 2022, van <https://www.statology.org/breusch-pagan-test/>

*Understanding the t-Test in Linear Regression.* (2021, 4 oktober). Statology. Geraadpleegd op 24 juni 2022, van <https://www.statology.org/t-test-linear-regression/>

Xia, Y. (2020). Correlation and association analyses in microbiome study integrating multiomics in health and disease. *Progress in Molecular Biology and Translational Science*, 171, 309–491. <https://doi.org/10.1016/bs.pmbts.2020.04.003>  
<https://www.sciencedirect.com/topics/medicine-and-dentistry/wilcoxon-signed-ranks-test>

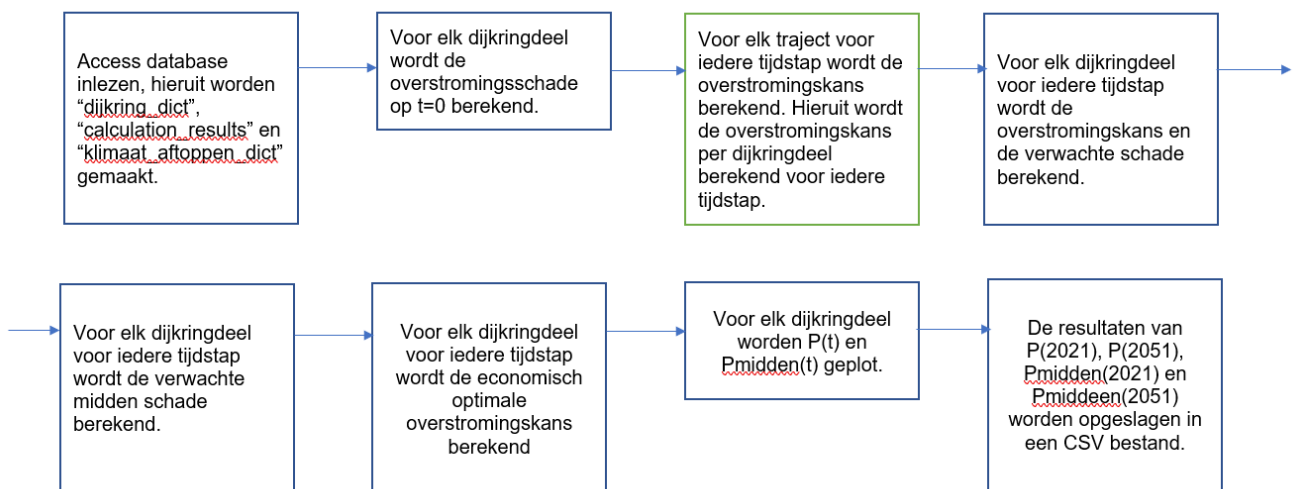
Zwaneveld, P., Verweij, G., & Van Hoesel, S. (2018). *Safe dike heights at minimal costs: An integer programming approach.* *European Journal of Operational Research*, 270(1), 294–301. <https://doi.org/10.1016/j.ejor.2018.03.012>  
<https://www.cpb.nl/sites/default/files/publicaties/download/cpb-discussion-paper-277-safe-dike-heights-minimal-costs-integer-programming-approach.pdf>



# Bijlagen

## Bijlage 1: Schematische weergave code

1. Een deel van de Access database is ingelezen als een pandas dataframe, dit dataframe wordt opgeslagen onder de naam "data".
2. De incorrecte dijkringen worden verwijderd uit "data".
3. Het oude CSV bestand is ingelezen als pandas dataframe, dit dataframe wordt opgeslagen onder de naam "df" en in de laatste kolom worden alle waarden op "NA" gezet.
4. Per rij uit "data" is opgezocht waar de dijkkring, het dijkkringdeel en eventueel het dijkkringtraject overeen komen met "df". Op deze plaats wordt in de laatste kolom de bijbehorende informatie uit "data" gezet. Bijvoorbeeld de overstromingskans op tijdstip 0.
5. Er is een check uitgevoerd of het aantal rijen dat gevuld is in "df", dat wil zeggen het aantal rijen dat geen "NA" waarde heeft, overeenkomt met het aantal rijen uit "data". Als dit het geval is, ga naar stap 6. Als dit niet het geval is, ga naar stap 7.
6. "df" wordt weggeschreven naar een CSV. Sla stap 7 over.
7. Er wordt een foutmelding op het scherm getoond.



Figuur 1: Schematische weergave van het programma "Berekenen\_economisch\_optimale\_overstromingskansen.py".



Figuur 2: Schematische weergave van het berekenen van de overstromingskansen in het programma "Berekening\_economisch\_optimale\_overstromingskansen.py"

## Bijlage 2: Dictionary's

Dijkkring\_dict:

Key	0	1	2	3	4
Vb.: "10;1"	Schade	Slachtoffers	Getroffenen	Psi	Zeta

Tabel 1: Schematische weergave van "dijkkring\_dict" in het programma Berekening economisch optimale overstromingskansen.

Klimaat\_aftoppen\_dict:

Key	0	1	2	3	4
Vb.: "10;1;1"	P(0)	Alpha	Eta	Hoogte	Lijst met P(t) per jaar

Tabel 2: Schematische weergave van "klimaat\_aftoppen\_dict" in het programma Berekening economisch optimale overstromingskansen.

Calculation\_results:

Key	0	1
Vb.: "10;1"	V(0)	Dictionary

Tabel 3: Schematische weergave van "calculation\_results" in het programma Berekening economisch optimale overstromingskansen.

Key	0	1	2	3	4	5	6
Vb.: 2016	P(t)	Lijst met P(t) per traject	Eta	V(t)	S(t)	Smidden(t)	Pmidden(t)

### Bijlage 3: Resultaten OptimaliseRing Discreet

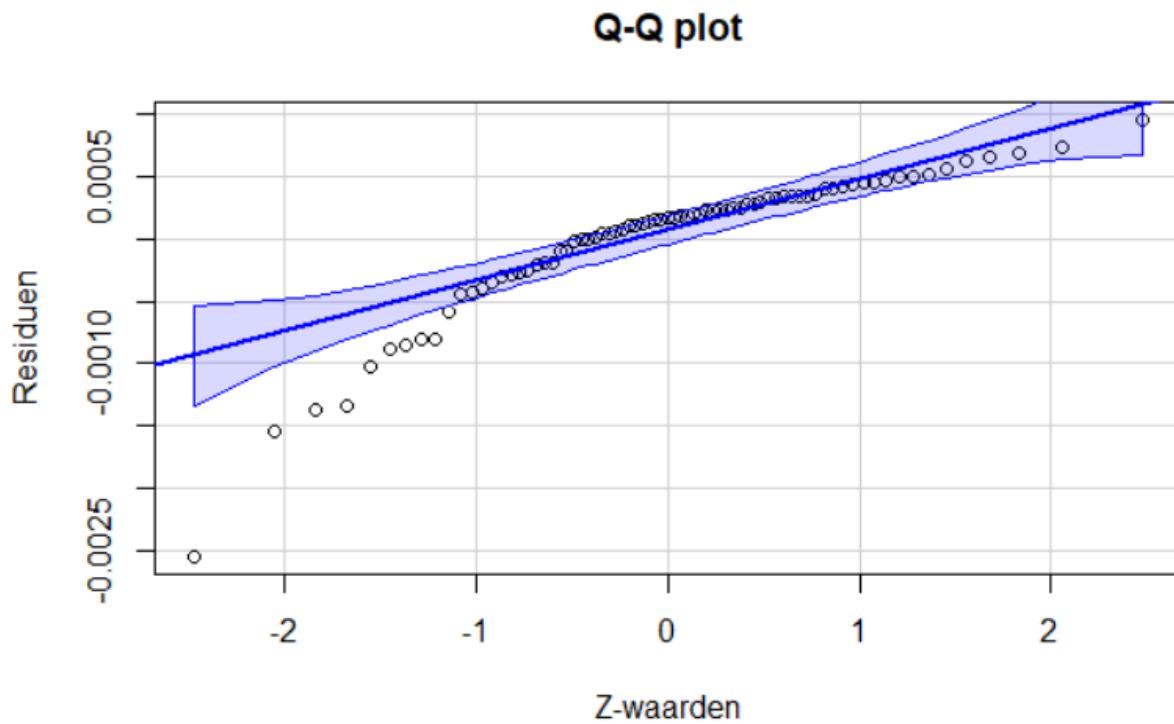
De dijkkring en het dijkkringdeel is gescheiden door een “;”. Voorbeeld: Dijkkring “10;1” wil zeggen dijkkring 10 deel 1.

	A	B	C	D	E	F	G
1	dijkkring	P_2021	P_2051	Pmidden_2021	Pmidden_2046	Pmidden_2051	total_cost
2	1;1	0.000612	0.001677	0.003362549	0.002091119	0.001901607	3.606007
3	2;1	0.000614	0.001709	0.004960533	0.003084881	0.002805308	9.888002
4	3;1	0.000637	0.002146	0.003606192	0.002242637	0.002039393	10.98391
5	4;1	0.000638	0.002152	0.003191553	0.001984779	0.001804904	1.66723
6	5;1	0.000306	0.000838	0.003140897	0.001953277	0.001776257	14.58498
7	6;1	0.00031	0.00091	0.002063947	0.001134725	0.001006767	15.75754
8	6;2	0.000311	0.000919	0.003004031	0.001679708	0.001495339	96.37511
9	6;3	0.000299	0.000738	0.002558084	0.001442489	0.001286329	52.88996
10	6;4	0.000392	0.001627	0.004705438	0.002633666	0.002345054	25.58572
11	7;1	0.000576	0.000141	0.000410631	0.000171526	0.000155432	145.9667
12	8;1	0.000286	5.23E-05	0.000238421	8.22E-05	7.33E-05	175.6086
13	8;2	9.93E-06	2.09E-05	3.41E-05	2.10E-05	1.90E-05	352.9163
14	9;1	9.71E-05	8.17E-06	0.000195418	0.000121528	8.70E-05	164.7863
15	10;1	0.001383	0.000184	0.000617772	0.000243189	0.000221149	181.045
16	11;1	0.001328	8.52E-05	0.000677997	0.000271543	0.000246934	142.0611
17	12;1	0.000671	0.000119	0.000435597	0.000211782	0.000154548	85.22221
18	13;1	0.000121	0.000314	0.001828625	0.001031151	0.000919522	21.4979
19	13;2	0.000134	0.00014	0.000296097	0.000119533	0.0001087	111.4543
20	13;4	0.00012	0.000294	0.00046944	0.000291937	0.00026548	42.13539
21	13-b;1	0.002081	0.002538	0.003225291	0.002005761	0.001823984	7.370583
22	14;1	0.000104	0.000129	0.000199494	0.000113448	0.000101337	115.1008
23	14;2	0.000115	0.000229	0.001023429	0.000576854	0.00051436	4.809579
24	14;3	0.000122	2.28E-05	0.000177492	9.98E-05	8.90E-05	205.6766
25	15;1	1.39E-05	7.20E-05	3.89E-05	2.42E-05	2.20E-05	599.9824
26	16;1	1.66E-05	7.09E-05	9.14E-05	5.68E-05	5.17E-05	1197.056
27	17;1	0.000648	0.000113	0.000330952	0.000136846	0.000122359	293.3373
28	18;1	0.000235	1.18E-05	0.000113167	3.33E-05	3.03E-05	44.47495

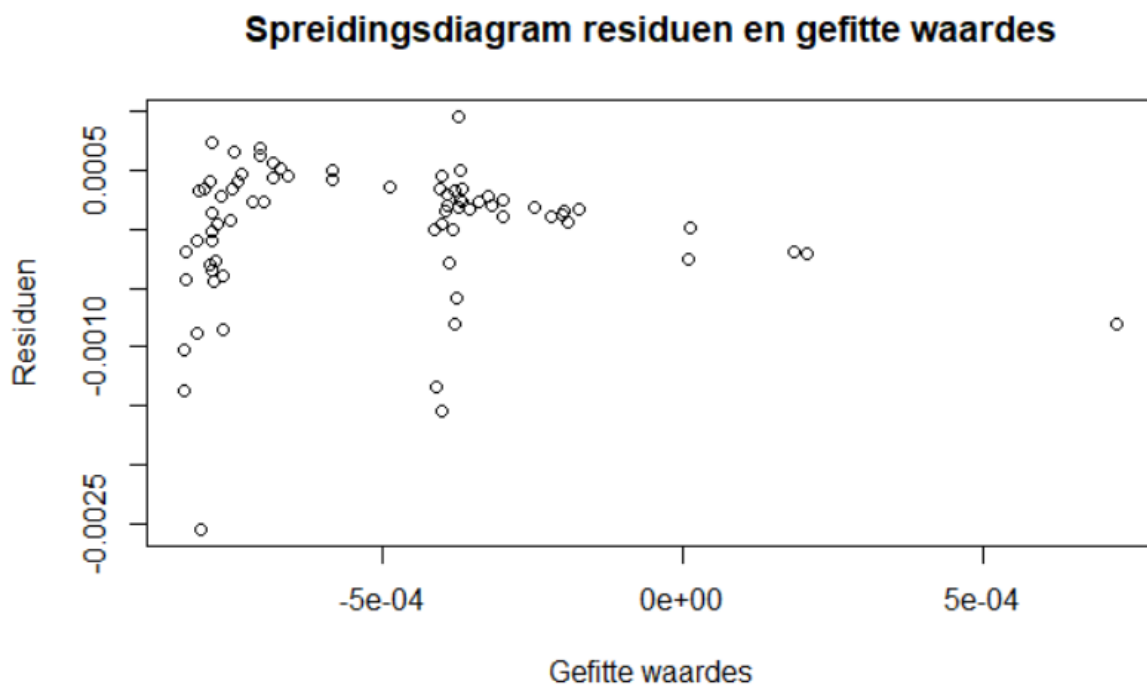
29	19;1	0.000232	0.000488	0.002314113	0.001439112	0.00130869	9.530522
30	20;1	4.12E-05	9.24E-05	0.000104126	4.87E-05	4.33E-05	126.6329
31	20;2	0.000357	8.84E-05	0.000493759	0.000213472	0.000189703	33.47758
32	20;3	3.91E-05	0.000106	9.01E-05	5.13E-05	4.58E-05	99.62749
33	21;1	0.001377	0.000299	0.002261313	0.00097452	0.000883308	119.7062
34	24;1	0.001445	0.000131	0.000434473	0.000140188	0.000127483	210.552
35	25;1	0.000282	0.00052	0.001151391	0.000654771	0.000584875	5.140338
36	25;2	0.000689	0.003437	0.009530243	0.005388063	0.004807268	11.3067
37	26;1	0.000378	0.000131	0.000630711	0.000195545	0.000174376	16.8468
38	26;2	0.000756	0.000224	0.000259399	0.000136582	0.000121591	63.94168
39	27;1	0.00037	9.93E-05	0.000748916	0.000256088	0.000228365	47.30854
40	28;1	0.000362	0.00231	0.001607439	0.000906426	0.000808299	12.80939
41	29;1	0.000292	0.000451	0.002856586	0.001624478	0.001433521	6.231783
42	29;2	0.00029	0.000615	0.000652688	0.000368047	0.000328203	76.23032
43	30;1	0.000339	0.000627	0.002606871	0.001472983	0.001313522	74.96739
44	31;1	0.00039	0.000493	0.002845431	0.001587804	0.001430821	58.48826
45	32;1	0.000289	0.000525	0.004860942	0.002741058	0.002444318	14.68485
46	32;2	0.000277	0.000467	0.002859663	0.001612548	0.001437978	31.07871
47	34;1	0.002492	0.00058	0.002536241	0.000718581	0.00064056	166.6443
48	34-a;1	0.000104	0.0003	0.000195405	0.00012152	0.000110507	49.24572
49	35;1	8.15E-05	0.000158	0.000170948	0.00010631	9.67E-05	189.5848
50	36;1	8.93E-05	1.13E-05	0.00010338	4.41E-05	4.01E-05	629.9041
51	36-a;1	0.002624	0.000406	0.002241561	0.001001379	0.000910627	4.589598
52	37;1	0.002559	0.00029	0.00261342	0.000871463	0.000792485	2.274811
53	38;1	3.86E-05	0.00014	6.44E-05	4.00E-05	3.64E-05	290.4825
54	38;2	4.50E-05	0.000154	9.22E-05	5.74E-05	5.22E-05	156.1597
55	39;1	0.002677	0.000647	0.001908919	0.000985784	0.000896445	22.5825
56	40;1	1.39E-06	4.27E-06	1.01E-05	6.28E-06	5.71E-06	21.47951

57	40;2	0.004898	0.000378	0.00241635	0.000998959	0.000908426	25.74693
58	41;1	4.79E-05	0.000158	6.82E-05	4.24E-05	3.86E-05	469.6729
59	41;2	6.83E-05	0.000222	0.000221034	0.000137458	0.000125001	263.8876
60	42;1	0.000109	0.000308	0.000332178	0.000206577	0.000187855	387.5883
61	43;1	0.000108	0.000343	0.000201406	0.000125252	0.0001139	1535.404
62	44;1	7.02E-06	1.19E-05	1.21E-05	7.50E-06	6.82E-06	168.7858
63	44;2	0.002103	0.002701	0.001677838	0.001043422	0.00094886	43.23293
64	45;1	1.76E-06	2.99E-06	3.03E-06	1.88E-06	1.71E-06	32.64868
65	45;2	0.002201	0.003557	0.002874586	0.001787662	0.001625652	25.374
66	46;1	0.00214	0.000127	0.001252575	0.000167527	0.000152344	28.3934
67	47;1	1.48E-05	3.51E-05	0.000107437	6.68E-05	6.08E-05	148.3861
68	48;1	5.00E-05	0.000151	9.98E-05	6.21E-05	5.65E-05	543.3052
69	48;2	1.28E-05	5.00E-05	3.94E-05	2.45E-05	2.23E-05	193.3136
70	49;1	0.002602	0.000209	0.001357769	0.000449384	0.000408658	72.10411
71	50;1	1.49E-05	6.03E-05	4.31E-05	2.68E-05	2.44E-05	59.15015
72	51;1	0.002942	0.000129	0.00413369	0.000248931	0.000226371	41.72094
73	52;1	9.44E-05	0.000389	0.000407857	0.00025364	0.000230654	269.3457
74	65;1	0.000327	0.000897	0.000720142	0.000447845	0.000407259	26.53986
75	68;1	0.000572	0.001591	0.000913035	0.000567803	0.000516345	95.28358
76	86;1	0.000914	0.002499	0.002078481	0.001292576	0.001175434	4.470563
77	87;1	0.000268	4.16E-05	0.000387445	0.000240946	0.000156192	29.15769

## Bijlage 4: Validatie regressiemodel



Figuur 1: Q-Q plot van de residuen met bijbehorende Z-waarde van het meervoudige regressiemodel



Figuur 2: Spreidingsdiagram residuen en gefitte waarden van het meervoudige regressiemodel