



Dienst Getijdewateren
Rijksinstituut voor Kust en Zee RIKZ

De basispeilen langs de Nederlandse kust Statistisch onderzoek

Rapport DGW-93.023
Deel 1 - Tekst

Ir. D. Dillingh
Rijkswaterstaat
Dienst Getijdewateren
's-Gravenhage

Prof. dr. L. de Haan
Erasmus Universiteit
Rotterdam

Dr. R. Helmers
Centrum voor Wiskunde en Informatica
Amsterdam

Dr. G.P. Können
Koninklijk Nederlands Meteorologisch
Instituut
De Bilt

Ir. J. van Malde
Rijkswaterstaat
Dienst Getijdewateren
's-Gravenhage

's-Gravenhage, april 1993

Inhoudsopgave

1. **Inleiding 7**
2. **Beschrijving van het verschijnsel 11**
 - 2.1 **Het verticale getij 11**
 - 2.1.1 Het astronomisch getij en zijn componenten 11
 - 2.1.2 Het optredende getij en de autocorrelatie tussen HW-opzetten 14
 - 2.2 **Meteorologische gezichtspunten 17**
 - 2.2.1 Het verband tussen circulatiepatronen en hoge HW-opzetten 17
 - 2.2.2 De persistentie van circulatietypen 20
 - 2.2.3 Bovengrens hoogwaterstanden 20
 - 2.3 **Veranderingen in „uitwendige omstandigheden” 22**
 - 2.3.1 Relatieve zeespiegelrijzing en invloeden grote werken 22
 - 2.3.2 Het effect van klimaatschommelingen op de stormvloedfrequenties 23
3. **Aanpak 25**
 - 3.1 **De werkwijze der Deltacommissie 25**
 - 3.1.1 Korte beschrijving van de methode 25
 - 3.1.2 Enkele kanttekeningen bij de werkwijze der Deltacommissie 28
 - 3.2. **Vergelijkbare onderzoeken in het buitenland 30**
 - 3.3 **De huidige aanpak 32**
 - 3.3.1 De in beschouwing genomen locaties en hun waarnemingsperioden 32
 - 3.3.2 De te verwerken veranderlijke(n) 32
 - 3.3.3 Beoordeling van HW en van HW-opzet als stochastische grootheden 34
4. **De voorberekingen van het waarnemingsmateriaal 37**
 - 4.1 **Keuze stations en waarnemingsperioden 37**
 - 4.2 **Samenstelling bestanden 38**
 - 4.2.1 Het HW-bestand 38
 - 4.2.2 Bestanden van astronomische hoogwaterstanden en HW-opzetten 40
 - 4.3 **De uitgevoerde selecties 41**
 - 4.3.1 Probleemstelling 41
 - 4.3.2 Eerste homogeniseringsselectie: invoering drempels 42
 - 4.3.3 Tweede homogeniseringsselectie: vaststelling stormseizoenen 43
 - 4.3.4 Selectie tegen autocorrelatie van HW-opzet 46
 - 4.3.5 De homogeniteit van het stormklimaat in de beschouwde waarnemingsperiode 50

- 4.4 Toetsing van de onafhankelijkheid van de astronomische hoogwaterstanden en HW-opzetten 51
 - 4.4.1 De Spearman-toets 52
 - 4.4.2 Gelijke waarnemingen 53
 - 4.4.3 Toepassing op de astronomische hoogwaterstanden en HW-opzetten 53
- 4.5 Herleiding van HW's naar de toestand 1985 54
- 4.6 De kansdichtheidfuncties der astronomische hoogwaterstanden in het lange stormseizoen 55
- 4.7 Histogrammen van het bewerkte waarnemingsmateriaal 57

- 5. **Toegepaste methoden 59**
 - 5.1 Inleiding 59
 - 5.2 Verdelingsvrije benadering 61
 - 5.3 Parametrische aanpak 67
 - 5.4 Welke statistische methode verdient de voorkeur? 73

- 6. **Berekeningsresultaten 75**
 - 6.1 Inleiding 75
 - 6.2 Gebruik van de empirische verdelingsfuncties 78
 - 6.3 Verdelingsvrije methode VVM-c 80
 - 6.4 Verdelingsvrije methode VVM-0 82
 - 6.5 GPV-model 83
 - 6.6 Convolutiemethode (CON) 84
 - 6.7 Het GEV-s-model 85
 - 6.8 Samenvattende tabel van de resultaten 86

- 7. **Evaluatie van de berekeningsresultaten 87**
 - 7.1 De empirische frequentieverdelingen 87
 - 7.2 Bespreking van de resultaten van de toegepaste statistische modellen 88
 - 7.3 Betrouwbaarheidsintervallen 90
 - 7.4 Bijzondere selecties 92
 - 7.5 Gevoeligheidsonderzoek m.b.t. de selectiemethode 94
 - 7.6 Tijdens het onderzoek afgevallen methoden 95
 - 7.7 Vergelijking met een aanpak analoog aan die van de Deltacommissie voor Hoek van Holland 96
 - 7.8 Resulterende overschrijdingslijnen en betrouwbaarheidsintervallen 96
 - 7.9 Fysische aspecten en ruimtelijke samenhang 98
 - 7.10 West-Terschelling en IJmuiden 99

- 8. **Historische bevindingen, zeeberen en zeebevingen 101**
 - 8.1 Historische zware stormvloed 101
 - 8.2 Bijzondere waterstandsverheffingen 103

- 9. **Samenvatting, conclusies en aanbevelingen 107**
 - 9.1 Samenvatting 107
 - 9.1.1 Probleemstelling en voorberekingen van het waarnemingsmateriaal 107
 - 9.1.2 De toegepaste methoden en hun uitkomsten 109
 - 9.1.3 Evaluatie en afsluitende opmerkingen 111
 - 9.2 Gevolgtrekkingen en aanbevelingen 113

10. Literatuur 117

Aanhangsels

- 1 Clustering van hoge HW-opzetten in het korte stormseizoen 122
- 2 De gebruikelijke wijze van schatten van betrouwbaarheidsintervallen bij het GPV-model 125
- 3 Bootstrap-betrouwbaarheidsintervallen voor het GPV-model en het VVM-0 model 127

Verwijzingslijst van belangrijke begrippen 131

1. Inleiding

Kort na de stormramp van 1 februari 1953 stelde de Minister van Verkeer en Waterstaat de Deltacommissie in; zij kreeg tot taak voor het Nederlandse getijgebied een plan op te stellen ter beveiliging van het land tegen overstromingen. In dit kader stelde de commissie in haar eindverslag (d.i. deel 1 van het „*Deltarapport*“) voor elk van een groot aantal peilmeetstations in het Nederlandse kustgebied en langs de Westerschelde twee buitenwaterstanden vast, t.w. het zgn. *basispeil* (overschrijdingsfrequentie: 10^{-4} p.j.) en het zgn. *ontwerppeil* [Deltacommissie 1960, *bl. 34 en 35]. Deze ontwerppeilen dienden evenzovele uitgangspunten te vormen voor de vereiste verzwaring der zeekeringen en hoogwaterkeringen in dit gebied; zij wijken voor veruit de meeste plaatsen af van de basispeilen vanwege:

- enerzijds verwachte waterstandsverhogingen als gevolg van de voorgenomen afdammingen der zeearmen,
- anderzijds een verlaging van enige dm (de zgn. economische reductie), toegepast voor de gebieden bezuiden de Nieuwe Maas c.a. en beoosten Den Helder.

Bij deze opgave tekende de commissie aan dat in deze peilen vrij veel onzekerheid schuilt en dit in het bijzonder voor de Westelijke Waddenzee, waarvoor als gevolg van de afsluiting der Zuiderzee (mei 1932) slechts korte homogene waarnemingsreeksen ter beschikking stonden. Zij beval dan ook voortzetting der studies onder verwerking van ter beschikking komende nieuwe gegevens aan [Deltacommissie 1960, bl. 36, 104].

Zowel bij het Koninklijk Nederlands Meteorologisch Instituut (KNMI) als bij verschillende diensten van de Rijkswaterstaat zijn zulke studies ook verricht; daarvan verdienen er hier ten behoeve van de probleemstelling twee vermelding.

Allereerst een onderzoek uit 1976/77 naar de invloed van bepaalde vooronderstellingen, die aan de gevolgde werkwijze ter vaststelling van de basispeilen ten grondslag lagen. Deze studie [Van Urk en Van Malde 1977] liep uit op een aantal voorstellen voor - sterk aanbevolen - verder onderzoek waaronder dat naar de geldigheid van de zgn. *grenspeilen* (dat zijn de vastgestelde hoogwaterstanden met een gemiddelde overschrijdingsfrequentie van $5 \cdot 10^{-1}$ p.j.); deze grenspeilen waren voor het overgrote deel der stations in 1938 bepaald. Dit onderzoek is aangevat m.b.v. het waarnemingsmateriaal voor de Nederlandse kust en de Westerschelde uit de periode 1971...1980 [De Ronde 1984a]. Inderdaad bleken in het bedoelde decennium de hoogwaterstanden met een geschatte overschrijdingsfrequentie $> 10^{-1}$ p.j. vergeleken met die in het „*Deltarapport*“ [Wemelsfelder 1960a] aanmerkelijk hoger te liggen, wat vooral gevolg is van de optredende relatieve zeespiegelrijzing en van vergroting van getijverschillen. De grenspeilen moesten dan ook in 1985 overeenkomstig de bevindingen van het laatstgenoemde onderzoek worden verhoogd.

* de literatuurverwijzingen staan tussen [].

De uitkomsten van deze studies gaven volop aanleiding te besluiten voor de belangrijkste kuststations een diepergaand en uitgebreider onderzoek in te stellen naar de *hoogwateroverschrijdingslijnen* voor lage overschrijdingsfrequenties (deze lijnen geven voor elk station afzonderlijk het verband aan tussen de hoogwaterstanden en hun overschrijdingsfrequenties) en met name ook naar de basispeilen. Vergeleken met het Deltarapport beschikte men daarvoor behalve over 30 extra jaren van waterstandsgegevens ook over nieuwe statistische inzichten en theorieën. Hiertoe werd door de toenmalige Directie Waterhuishouding en Waterbeweging van de Rijkswaterstaat samenwerking gezocht met het KNMI te De Bilt en het Centrum voor Wiskunde en Informatica (CWI, vroeger: Mathematisch Centrum) te Amsterdam, welke beide instellingen destijds belangrijke bijdragen hadden geleverd aan het Deltarapport. Medio 1984 kon het onderzoek een aanvang nemen; daarbij was van meet af aan ook nauw betrokken dr. L. de Haan, hoogleraar aan de Erasmus Universiteit Rotterdam en adviseur van het CWI.

Het ingestelde onderzoek is geheel gericht geweest op statistische verwerking van waterstandsgegevens van verschillende aard; in dit kader is niet alleen gebruik gemaakt van beschikbare theorieën en benaderingswijzen, maar konden ook nieuwe methoden worden ontwikkeld en toegepast. Om op zo verantwoord mogelijke wijze reeksen van waterstandsgegevens te kunnen vaststellen diende overigens ook veel werk op meteorologisch, hydrologisch en hydrologisch-statistisch gebied te worden verzet. Daarbij zijn de verwerkte gegevens, die deels teruggaan tot 1881, alle herleid tot de toestand 1985; *de uitkomsten van het onderzoek hebben dan ook betrekking op de (klimatologische, morfologische en algemeen-hydrologische) situatie van 1985.*

Er is dus in het raam van dit onderzoek van afgezien om ook langs andere weg inzicht te verkrijgen in de vermelde problematiek. Evenwel kan de statistische aanpak van het gestelde probleem geen aanspraak maken op een geheel sluitende bewijsvoering: aan de vooronderstellingen, waarop hij is gegrond, kan namelijk niet geheel worden voldaan terwijl de vraagstelling een zeer grote extrapolatie vanuit het waarnemingsgebied vereist, wat uit de aard der zaak met onzekerheden is behept. Het verdient daarom ongetwijfeld aanbeveling ook andere methoden, zoals modelonderzoekingen voor b.v. het Waddenzeegebied, toe te passen ten einde een zo goed mogelijk beeld van mogelijke extreme stormvloedstanden te verkrijgen.

Over de voortgang van het onderzoek zijn twee tussentijdse verslagen uitgebracht; het eerste, over de periode juni 1984 ... december 1985, verscheen in juni 1986, het tweede (over de periode tot medio 1988) in februari 1989 [Samenwerkingsverband Basispeilen 1986 resp. Van der Made 1989]. Met het voorliggende rapport wordt de studie afgesloten.

In de loop van het onderzoek zijn geregeld voortgangsbesprekingen gevoerd, waaraan van de betrokken instellingen een of meer vaste vertegenwoordigers (incidenteel vergezeld van een medewerker) deelnamen. Dit „Samenwerkingsverband Frequentielijnen Kust” leed een gevoelig verlies door het onverwachts overlijden van dr.ir. J.W. van der Made in oktober 1991.

De leden van het bovenbedoelde samenwerkingsverband waren:
namens het CWI: dr. R. Helmers; drs. A.L.M. Dekkers (tot 1 januari 1989); prof. dr. L. de Haan;
namens het KNMI: H. Daan (tot 1 januari 1989); dr. G.P. Können (vanaf 1 september 1988);
namens de
Rijkswaterstaat*: ir. D. Dillingh (vanaf 1 april 1986); dr.ir. J.W. van der Made († 6 oktober 1991); ir. J. van Malde; ir. J.G. de Ronde (tot 1 januari 1989).

Ten behoeve van het onderzoek hebben medewerkers van de betrokken instellingen vele werkzaamheden verricht. De dank van de leden van het samenwerkingsverband gaat wat dit betreft in het bijzonder uit naar de heren J. Doekes van de Dienst Getijdewateren, drs. G. Draisma van de Erasmus Universiteit Rotterdam en R. van der Horst van het CWI.

Volledigheidshalve zij tenslotte vermeld dat dit eindrapport onder de verantwoordelijkheid der aanvang 1993 zittende leden verschijnt.

De opzet van dit eindrapport is als volgt:

Eerst komen vrij uitgebreid ter sprake de beide verschijnselen, die in hun samengaan hoge tot zeer hoge waterstanden aan onze kusten veroorzaken. Deze verschijnselen zijn het door de aantrekkingskrachten van maan en zon opgewekte getij en de weersomstandigheden, die belangrijke waterstandsverhogingen aan onze kusten bewerkstelligen. Op dit hoofdstuk 2 volgt in hoofdstuk 3 een beschrijving met kanttekeningen van de werkwijze der Deltacommissie met betrekking tot de vaststelling van de basispeilen en van de hoogwateroverschrijdingslijnen voor Nederlandse getijstations, een korte beschrijving van vergelijkbare onderzoeken in omliggende landen en een verantwoording van de keuzen van de te onderzoeken stochastische veranderlijken.

Het beschikbare waarnemingsmateriaal is pas geschikt voor statistische analyse nadat het een aantal voorbereidingen heeft ondergaan. Een bespreking met verantwoording van deze voorbereidingen, waaronder een drietal belangrijke selecties, treft men aan in hoofdstuk 4. Hoofdstuk 5 is geheel gewijd aan de theoretische achtergronden van de toegepaste statistische analyses. Een voor niet-wiskundigen bedoelde beschrijving van deze theorieën vormt de inleiding van hoofdstuk 6, dat verder is gewijd aan hun toepassing en aan de gevonden uitkomsten. De hoofdstukken 7 en 8 bevatten aanvullende onderzoeken en een bespreking met evaluatie van de resultaten van de studie respectievelijk door andere oorzaken teweeggebrachte waterverheffingen aan onze kusten. Hoofdstuk 9 ten slotte biedt een samenvatting van alsook enkele conclusies uit het onderzoek en enige aanbevelingen, terwijl hoofdstuk 10 de literatuur opsomt, waarnaar in het voorgaande is verwezen. Enkele moeilijk in de voortgang van het betoog in te passen beschouwingen zijn in aanhangsels opgenomen. Een verwijzingslijst van belangrijke begrippen besluit het geheel. Het rapport bestaat uit drie delen. Het voorliggende eerste deel behelst de tekst, deel 2 de bijlagen terwijl deel 3 alle relevante gegevens van het verwerkte waarnemingsmateriaal bevat.

* tot 1 juli 1985: Directie Waterhuishouding en Waterbeweging; per 1 juli 1985: Dienst Getijdewateren

2. Beschrijving van het verschijnsel

2.1 Het verticale getij

2.1.1 Het astronomisch getij en zijn componenten

De getijbeweging aan kusten en in zeearmen wordt allereerst bepaald door twee factoren, te weten:

- de getijverwekkende kracht en
- de configuratie en bodemgesteldheid van zee, kust en zeearmen.

De getijverwekkende kracht vloeit voort uit:

- de aantrekkingskrachten van maan en zon op de aarde en
- de krachten samenhangende met de beweging van het samenstel aarde-maan ten opzichte van zijn zwaartepunt.

Als gevolg van de bewegingen binnen het systeem zon-aarde-maan en de draaiing van de aarde om haar as varieert de getijverwekkende kracht in de tijd met periodiciteiten, die samenhangen met deze bewegingen.

Men kan bewijzen dat de door een hemellichaam teweeggebrachte getijkracht evenredig is met zijn massa en omgekeerd evenredig met de derde macht van zijn afstand tot de aarde; de verhouding tussen de getijkrachten van de twee relevante hemellichamen maan en zon is voor hun gemiddelde afstanden tot de aarde 1 : 0,46 [Van Urk en De Ronde 1982, bl. 13/14]. Verder doet de invloed van de getijverwekkende kracht zich behalve op de grote watermassa's op aarde ook gelden op de aardbodem, maar dit zgn. aardgetij kan hier verder buiten beschouwing blijven.

De beide hierboven genoemde factoren bepalen samen het „*astronomisch getij*“. Het optredende getij wijkt hiervan veelvuldig en soms aanzienlijk af als gevolg van de meteorologische omstandigheden boven zee, kusten en zeearmen.

Het getij manifesteert zich als heen en weer trekkende stromingen (horizontaal getij) en door op- en neergaande waterspiegels, het zgn. verticale getij dat in dit rapport verder kortweg *het getij* wordt genoemd. Dit getij plant zich als golf langs de Nederlandse kust voort van zuid naar noord en op de zeearmen in landwaartse richting.

Daarbij veranderen de hoog- en laagwaterstanden en daarmee hun verschil. Dit komt tot uitdrukking in het gemiddelde *tijverschil* dat langs de kust van zuid naar noord gaande eerst afneemt om voorbij Petten weer toe te nemen; op de zeearmen daarentegen is in landwaartse richting alleen sprake van een toenemend tijverschil. Verder overheerst in de Nederlandse getijdewateren de maansinvloed, waardoor de gemiddelde tijdsduur tussen twee opeenvolgende hoogwaters c.q. laagwaters 12 uur 25 min. belooft; hierdoor treden op een willekeurige plaats op gemiddeld 6 dagen per week twee hoog- en twee laagwaters op („tweemaaldaags getij“), terwijl op de resterende dag bij wijze van spreken één hoogwater dan wel één laagwater ontbreekt.

Het astronomische getij vertoont een groot aantal periodieke schommelingen, waarvan hier vermelding verdienen:

- a de **dagelijkse ongelijkheid**: de twee opeenvolgende hoogwater- c.q. laagwaterstanden zijn in het algemeen ongelijk. De veruit belangrijkste oorzaak hiervan is de **maansdeclinatie**, dat is de hoek, die de lijn middelpunt maan tot middelpunt aarde maakt met het equatorvlak, d.i. het vlak bepaald door de evenaar (deze hoek is overigens niet constant, maar onderhevig aan een tweetal schommelingen, waarvan de kortste met een periode van 27,21 dagen gevolg is van de draaiing van de maan om de aarde);
- b de „**twewekelijkse**“: tijdens een maanmaand (duur: 29,53 dagen) bereikt de getijverwekkende kracht maxima bij nieuwe en volle maan (NM resp. VM) en minima bij de kwartierstanden der maan. Gevolg: sterk resp. zwak ontwikkelde getijden (**spring-** en **doodtij**en), die in onze streken zo'n 2 dagen na bedoelde schijngestalten, d.w.z. met een gemiddelde periode van 14,76 dagen, optreden. Overigens verschillen tijdens iedere maanmaand de springtijden na NM van die na VM, wat evenzeer geldt voor de doottijden na EK t.o.v. die na LK;
- c de **half-jaarlijkse**: de maansbaan om de aarde is een ellips (excentriciteit $e = 0,055$), zodat de afstand aarde-maan varieert. Bij de minimale afstand staat de maan in het zgn. **perigeum**, bij de maximale afstand in het zgn. **apogeum**. In het eerste geval is de getijkracht van de maan bijna 35,3% groter dan in het tweede; doet dit eerste geval zich voor bij VM of NM dan zijn de springtijden bij gevolg extra krachtig. Daar de maansbaan langzaam draait (periode: 8,85 jaar) ligt tussen springtijden met de maan in het perigeum niet 0,5 jaar, maar 0,56 jaar;
- d de **jaarlijkse**: bij dag-en-nachtevening (equinox), dat wil zeggen op of nabij 21 maart en 22 september, bevindt de zon zich in het equatorvlak (dit vlak maakt een hoek van $23,5^\circ$ met de ecliptica, d.i. het „zonsbaanvlak“). De zonsdeclinatie is dan dus 0° , wat op zich de springtijden versterkt en de doottijden verzwakt - bij de zonnewenden rond 21 juni en 21 december is de zonsdeclinatie daarentegen het grootst ($\pm 23,5^\circ$ resp. $23,5^\circ$), wat de springtijden verzwakt en de doottijden versterkt [o.a. Pugh 1987, bl. 69/70 en 83/84]. Verder is ook de baan van de aarde om de zon licht ellipsvormig ($e = 0,017$); de grootste getijkracht van de zon, in onze tijd optredend omstreeks 4 januari, is hierdoor ruim 10% groter is dan de kleinste;
- e de **18,61-jarige cyclus** (het „nodale tij“): het maansbaanvlak maakt een hoek van 5° met de ecliptica en wentelt in 18,61 jaar om een as daar loodrecht op. Hierdoor varieert de onder a genoemde schommeling der maansdeclinatie tussen 57° (d.w.z. van $+28,5^\circ$ naar $-28,5^\circ$) en 37° (van $+18,5^\circ$ naar $-18,5^\circ$) en daarmee die van de uitwerking der aantrekkingskracht van de maan. Aan onze kust zijn de getijden in het eerste geval iets zwakker en in het tweede iets krachtiger dan gemiddeld. De snijlijn van maansbaanvlak en ecliptica is de **knopenlijn**, die om de 9,3 jaar in het equatorvlak ligt; die gemiddelde situatie halverwege deed zich voor het laatst voor op 2 juli 1992 en daarvoor in maart 1983.

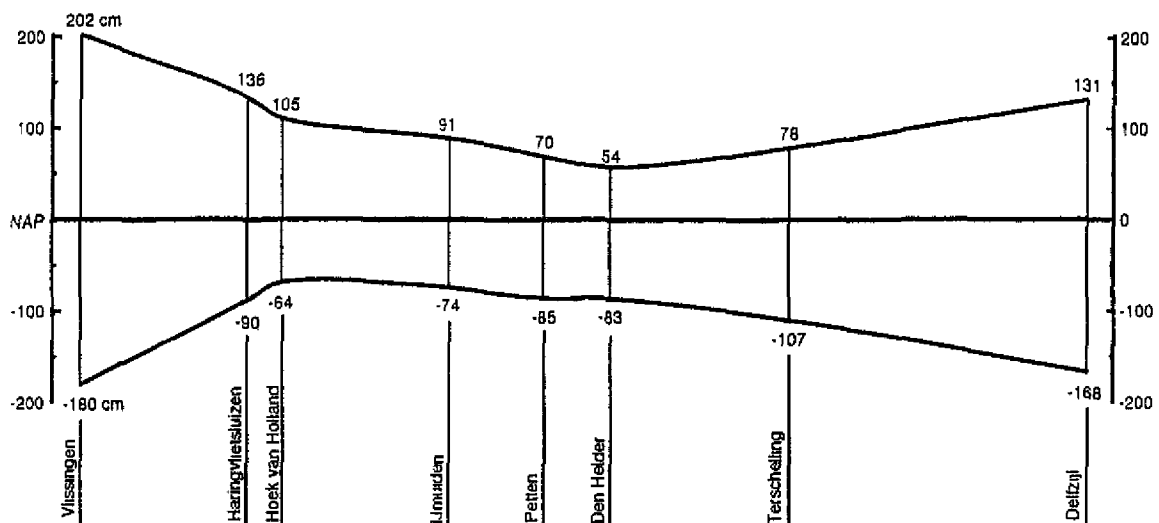
Een nadere toelichting op deze periodiciteiten is elders gegeven [TO 1971-1980, par. 5.1 en 5.3; Van Urk en De Ronde 1982, par. 3.1 en 3.2 (waarin ook nog enkele andere, de getijverwekkende kracht

beïnvloedende, uiterst langzame astronomische periodieke schommelingen zijn genoemd]). Volledigheidshalve zij nog opgemerkt dat de aarde onderhevig is aan een geringe, maar grillige aswiebeling (gemiddelde periode ≈ 436 dagen [Pugh 1987, bl. 316]); deze aswiebeling bewerkstelligt het zgn. poolgetij of Chandler-effect, dat aan onze kusten overigens van ondergeschikt belang is [Van Malde 1992, par. 2.6].

Voor een uitvoeriger overzicht van de relevante astronomische aspecten raadplege men de handboeken.

Het astronomisch getij laat zich bij benadering uit waterstandswaarnemingen berekenen; zogenaamde getijtafels bevatten hiervan de uitkomsten voor wat betreft de hoog- en laagwaterstanden en -tijdstippen ter plaatse van een aantal getijstations. De getijtafels voor Nederland werden tot en met die van 1985 berekend met de zogeheten *culminatiemethode* [Van Urk en De Ronde 1982]. Vanaf de getijdetafels 1986 past men de zgn. *harmonische analyse* toe. Bij de culminatiemethode verwerkt men de opgetreden hoog- en laagwaters uit de analyseperiode onder eliminering per kalendermaand van door weersinvloeden sterk afwijkende gevallen [o.c., par. 5.3.2]. Dit betekent dat de berekende getijgegevens als het ware betrekking hebben op een onbekende toestand met „middelbaar weer” en geen „zuivere astronomische getijgegevens” zijn. Dit laatste geldt, zij het op andere wijze, overigens ook voor getijgegevens berekend met de harmonische analyse, bij welke theoretisch meer bevredigende methode geen waterstanden worden geëlimineerd. Een tweede voordeel ervan boven de culminatiemethode is dat ze in staat stelt het gehele verloop van de „astronomische” waterstanden (d.w.z. de opeenvolgende „astronomische” *getijkrommen*) te bepalen.

Figuur 2.1:
Gemiddelde hoog- en laagwaterstanden langs de Nederlandse kust in cm t.o.v. NAP (slotgemiddelden 1981.0).



In de wandeling duidt men de in getijtafels afgedrukte waarden en het eventueel erbij berekende getijverloop toch aan als astronomische getijgegevens; dit spraakgebruik is verder in dit rapport gevolgd.

Een eerste indruk van het getijregime aan een getijstation over een zeker kalenderjaar bieden de gemiddelden van ettelijke getijgrootheden over dat jaar. De hoogtematen hiervan zijn in de tijd zowel aan toevallige fluctuaties en periodieke schommelingen als aan langzame, meer trendmatige veranderingen onderhevig. Van elk dezer grootheden wordt nu om de tien jaar onder eliminering van de periodieke effecten een vereffende waarde vastgesteld, die mede een mogelijke trend markeert. Zo'n waarde heet **slotgemiddelde** en is bij grote benadering een astronomisch getijgegeven in de gangbare betekenis van het woord; de slotgemiddelden 1981.0 hebben betrekking op de toestand per 1 januari 1981.

Fig. 2.1 brengt nu voor onze kust in beeld het verloop van de gemiddelde hoog- en laagwaterstanden volgens de slotgemiddelden 1981.0. Het gemiddelde tijverschil varieert volgens dit prentje van ruim 380 cm bij Vlissingen tot ruim 135 cm bij Den Helder (op de Westerschelde neemt het toe van 380 cm tot 475 cm). Een indruk van de plaatselijke veranderlijkheid van het astronomisch getij ten gevolge van de dagelijkse ongelijkheid en de schommelingen met perioden ≤ 1 jaar biedt tabel 2.1.

Tabel 2.1:

Variaties van hoog- en laagwaterstanden (in cm t.o.v. NAP) en in tijverschil (in cm) van het astronomisch getij in 1982 te Vlissingen en Den Helder.

	Vlissingen						Den Helder					
	max. getalwaarden	STij	GTij	DTij	min. getalwaarden	max. getalwaarden	STij	GTij	DTij	min. getalwaarden		
HW	19-9, 272	239	202	153	122, 17- 4	27-1, 89	62	54	44	15, 17- 5		
LW	11-2, -250	-203	-180	-149	-110, 27-10	11-2, -124	- 92	- 83	- 69	-47, 26- 9		
TV	25-4, 514	442	382	302	237, 26-10	26-2, 193	154	137	133	85, 26-10		

HW:	hoogwater,	stand & evt.	datum	STij:	gemiddeld springtij	} slotgemiddelden 1981.0
LW:	laagwater,	stand & evt.	datum	GTij:	gemiddeld tij	
TV:	tijverschil,	waarde & evt.	datum	DTij:	gemiddeld doortij	

2.1.2 Het optredende getij en de autocorrelatie tussen HW-opzetten

In de jaren tachtig der vorige eeuw is men in Nederland ertoe overgegaan voor alle belangrijke relevante plaatsen de actuele getijkrommen vast te leggen met behulp van peilschrijvers. Door de inrichting van de betrokken meetopstellingen worden daarbij golven met perioden kleiner dan 1 minuut (zoals windgolven, deining, scheepsgolven en korte golven als gevolg van buiigheid) uitgefilterd.

Het zo geregistreerde getij wijkt vaak af van het astronomische; de oorzaken ervan zijn wind en luchtdruk. Wind sleept als het ware waterdeeltjes mee; zodoende verlaagt afluende wind de waterstand aan de kust, terwijl aanlandige wind hem verhoogt. Hoge en lage luchtdruk hebben, zij het veel minder sterk, achtereenvolgens dezelfde uitwerkingen. Voor het gestelde doel zijn alleen van belang aanmerkelijke waterstandsverhogingen die aan onze kust voornamelijk bij westelijke tot noordelijke winden optreden. Wanneer daarbij aan één of meer der zgn. *basisstations* Vlissingen, Brouwershaven (tot nov. 1964), Hoek van Holland, IJmuiden, Den Helder, Harlingen en Delfzijl het hoogwater een stand gelijk aan of hoger dan het *grenspeil* (overschrijdingsfrequentie $5 \cdot 10^{-1}$ p.j.) bereikt(e) spreekt men van een *stormvloed*.

De opwaaiing, de waterstandsverhoging die aanlandige wind teweegbrengt, is des te groter naarmate de waterdiepte kleiner is. Op niet-diep water kan waterstandsverhoging tot een aanmerkelijke versnelling van de voortplanting van het getij leiden, wat dan veelal een vervroeging van het hoogwater en/of van het laagwater meebrengt. Verder wekt die aanlandige wind stromingen op, die ook van invloed zijn op de getijvoortplanting. Onder die omstandigheden wordt het optredende waterstandsverloop bepaald door de volgende drie factoren:

- a het astronomisch getij;
- b de opwaaiing met het luchtdrukeffect;
- c de wisselwerking tussen astronomisch getij en weereffecten.

Het gevolg hiervan is weer dat het ogenblikkelijke verschil tussen het optredende en het astronomische getij, de zgn. (rechte) *opzet*, in de tijd belangrijk anders kan verlopen dan men op grond van de weersomstandigheden zou verwachten. Een verdere complicatie kan nog de traagheid van het water vormen, die een extra verhoging (een „opzweeping“) kan bewerkstelligen.

Klaarblijkelijk geeft het verloop van de (rechte) opzet aan een station voor een bepaalde stormvloed niet de grootste uitwerking aan, die de opgetreden weersomstandigheden op zichzelf genomen ter plaatse hebben gehad. Om deze grootste uitwerking te kunnen schatten en tevens om van die stormvloed het verloop van de uitwerking van de weersomstandigheden (het *stormeffect*) aan verschillende stations te kunnen vergelijken kan men per station de betrokken (gestyleerde) getijkrommen proberenderwijze ontleden in

- een vloeiend verloopende waterstandsverheffing (het stormeffect) en
- een gewijzigde, aangepaste astronomische getijbeweging.

De zo geconstrueerde grafieken van het stormeffect dienen voor opeenvolgende getijstations een vanuit hydrologisch en meteorologisch oogpunt logische samenhang te vertonen.

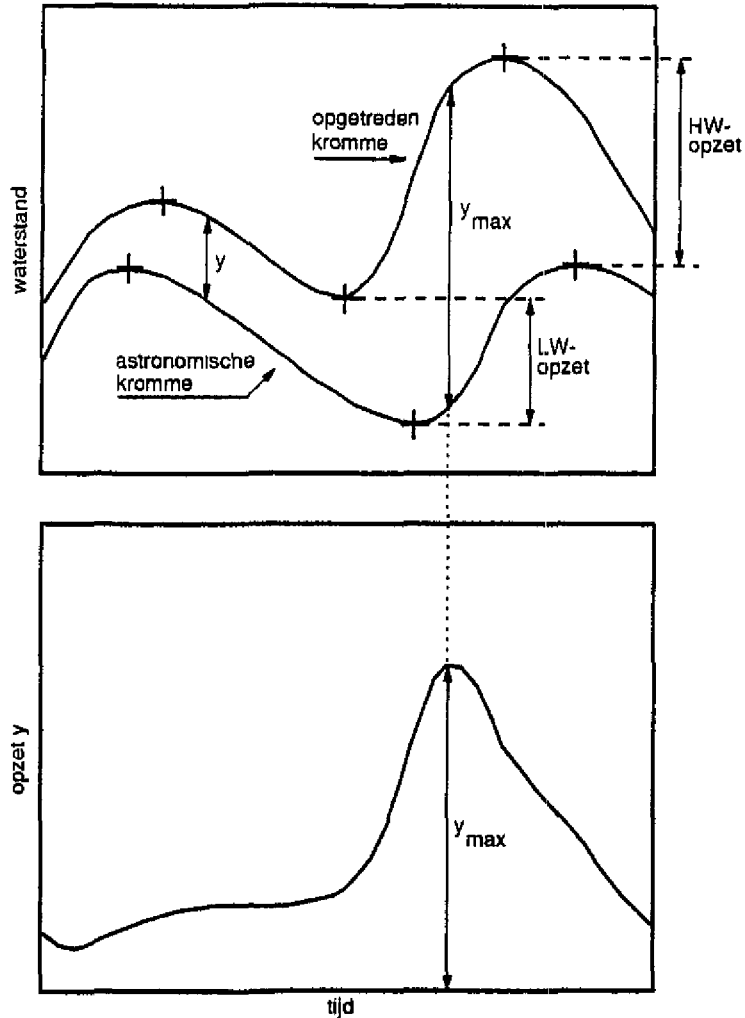
Deze procedure is b.v. gevolgd voor de stormvloed van 1953 [Rijkswaterstaat en KNMI 1961, par. 2.3, i.h.b. Fig. 2.12...2.21, bl. 111...120]; in § 2.2.3 zal voor Hoek van Holland op deze procedure worden teruggekomen. Voor de goede orde zij nog vermeld dat ook een verlating van het hoogwater mogelijk is (b.v. bij het opsteken van een krachtige noordwestelijke wind of het ruimen van een westelijke wind omstreeks het hoogwater).

Voor de karakterisering van een stormvloed aan een station gebruikt men intussen in de dagelijkse praktijk de opgetreden hoogwaterstand met daarnaast niet het stormeffect, maar twee belangrijke „scheve“

opzetten, t.w. de hoogwateropzet **HW-opzet** (d.i. het hoogteverschil tussen de opgetreden en de astronomische hoogwaterstand, ongeacht een eventueel tijdsverschil tussen beide) en de laagwateropzet **LW-opzet** (het overeenkomstige verschil tussen beide laagwaterstanden). In figuur 2.2 staan beide scheve opzetten alsook de (rechte) opzet y aangegeven.

Figuur 2.2:

Voorbeeld verloop rechte opzet y met (scheve) HW- en LW-opzet.



N.B. Opgetreden of door middel van statistische verwerkingen vastgestelde hoogwaterstanden, worden in dit rapport als **HW's** aangeduid, astronomische hoogwaterstanden worden steeds met volledige omschrijving vermeld.

Opeenvolgende hoogwaterstanden zijn duidelijk gecorreleerd, deels door het aandeel van het deterministisch bepaalde astronomische getij en deels door samenhang tussen de HW-opzetten. Deze autocorrelatie vloeit dus ten dele voort uit het „gedrag van het water”, d.w.z. uit astronomische en hydraulische wetmatigheden, en is voor een ander deel het gevolg van de correlatie, die bestaat tussen opeenvolgende meteorologische toestanden.

De invloed van de eerste factor, *het gedrag van het water*, uit zich vooral op de korte termijn: bij gelijkblijvende meteorologische omstandigheden beïnvloedt een hoge HW-opzet de opzet bij het volgende hoogwater. Een ander effect kan optreden bij belangrijke

windveranderingen, b.v. bij het tijdelijk wegvallen van de wind: het teruglopende water keert na een zekere periode, zij het gedempt tot ong. 1/4 van de oorspronkelijke hoogte, terug en kan op deze wijze een extra bijdrage leveren aan de HW-opzet zo'n 36 uur later. Dit effect, „terugslingering” of „resonantie” genoemd, deed zich b.v. voor bij de „tweelingstorm” van december 1954 [Groen 1961, bl. 235/6; Weenink 1960^a]. Overigens zou volgens recent onderzoek met behulp van het Noordzeemodel CSM-16 van de Dienst Getijdewateren van de Rijkswaterstaat [Bijlsma 1989] het getij de terugslingering in aanmerkelijk sterkere mate dempen dan men eerder meende. De invloed van de tweede factor, *correlatie tussen opeenvolgende meteorologische situaties*, is veel complexer van aard en strekt zich vaak ook over een aantal getijcycli uit; ze komt in § 2.2 nader ter sprake.

Tenslotte verdient nog vermelding dat ook andere dan de hierboven genoemde oorzaken waterstandsverhogingen van enige betekenis aan onze kusten kunnen teweegbrengen; zij komen in § 8.2 nader ter sprake. Voor zover zij in de beschouwde periode HW en HW-opzet hebben beïnvloed zijn ze in de verwerkingen „meegenomen”.

2.2 Meteorologische gezichtspunten *

Aan onze kusten doen zich grote opwaaiingen voor bij stormen uit westelijke tot noordelijke richtingen. Zulk weer wordt veroorzaakt door grote luchtdrukgradiënten om één of meer voor deze weerssituatie gunstig gelegen lagedrukgebieden of depressies. Depressies verplaatsen zich, waarbij de luchtdrukgradiënten kunnen veranderen; de route die een depressie volgt heet *depressiebaan*; deze wordt bepaald door het circulatiepatroon in de hogere luchtlagen. Bijdragen tot de in § 2.1.2 genoemde correlatie tussen opeenvolgende meteorologische situaties worden geleverd door de volgende oorzaken:

- a) één stormveld kan meer dan één hoogwater beïnvloeden;
- b) één depressie kan meer dan één stormveld veroorzaken;
- c) één circulatie-patroon kan meer dan één depressie sturen;
- d) circulatie-patronen hebben een zekere levensduur en hun overgang naar andere patronen kent voorkeuren.

In deze paragraaf wordt ingegaan op enkele eigenschappen van de circulatiepatronen en de samenhang ervan met de HW-opzetten te Hoek van Holland.

2.2.1 Het verband tussen circulatiepatronen en hoge HW-opzetten

De ontwikkeling van het luchtdrukpatroon aan het aardoppervlak wordt „gestuurd” door de luchtstromingen op het 500-mbar vlak, d.i. het *circulatiepatroon*. Dit vlak bevindt zich over het algemeen tussen 5 en 6 km boven het aardoppervlak („halverwege” de atmosfeer). De stroming op 500-mbar wordt bepaald door de geografische verdeling van de hoogte van dit vlak. Dat deze verdeling gecorreleerd is met de HW-opzet als gevolg van meteorologische effecten op de hoogwaterstand wordt duidelijk geïllustreerd in figuur 2.3. Hierin is een

* in deze paragraaf is alleen het zgn. korte stormseizoen (1 november ... 31 januari - zie § 4.3.3) beschouwd!

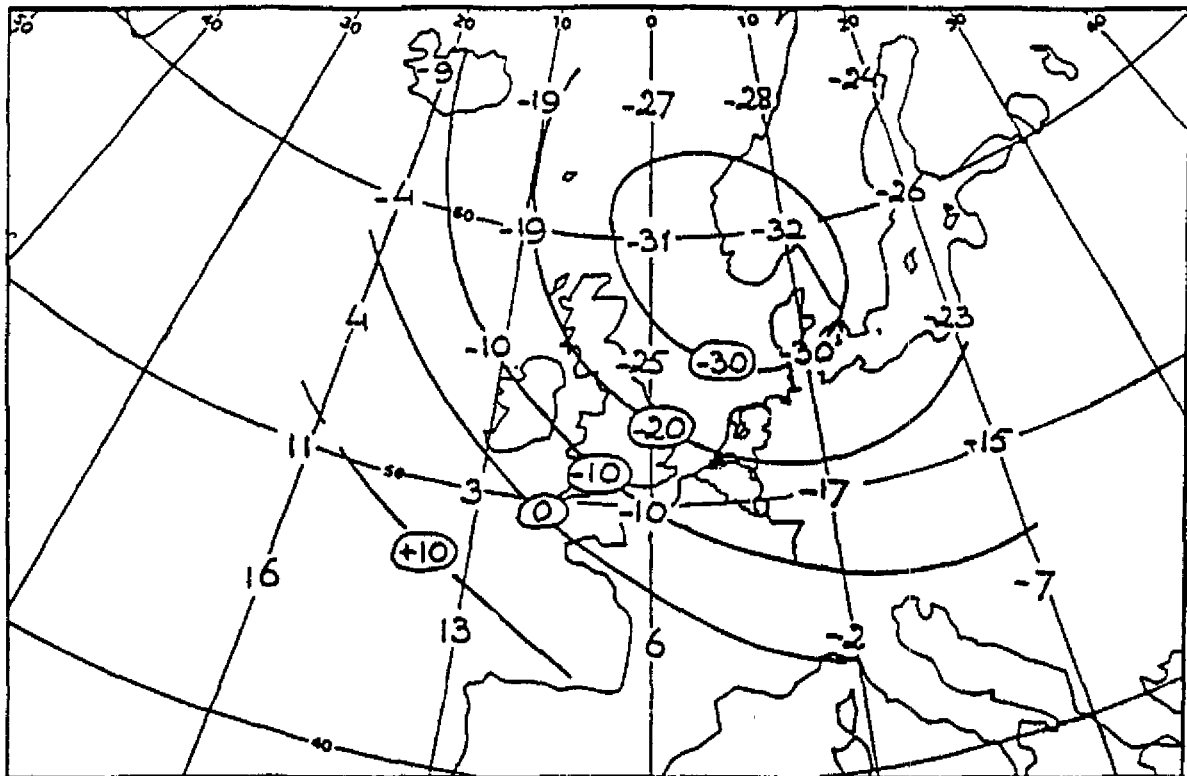
een schets gegeven van de geografische verdeling der correlatiecoëfficiënt van de 500-mbarhoogte ter plekke met de HW-opzet in Hoek van Holland.

N.B.: De correlaties zijn berekend over de periode 1/11/1949 ... 31/1/1980; voor alle HW's met een opzet < 50 cm is HW-opzet = 0 gesteld. De correlatiecoëfficiënten zijn weliswaar niet groot, maar vertonen een duidelijk ruimtelijk patroon.

Uit de figuur blijkt een negatief verband tussen de 500-mbarhoogte boven Zuid-Noorwegen en de HW-opzet te Hoek van Holland. De afbeelding geeft het voorkeurs-circulatiepatroon voor hoge opzetten weer.

Figuur 2.3:

Correlatie van de 500-mbarhoogte boven West-Europa met de HW-opzet te Hoek van Holland (in %).



Het bestaan van voorkeurscirculaties kan ook gevonden worden uit een analyse van de circulatietypen [Kruizinga 1978]. Deze classificatie is gebaseerd op een zgn. hoofdcomponentenanalyse van het 500-mbarpatroon, waarbij de eerste (d.w.z. de belangrijkste) 3 hoofdcomponenten in hoge mate bepalend bleken voor de circulatie. Deze drie zijn als volgt te karakteriseren:

- 1^e hoofdcomponent: sterkte van de west-stroming;
- 2^e hoofdcomponent: sterkte van de zuid-stroming;
- 3^e hoofdcomponent: cyclonaliteit (rotatie) van de stroming.

Door de scores op elk van deze hoofdcomponenten in tercielen te verdelen ontstaan door combinaties 27 circulatietypen, die tezamen de zgn. P27-classificatie vormen. De typen worden aangeduid met een nummer uit de reeks 1 ... 27. De classificatie is verder zo gekozen, dat alle circulatie-typen over het gehele jaar genomen dezelfde frequentie van voorkomen hebben (t.w. 3,7%). Deze eigenschap geldt echter niet wanneer alleen naar één bepaald jaargedeelte gekeken wordt (zie de waarden van F_c in tabel 2.2)

Een samenvatting van de frequenties van HW-opzetten > 50 cm en > 100 cm per circulatie-type over de eerdergenoemde periode is opgenomen in Tabel 2.2. Type 3 (dat weliswaar in de wintermaanden relatief veel voorkomt: ruim 7% van alle dagen tegen 3,7% over het gehele jaar) blijkt samen te vallen met meer dan de helft van alle HW-opzetten > 100 cm. Dit type is gekarakteriseerd door:

- sterke west-stroming;
- noordelijke stroming;
- positieve cyclonaliteit (stroming rond een lagedrukgebied).

Tabel 2.2:

Frequentie van hoge HW-opzetten te Hoek van Holland bij de relevante circulatietypen (seizoen: 1/11 ... 31/1; periode 1949-1980).

C	Gegevens circ.type		Gegevens HW-opzetten met					
	N_c	F_c %	G = 50 cm			G = 100 cm		
			T_g	F_g %	P%	T_g	F_g %	P%
1	213	7,5	28	6	7	5	7	1
2	159	5,6	60	13	19	12	16	4
3	205	7,2	132	28	32	41	53	13
4	128	4,5	15	3	6	2	3	1
5	104	3,6	19	4	9	2	3	1
6	135	4,7	72	16	27	8	11	3
7	140	4,9	14	3	5	1	1	0
8	107	3,8	28	6	13	1	1	0
9	147	5,2	35	8	12	-	-	-
11	62	2,2	5	1	4	-	-	-
12	111	3,9	30	7	14	3	4	1
15	86	3,0	7	2	4	-	-	-
16	79	2,8	1	0	1	-	-	-
18	109	3,8	4	1	2	-	-	-
20	101	3,5	1	0	1	-	-	-
21	106	3,7	9	2	4	1	1	0
(Totalen)	1992	69,9	460	100		76	100	

- C = nummer circulatietype;
- N_c = aantal kalenderdagen met type C in nov. ... jan.;
- N = 2852 = totaal aantal beschouwde kalenderdagen;
- F_c % = $(N_c/N) \cdot 100\%$ = relatieve frequentie van type C;
- G = drempelwaarde HW-opzet;
- T_g = aantal HW-opzetten per circulatietype;
- T = totaal aantal HW-opzetten (T = 460 resp. 76, voor G = 50 resp. 100 cm);
- F_g % = $(T_g/T) \cdot 100\%$; verdeling van de HW-opzetten over de circulatietypen;
- P% = $(\frac{1}{2}) \cdot (T_g/N_c) \cdot 100\%$, frequentie van de HW-opzetten per circulatietype, waarbij in de factor $\frac{1}{2}$ het optreden van twee HW's per dag is verdisconteerd. De typen 2, 3 en 6 geven elk een significant hoge kans.

Verder blijken hoge HW-opzetten samen te gaan met een beperkt aantal circulatietypen. Naast type 3 leveren slechts de typen 2 en 6 een „bovennormale” kans op een hoge HW-opzet. Deze typen wijken slechts van type 3 af door een neutrale cyclonaliteit, resp. een niet-uitgesproken noordelijke component.

Tabel 2.3:

Frequenties van hoge HW-opzetten bij verschillende tercielen van de hoofdcomponenten H.

H omschrijving	N _c	F _c %	G = 50 cm			G = 100 cm		
			T _g	F _g %	P%	T _g	F _g %	P%
I sterk west	1338	47	403	88	30	72	95	5
matig west	759	27	47	10	6	3	4	0
zwak west	755	26	10	2	1	1	1	0
II noordelijk	1185	42	265	58	22	62	81	5
zwak	779	27	113	24	15	12	16	2
zuidelijk	888	31	82	18	9	2	3	0
III cyclonaal	1033	36	289	63	28	53	69	5
neutraal	769	27	113	24	15	15	20	2
anticyclonaal	1050	37	58	13	6	8	11	1

Een verdeling van de frequenties over de tercielen van de hoofdcomponenten is voor dezelfde 2852 kalenderdagen gegeven in Tabel 2.3. Hieruit blijkt de belangrijkheid van de sterkte der west-circulatie voor het optreden van de hoge HW-opzetten.

2.2.2 De persistentie van circulatietypen

Circulatietypen hebben een aanwijsbare *persistentie*: wanneer vandaag type 3 optreedt, dan is de kans op type 3 voor morgen duidelijk groter dan uit een willekeurige toedeling van typen over de dagen zou volgen. De mate van persistentie kan worden uitgedrukt in de gemiddelde duur van een periode met een bepaald weertype. Voor de typen met een verhoogde kans op een hoge HW-opzet zijn de gegevens hieronder vermeld.

Type	gem. duur (dagen)	gem. duur bij willekeurige toedeling
3	1,74	1,08
2	1,43	1,06
6	1,29	1,05

Ten opzichte van type 3 zijn de typen 2 en 6 aangrenzende circulatietypen: dit houdt in dat zij maar op één hoofdcomponent van type 3 verschillen. Het is dan ook niet verrassend dat veel overgangen van 3 naar 2 en 6 en omgekeerd voorkomen. Wanneer we deze drie typen als één type beschouwen, dan blijkt dat de gemiddelde duur stijgt naar 2,41 dagen tegenover 1,21 bij een willekeurige toedeling. Deze gemiddelde persistentie van 2,41 dagen ($\approx 4,6$ getijcycli) geeft een indicatie van de tijdschaal van de autocorrelatie van HW-opzet als gevolg van meteorologische oorzaken.

Kwantificering van de autocorrelatie in HW-opzet en de wijze, waarop ze is geëlimineerd, worden besproken in hoofdstuk 4.

2.2.3 Bovengrens hoogwaterstanden

Het ligt in de rede dat er een door de huidige klimatologische en morfologische omstandigheden bepaalde, maar onbekende bovengrens is aan de hoogwaterstanden aan onze kusten. In het „Deltarapport” is reeds aangetoond dat deze grens in 1953 niet is bereikt; de daar gegeven redenering is hieronder met enige aanpassingen en aanvullingen gevolgd.

De buitengewoon hoge waterstanden, die in de nacht van 31 januari op 1 februari 1953 in Zuidwest Nederland optraden (b.v. Vlissingen NAP + 455 cm, Hoek van Holland NAP + 385 cm) deden zich voor tijdens een springtij waarbij de maan zich dicht bij het apogeum bevond (VM: 30 jan., 0.44 h; maan in apogeum: 1 feb., 13 h); dienovereenkomstig was het springtij relatief zwak. Volgens de, voor dit rapport herziene, getijtafels (zie § 4.2.2) waren te Vlissingen de hoogste astronomische springtijhoogwaterstanden van medio januari en medio februari een 3½ dm hoger dan die van 31 januari op 1 februari; te Hoek van Holland was dit medio januari ook het geval (zie tabel 2.4 - volgens de oude getijtafels beliep het verschil daar 44 cm [Deltacommissie 1960, bl. 28]).

Aangezien er geen verband is tussen het optreden van zware stormvloed en de fase van de maan of de maansafstand [Rijkswaterstaat en KNMI 1961, par. 2.1.4 ... 2.1.6] moeten in ZW-Nederland dus aanmerkelijk hogere waterstanden mogelijk zijn dan die van 1 februari 1953. Overigens zijn naar in § 4.6 zal blijken aan beide stations tot 2½ dm hogere astronomische hoogwaterstanden mogelijk dan de hoogste standen van Tabel 2.4.

Tabel 2.4:

Enige astronomische getijgegevens
1953 (tijden in MET, standen in cm
t.o.v. NAP - SHW =
springtijhoogwater).

	VLISSINGEN	DATUM	HOEK VAN HOLLAND	
	+ 255 15h 14	17/1	16h 07 + 124	
slotgem.	+ 250 3h 36	18/1	4h 36 + 109	slotgem.
1951.0:	+ 255 16h 01		16h 53 + 123	1951.0:
HW + 190				+ 90 HW
SHW + 217	+ 217 15h 08	31/1	16h 07 + 103	+ 104 SHW
Top 1953:	+ 220 3h 25	1/2	4h 32 + 92	Top 1953:
+ 455				+ 385
	+ 253 14h 59	15/2	15h 47 + 108	
	+ 255 3h 19	16/2	4h 22 + 98	
	+ 254 15h 41		16h 31 + 109	
	+ 256 3h 58	17/2	5h 06 + 103	

Men kan zich vervolgens afvragen of zich aan onze Zuidwestkust ook een belangrijk hogere HW-opzet kan voordoen dan in 1953 is opgetreden. Het is namelijk alleszins aannemelijk dat hydrodynamische effecten het onmogelijk maken dat het water in de zuidelijke Noordzee een bepaalde stand te boven gaat; ook is de sleepkracht van de wind, die de opwaaiing teweegbrengt, aan een (onbekend) maximum gebonden. Een eerste indruk van de mogelijkheden van een hogere HW-opzet boden handmatige berekeningen naar de invloed van „manipulaties” met de depressie van 31 januari 1953, t.w. verplaatsing al dan niet tezamen met een verdraaiing en/of een verlenging van de stormduur. Al naar de combinatie zou hierdoor te Hoek van Holland een maximale opwaaiing van 42 dm tot („zeer onwaarschijnlijk”) 65 dm optreden [Weenink 1960^b]. In elk geval achtte men „langs onze kust een maximaal stormeffect van meer dan 5 m” niet onmogelijk, zij het ook met een „uiterst klein(e) kans van optreden” [Deltacommissie 1960, bl. 28]. Latere modelsimulaties met een tweetal verplaatste depressies [Timmerman 1977] komen wat lager uit, t.w. op 3,9 m te Hoek van Holland*, wat altijd nog aanzienlijk meer is dan daar op 1 februari

* waarmee daar mogelijk zou worden een stand van NAP + 540 à 545cm (overschrijdingskans, alhoewel onbekend, buitengewoon klein)

1953 aan grootste stormeffect (≈ 325 cm) en HW-opzet (293 cm) is opgetreden. Van de laatstgenoemde twee waarden is het grootste stormeffect, afgeleid m.b.v. alleen verschoven astronomische getijkrommen van 31 januari en 1 februari van dit station (zie § 2.1.2), gelijk aan dat in het „Deltarapport“ [Deltacommissie 1960, bl. 28].

2.3 Veranderingen in „uitwendige omstandigheden“

In het voorgaande zijn de uitwendige omstandigheden, die evenzevele randvoorwaarden vormen, stilzwijgend als onveranderd in de afgelopen 100 jaar beschouwd. Dit is echter onjuist voor wat betreft de gemiddelde zeeniveaus en de configuratie van bepaalde kustgedeelten; verder houdt men in brede wetenschappelijke kringen klimaatsveranderingen in de voor ons liggende decennia of eeuw voor waarschijnlijk [IPCC 1990]. In deze paragraaf komt e.e.a. nader ter sprake; de vraag of het stormklimaat in de afgelopen 100 jaar als onveranderd mag worden beschouwd wordt in § 4.3.5 behandeld.

2.3.1 Relatieve zeespiegelrijzing en invloeden grote werken

Sinds het begin van deze eeuw vormt in Nederland de relatieve rijzing van de gemiddelde zeestand t.o.v. het NAP-peil onderwerp van onderzoek. De Deltacommissie beval aan voor wat dit betreft rekening te houden met 15 à 20 cm/eeuw [Deltacommissie 1960, bl. 87,88]. Gerekend vanaf 1900 loopt deze rijzing voor onze belangrijkste getijstations uiteen van 15 tot 22 cm/eeuw [TO 1971-1980, par. 5.2.2 bl. XLII e.v.]. Hiervan is een klein, langs onze kust nog wisselend deel gevolg van bodemdaling [Lorenz et al. 1991]; het onderzoek hiernaar is echter nog niet afgerond. Verder is gebleken dat aan bepaalde Nederlandse peilmeetstations, waaronder *Vlissingen* en (althans na omstreeks 1935) *Harlingen*, de gemiddelde hoogwaterstanden relatief aanmerkelijk (een 50% of meer) sneller rijzen dan de gemiddelde zeestanden. Dit hangt rechtstreeks samen met een toeneming van gemiddelde tijverschillen; het is aannemelijk dat uitgevoerde waterbouwkundige werken (w.o. baggerwerken) hierop grote invloed hebben gehad [Van Malde 1984], maar ook zullen natuurlijke morfologische veranderingen daarbij een rol gespeeld hebben. Bijlage 1 toont het verloop in de tijd van de jaargemiddelden van 4 getijgrootheden te Vlissingen. In elk der gevallen is op de tijdreeks toegepast een vereffening naar een rechte (de trend) plus een sinusoïde met een periode van 18,6 jaar, die de invloed weergeeft van het nodale getij - de enige langperiodieke schommeling die zich duidelijk manifesteert in deze grafieken en, zij het minder uitgesproken, in overeenkomstige grafieken voor andere Nederlandse getijstations. In voorkomende gevallen dient men dan ook op zijn mogelijke invloed bedacht te zijn. Aandacht verdient nog de opmerkelijke sprong omstreeks 1887, die de bovenste drie grafieken van bijlage 1 vertonen; de oorzaak van die sprong, die voor HW 14 cm bedraagt, is niet duidelijk.

Uitvoering van grote waterbouwkundige werken heeft in bepaalde gevallen het getij min of meer abrupt beïnvloed. Zo heeft de *afsluiting der Zuiderzee* (28 mei 1932 een feit) de waterstanden en stromingen in de Westelijke Waddenzee aanmerkelijk beïnvloed. Daarbij is te Den Helder en Harlingen in een kort tijdsbestek het jaargemiddelde van

HW achtereenvolgens 7 en 16 cm hoger en dat van LW 8 resp. 24 cm lager geworden [o.a. Van Malde 1992, par. 2.3]. Verder komen in dit gebied na de afsluiting belangrijk hogere waarden van HW-opzet voor dan daarvoor; het verschil kan bij zeer zware stormvloedden te Den Helder oplopen tot ruim 50 cm en te Harlingen tot ruim 100 cm [Dronkers 1955]. Bovendien vond in de periode 1932-1980 een gestage aanpassing van het getij aan de nieuwe toestand plaats. Dit blijkt uit de voor Harlingen aangetoonde vervroegingstrend voor het hoogwater gepaard gaande met een verkorting van de rijzingsduur; van een aldaar ten opzichte van andere stations afwijkend verloop van het jaargemiddelde van HW is verder niets gebleken [Doekes 1985]. Na 1980 lijkt de genoemde trend voor het hoogwater aan te houden, echter zonder verdere verkorting van de rijzingsduur. Voor de beide stations Harlingen en Den Helder is het ondoenlijk de hoge HW's van voor 1932 te herleiden tot de toestand na 1932; bij gevolg kunnen bij het statistisch onderzoek voor *Den Helder* en *Harlingen* alleen de waarnemingen na mei 1932 in de beschouwingen worden betrokken. Voor de andere stations zijn de invloeden van uitgevoerde werken minder ingrijpend: Te *Hoek van Holland* begon omstreeks 1965 een relatief snelle rijzing van de gemiddelde hoogwaterstand, die ongeveer 15 jaar aanhield en een 1½ dm beliep; tezelfder tijd stagneerde de rijzing van de gemiddelde laagwaterstand.

Oorzaken van een en ander zijn de uitvoering van het Europoortproject, baggerwerken, voortgang en voltooiing van Deltawerken, zoutweringswerken in de Nieuwe Waterweg [TO 1961-1970, par. 4.2.4, bl. XL-XLIII; TO 1971-1980, par. 5.2.3, bl. XLVI]. Te *Delfzijl* vertoont het verloop der gemiddelde hoog- en laagwaterstanden twee discontinuïteiten n.l. rond 1960 (te verklaren uit het verdiepen der vaargeul naar Emden) en rond 1978 (veroorzaakt door de afsluiting van de oude westelijke havenmond ter plaatse). Hierdoor werd de gemiddelde hoogwaterstand met achtereenvolgens ong. 3 en 6 cm verhoogd en de gemiddelde laagwaterstand met (uiteindelijk) een 10 cm en 4 cm verlaagd [De Ronde 1984^b]. Het is aannemelijk dat al deze veranderingen te gering zijn geweest om de waarden van HW-opzet in enige mate van betekenis te hebben kunnen beïnvloeden. De uitkomsten van de toetsing van de onafhankelijkheid van astronomische hoogwaterstanden en HW-opzetten (§ 4.5) leveren een bevestiging van dit vermoeden.

Voor de vaststelling van tijdreeksen van HW voor een getijstation zijn correcties op trends en min of meer abrupte veranderingen ter plaatse uiteraard noodzakelijk. Voor de vaststelling van HW-opzet c.q. LW-opzet vereisen deze verschijnselen de nodige zorgvuldigheid bij de berekening van astronomische standen (zie § 4.2.2).

2.3.2 Het effect van klimaatschommelingen op de stormfrequentie

Uit historische meetreeksen en uit paleoklimatologisch onderzoek is het bekend dat het klimaat fluctueert op alle denkbare tijdschalen. Deze klimaatschommelingen hebben tot aanzienlijke fluctuaties in het zeewaterniveau geleid. De redenen zijn veranderingen in het mondiale volume van het landijs en thermische expansie van oceaanwater [Tennekens en Können 1990, bl. 19-26]. De huidige rijzing van het mondiale zeewaterniveau (15 cm/eeuw) heeft zich pas in de 19e eeuw ingezet als gevolg van het aflopen van de zogeheten „Kleine IJstijd“.

Klimaatschommelingen kunnen ook leiden tot een verandering in de stormfrequentie boven de Noordzee. De combinaties van zeespiegelstand en stormfrequentie bepalen de kans op hoge waterstanden. Er is consensus dat de uitstoot van CO₂ en andere broeikasgassen zal leiden tot een geforceerde opwarming van de aarde. Deze opwarming zal in de komende eeuw enkele graden bedragen en het grootst zijn in de poolgebieden. Op zich zal dit de zeespiegelrijzing 3-5 keer zo groot als nu doen worden [IPCC 1990; Tennekens en Können 1990, bl. 28-32]. Anderzijds is het onduidelijk hoe de verdeling der stormfrequenties bij een geforceerde klimaatverandering zal worden. De verkleining van de temperatuurgradiënt in de noord-zuid richting suggereert een afneming van de kans op extreme wind, hetgeen de kans op hoge waterstanden weer minder zou maken dan men op grond van alleen de voorziene zeespiegelrijzing zou verwachten, maar of dit ook daadwerkelijk zal gebeuren is niet zeker. De huidige computermodellen zijn nog niet in staat klimaatveranderingen op ruimtelijke schalen zoals die van Europa bevredigend te beschrijven. Gedetailleerde uitspraken over het verloop van het toekomstig klimaat bij CO₂-forcering en over het effect hiervan op de kans op hoge waterstanden zijn pas te verwachten als het inzicht in de wisselwerking tussen wolken en straling en die tussen oceaan en atmosfeer significant zal zijn toegenomen. Het zal nog jaren kunnen duren voordat de resultaten van dit onderzoek de bovenstaande, kwalitatieve uitspraken wezenlijk zullen kunnen doen verbeteren.

3. Aanpak

Alvorens voor zeekeringen enz. bepalend te stellen zeer hoge hoogwaterstanden langs statistische weg te berekenen dient men te zorgen dat het te gebruiken *gegevens*materiaal *homogeen* is (dat wil zeggen betrekking heeft op dezelfde klimatologische en hydraulische omstandigheden) *en autocorrelatievrij*, d.w.z. uit onderling onafhankelijke gegevens bestaat (waarvoor in den regel een selectieprocedure nodig zal zijn). Verder dient men gemotiveerde beslissingen te nemen over de te verwerken grootte of grootheden (HW en/of een opzet), de toe te passen (statistische) verdelingsfunctie(s) enz.

Deze aspecten zijn hier gememoreerd omdat zij in § 3.1 ter sprake komen; in § 3.3 worden zij uitvoeriger besproken.

3.1 De werkwijze der Deltacommissie

3.1.1 Korte beschrijving van de methode

Het meteorologische en wiskundig-statistische onderzoek ten behoeve van de bepaling van basispeilen en hoogwateroverschrijdingslijnen is in hoofdzaak gericht geweest op het station *Hoek van Holland*. Voor het homogeniseren en selecteren van het daarvoor beschikbare waarnemingsmateriaal zijn de gedachtengang en gedragslijn als volgt geweest [Van Dantzig et al 1960]:

- A. De oorzaak van een gevaarlijk hoogwater is een storm, opgewekt door een depressie, waarvan de kern zich verplaatst volgens de eigen, specifieke depressiebaan (zie ook § 2.2). Voor de periode 1898 ... 1956 zijn nu [Van der Ham 1960] voor *Hellevoetsluis* de HW- en LW-opzetten ≥ 160 cm opgezocht. Deze 69 scheve opzetten (zie § 2.1.2) bleken door 49 depressies (35 in de periode nov ... jan, 4 ervoor, 10 erna) te zijn veroorzaakt; 16 ervan brachten in totaal 19 maal te Hellevoetsluis een opzet ≥ 200 cm teweeg. Voor zover de depressiebanen van deze 49 depressies konden worden gereconstrueerd bleken van het westen komende de kernen dezer depressies zich ter hoogte van het Noordzeegebied als volgt te hebben verplaatst:
- op 10° W.L.: tussen 51° N.B. en 67° N.B. alle, t.w. 41 stuks;
 - op 0° W.L.: tussen 52° N.B. en 66° N.B. alle, t.w. 48 stuks;
 - op 7° O.L.: tussen 52° N.B. en 65° N.B. 46 stuks, d.i. alle op één na die vóór 7° O.L. zuidwaarts af boog.

Uit deze bevindingen van het KNMI is opgemaakt dat de kern van ledere potentieel gevaarlijke depressie oostwaarts gaande zich binnen de aangegeven aardrijkskundige breedteband (de „poort”) verplaatst*.

* de noordgrens van de „poort” staat op enkele plaatsen in het „Deltarapport” [Van Dantzig et al 1960, bl. 31; Van der Ham 1960, bl. 99] onjuist aangegeven.

- B. De statistisch te verwerken reeks werd nu samengesteld uit de hoogwater waarnemingen te *Hoek van Holland* door beperking [Van Dantzig et al 1960, bl. 21, 31 en 40] tot:
- de opwaaiingsgevallen veroorzaakt door de potentieel gevaarlijke depressies;
 - één hoogwaterstand (de hoogste) per depressie (dit ter vermindering van autocorrelatie);
 - hoogwaterstanden met een HW-opzet ≥ 50 cm;
 - hoogwaters in het stormrijke seizoen („winterperiode“).
- De beperking tot de „winterperiode“ was nodig om klaarblijkelijke inhomogeniteit door seizoensinvloeden te vermijden; gekozen is voor de maanden november ... januari. Vanwege het ontbreken van weerkaarten uit de Tweede Wereldoorlog bleef de verzameling verder beperkt tot de perioden 1888/89 ... 1938/39 en 1945/46 ... 1956/57 [o.c., bl. 40].
- Deze procedure staat bekend als de *selectie Van der Ham*. Waarom er tussen de daarbij gekozen perioden en de voor de vastlegging van „de poort“ beschouwde periode in tweeërlei opzicht verschil bestaat is niet duidelijk.
- Wel mag worden aangenomen dat de kennelijk verrichte reconstructie van depressiebanen uit de periode 1888-1898 op beperktere informatie berust dan per depressiebaan uit het aansluitende tijdvak ter beschikking stond.

De selectie Van der Ham omvat 63 stormseizoenen en 332 hoogwaterstanden, lopend van NAP + 97 cm t/m NAP + 385 cm [o.c., bl. 39, 48, 49]; ook dit materiaal was echter nog niet homogeen. Na statistische verwerkingen heeft het Mathematisch Centrum gekozen voor de laagste aanvaardbare ondergrens („drempel“), te weten NAP + 170 cm [o.c., bl. 22, 23, 51]. Aan de betrokken 166 waarnemingen (bijlage 2) is een exponentiële verdeling aangepast, aan de 69 jaarmaxima uit de volle periode 1888 ... 1956 bovendien (ter ondersteuning) een Gumbel-verdeling [o.c., bl. 19/20, 35/36 en 38]*. Dit leidde tot een basispeil van NAP + 513 cm (d.i. 14 cm hoger dan met de 129 waarnemingen met NAP + 180 cm als drempel). Verwerking van de HW-opzetten zou volgens een oriënterend onderzoek niet tot andere uitkomsten hebben geleid [o.c., bl. 20] terwijl een beperkte verkleining of vergroting der „winterperiode“ geen aanzienlijke invloed op de uitkomsten zou hebben [o.c., bl. 21]. Tenslotte is nog enige aandacht besteed aan de lognormale verdeling en een Pearson-verdeling zonder dat dit aanleiding gaf tot enigerlei bijstelling [o.c., bl. 27-29 en 53 resp. 42 en 56].

De uitgevoerde selectie leidt tot een steilere overschrijdingslijn dan die behorende bij alle waarnemingen. Elders in het „Deltarapport“ zijn lijnen van dit tweede type gepresenteerd. Daartoe behoort allereerst de „werklijn“ voor Hoek van Holland. Voor haar vaststelling [Deltacommissie 1960, bl. 28-32] zijn van dit station gebruikt alle HW's uit de periode 1888-1948, alle stormvloedstanden uit de periode 1864 ... 1958 en afgeleide stormvloedstanden uit de jaren 1859 ... 1863. Door de punten kon „met grote zekerheid een

* buiten de „winterperiode“ vallen 3 van 166 hoogwaterdata (1/2/1908, 31/10/1912 en 1/2/1953) en 18 van die der 69 jaarmaxima, t.w. 11 tussen 31/1 en 8/4 en 7 tussen 11/9 en 1/11 [o.c., bl. 38]

vloeiende overschrijdingslijn" worden getrokken. Voor het op zich onzekere verdere verloop is in eerste aanleg (dat wil zeggen nog voor het statistische onderzoek was begonnen) uitgegaan van NAP + 5 m als basispeil; overwegingen daarbij vormden een aantal kwantificeringen van mogelijke verhogingen der voor de stand van 1 februari 1953 bepalende componenten (vergelijk § 2.2.3). Met dit basispeil kon men later zijdens het Mathematisch Centrum - zij het schoorvoetend [a.w., bl. 30] - instemmen; het is vervolgens als uitgangspunt aangehouden. De „werklijn" zelf heeft verder geen rol gespeeld.

Behalve die „werklijn" zijn in het „Deltarapport" opgenomen bundels hoogwateroverschrijdingslijnen voor Nederlandse getijstations, vastgesteld door Rijkswaterstaat „na overleg met het Mathematisch Centrum en het KNMI" [a.w., bl. 30/32]. Aan deze lijnen liggen ten grondslag [Wemelsfelder 1960^a, bl. 63, 64]:

- a voor Hoek van Holland het basispeil;
- b beneden het grenspeil voor 9 stations (waaronder Vlissingen, Hoek van Holland en Delfzijl) de hoogwateroverschrijdingsfrequenties „van de ronde decimeterwaarden", bepaald voor het tijdperk 1901-1950. Voor 16 andere stations zijn met behulp van betrekkinglijnen overeenkomende frequentiegegevens afgeleid van één van die 9 stations. Voor ieder station is in de grafieken nog het slotgemiddelde van HW aangegeven [o.c., bl. 73] en wel als ondergrens van de betrokken overschrijdingslijn;
- c boven het grenspeil *alle* bekende stormvloedstanden. Daarvan zijn die
 - van voor 1900 met behulp van de per 10 jaar bepaalde gemiddelde HW's gecorrigeerd in verband met de relatieve rijzing van het gemiddelde HW. De periode, waarover deze stormvloedstanden bekend zijn varieert voor de beschouwde stations en beslaat maximaal 40 jaar [o.c., bl. 63];
 - aan de stations vanaf de Belgische grens t/m het Haringvliet vanwege het daar grote aandeel van het astronomisch getij herleid tot gemiddeld hoogwater ten einde „een meer homogene bewerking van het materiaal te verkrijgen" [o.c., bl. 64];
- d. een vloeiend verloop van die lijnen tot bij de zone der hoogste waarnemingen, alwaar de steilheid der lijnen „die ... aangegeven door de lijn van de geselecteerde populatie" (dat wil zeggen die voor Hoek van Holland zoals bepaald door het Mathematisch Centrum) „benadert". Voor deze lijngedeelten is rekening gehouden met de relevante hydraulische omstandigheden [o.c., bl. 73/74 resp. 76-80];
- e een vrijwel evenwijdig verloop der overschrijdingslijnen tussen de overschrijdingsfrequenties $P = 10^{-2}$ p.j. en $P = 10^{-4}$ p.j. [o.c., bl. 74/75].

Voor de stations aan de westelijke Waddenzee is een bijzondere gedraglijn gevolgd in verband met de - vanwege de afsluiting der Zuiderzee (1932) - korte beschikbare homogene reeksen [o.c., bl. 81/82].

Met betrekking tot de door de Rijkswaterstaat gevolgde werkwijze en gepresenteerde overschrijdingslijnen passen nog de volgende aanvullende opmerkingen:

- (i) de beide gepresenteerde overschrijdingslijnen voor Hoek van Holland met waarnemingsperioden 1889 resp. 1862-1950 [o.c., bl. 61 en 69] zijn vrijwel identiek en liggen tussen grenspeil en basispeil tot ruim 10 cm boven de eerdergenoemde „werklijn“;
- (ii) voor de ingetekende stormvloedstanden zijn geen selecties toegepast hoewel slechts $\frac{2}{3}$ ervan betrekking heeft op het „stormrijke seizoen“ (november...januari) en sommige kort op elkaar volgden [TO 1951-1960, Afdeling V]. De gedragsregels ad a, d en e zijn geacht dit bezwaar te hebben ondervangen;
- (iii) de toenmalige grenspeilen - voor de basisstations bepaald voor de periode 1921 ... 1936 (met aanpassingen voor Den Helder en Harlingen) en voor de overige stations van die der basisstations afgeleid met behulp van betrekkinglijnen 1901 ... 1940 [De Ronde 1984^a, bl. 3/4] - zijn op de betrokken overschrijdingslijnen met een „plus“ gemarkeerd. De meeste van deze plussen vallen niet op de lijn $P = 5 \cdot 10^{-1}$ p.j. Dit zou gevolg zijn van hun afronding op 5 cm [Wemelsfelder 1960, bl. 73], maar dit argument is niet overtuigend en gaat in elk geval niet op voor de op cm nauwkeurige grenspeilen der basisstations;
- (iv) in de grafieken zijn voorts aangegeven de stormvloedstanden van 1953 (voor de Noordnederlandse stations ook die van 1954), echter alleen voor de stations bezuiden Scheveningen met de bijbehorende overschrijdingskans. Volgens de korte toelichting op deze P-waarden [o.c., bl. 76/77] is in de grafieken een onderlinge samenhang van de punten voor 1953 nagestreefd. Het punt voor Hoek van Holland is met behulp van waarnemingen uit de periode 1862 ... 1951 bepaald [Rijkswaterstaat en KNMI 1961, par. 2.3.9];
- (v) de waarnemingsperioden (1901-1950 voor o.a. Vlissingen; 1862-1950 voor o.a. Hoek van Holland en Delfzijl - in beide tijdvakken zal het volle jaar 1950 zijn begrepen) staan alleen vermeld in of onder de diagrammen (echter dat voor Scheveningen, Katwijk en IJmuiden bevat ter zake niets).

Voor de goede orde zij nog vermeld dat men voor Hoek van Holland omstreeks 1960 als gemiddelde HW in „het stormrijke seizoen“ mag aanhouden NAP + 95 cm; het grenspeil was toen NAP + 242 cm.

3.1.2 Enkele kanttekeningen bij de werkwijze der Deltacommissie

Voor wat betreft de statistische behandeling van het waarnemingsmateriaal voor Hoek van Holland kan het volgende worden opgemerkt: In § 2.1.2 is er op gewezen dat voor elk getijstation de reeks hoge HW's een zekere autocorrelatie vertoont. Bij de „selectie Van der Ham“ laat men per depressie slechts één HW in aanmerking komen, wat de onderlinge afhankelijkheid binnen de beschikbare waarnemingsreeksen aanmerkelijk vermindert (vergelijk § 2.2). Toch bestaan tegen toepassing van deze depressiebaanselectie enkele bezwaren, die a.v. kunnen worden samengevat:

- a. Bij het bepalen van een depressiekern en van een depressiebaan speelt een subjectief element een rol. Toepassing van deze methode door een andere persoon zou zeer waarschijnlijk leiden tot een andere selectie.

- b. De gronden voor de definitie van „gevaarlijke depressies” zijn niet sterk. Het is zeker niet ondenkbaar dat in het „Deltarapport” depressies met bepaalde, zelden voorkomende banen ten onrechte zijn geëlimineerd.
- c. In een aantal gevallen zijn de standen bij opeenvolgende „tweeling-depressies” beide opgenomen, terwijl deze depressies niet onafhankelijk van elkaar optraden. Het „Deltarapport” zegt over de afhankelijkheid van stormen zelf: „Twee of drie dagen na een storm is de kans op het optreden van een volgende storm groter dan na enkele stormvrije dagen, en wel 2 à 3 maal zo groot” [Rijkoort 1960, bl. 143].
- d. De methode is bijzonder arbeidsintensief.

Een selectie op circulatie-patronen zou aan bovengenoemde bezwaren tegemoet komen; deze patronen zijn echter pas vanaf 1949 beschikbaar. In dit rapport is dan ook gekozen voor een andere selectiemethode (§ 4.3).

De bezwaren a ... c vormen uiteraard een punt van kritiek op de toegepaste selectie der HW's te Hoek van Holland. Bij de verdere door het Mathematisch Centrum gevolgde gedragslijn kunnen nog enige kritische kanttekeningen worden geplaatst bij

- de keuze van het stormseizoen;
- het achterwege blijven van correcties op de waarnemingen in verband met de relatieve zeespiegelrijzing enz;
- het kleine aantal beschouwde verdelingsfuncties.

Met betrekking tot de werkwijze, gevolgd door de Rijkswaterstaat, verdienen de volgende punten de aandacht:

- (i) Richtsnoer voor het vrijwel evenwijdige verloop der vastgestelde overschrijdingslijnen voor lage P (§ 3.1.1 ad d en e) was de lijn van het Mathematisch Centrum voor Hoek van Holland. Dit lijkt op zich niet consistent: die lijn berust op via vijf selecties geselecteerde HW's, de overschrijdingslijnen op geheel ongeselecteerde HW's (zie ook hier na onder (iv)). Afgezien daarvan is met dat richtsnoer stilzwijgend de uitgevoerde depressiebaanselectie als relevant voor de gehele Nederlandse kust beschouwd. Deze selectie nu stoelt op de grote HW- en LW-opzetten te Hellevoetsluis; echter grote HW- en/of LW-opzetten in het Zuidwesten des lands gaan lang niet altijd samen met soortgelijke grote opzetten in het Noorden en omgekeerd. De genoemde veronderstelde relevantie kan dus betwijfeld worden al laat de invloed van deze overweging zich niet kwantificeren [Van Urk en Van Malde 1977, par. 4.5.2 e.v.]. Men kan zich verder afvragen of de geschetste evenwijdigheid der overschrijdingslijnen te rijmen is met de verschillen in de kust configuratie bij de onderscheidene stations en met die in de morfologie van de betrokken kustgebieden.
- (ii) Niet consequent is dat (§ 3.1.1 ad d) correcties in verband met wijzigingen in de gemiddelde zeeniveaus (vermoedelijk naar de tienjarige gemiddelden 1901-1910) slechts zijn toegepast voor het verwerkte materiaal van voor 1900. Overigens zou volgens een oriënterend onderzoek [o.c., 1977, par. 4.4] toepassing van die correcties op alle HW's de basispeilen van Vlissingen en Delfzijl met slechts ong. ½ dm verhogen.

- (iii) De herleiding van de stormvloeden naar gemiddeld HW, toegepast voor de stations in ZW-Nederland (§ 3.1.1 ad d), houdt in dat de betrokken gedeelten van de HW-overschrijdingslijnen, althans ten naastebij, evenwijdig lopen aan de overeenkomstige gedeelten van de overschrijdingslijnen voor HW-opzet. Daarmee is theoretisch gesproken de procedure incorrect geworden omdat een verdeling van HW-opzet gekoppeld moet worden aan die van de astronomische hoogwaterstanden ter plaatse om de verdeling van HW vast te stellen.
- (iv) De geconstrueerde overschrijdingslijnen hebben een tweeslachtig karakter. Voor ieder getijstation geldt namelijk dat de overschrijdingslijn van de totale populatie aan HW-waarnemingen een andere lijn is dan die van door een geschikte en toereikende selectie ongecorreleerde HW's met flinke tot zeer hoge HW-opzetten. Beide lijnen vallen niet samen wat trouwens ook destijds werd onderkend [Wemelsfelder 1960^a, par. 4.0, bl. 73/74]. Toch is er naar gestreefd overschrijdingslijnen te construeren, die een combinatie vormen van beide, in beginsel niet verenigbare typen.

Volledigheidshalve zij ten slotte nog opgemerkt dat de overschrijdingslijnen in het „Deltarapport”, voor wat betreft het gebied der waarnemingen betrekking hebben op grofweg de situatie 1925; toekenning van een „situatiejaar” aan de basispeilen is minder eenvoudig.

3.2 Vergelijkbare onderzoeken in het buitenland

Onderzoek naar extreme waarden van hoogwaterstanden is en wordt wereldwijd gedaan. Voor wat betreft „Noordzeestations” is het buiten Nederland voornamelijk verricht in Engeland [o.a. Ackers & Ruxton; Pugh & Vassie; Graff; Tawn] en Duitsland [Führböter, Jensen]. *Ackers en Ruxton (1975)* verwerkten van Southend (linker Theemsoever nabij de mond) de hoogwaterstanden hoger dan hoog springtij over 41 jaar (vanaf 1929).

Daarbij bepaalden zij de astronomische hoogwaterstanden met behulp van een harmonische methode en daaruit de HW-opzetten. Vervolgens stelden zij voor beide grootheden afzonderlijk de verdeling vast (voor de HW-opzetten numeriek); door deze twee verdelingen als onafhankelijk te beschouwen (bij benadering juist) konden zij de verdeling der HW's afleiden. Deze verdeling levert voor kleine overschrijdingskansen overigens lagere standen dan uit de stormjaarmaxima (stormjaar: juni...mei, periode 1929...1973) zou volgen bij toepassing van de zgn. EV-verdeling = extreme-waarden-verdeling (waarvan de Gumbelverdeling een bijzonder geval is). In de studie is de relatieve zeespiegelrijzing verrekend.

Pugh en Vassie (1979) ontleedden het waarnemingsmateriaal van 7 Britse stations, waarvan drie - Southend, Aberdeen en Lerwick (Shetlandeilanden) - aan de Noordzee. Voor 5 stations beliep de periode der waarnemingen niet meer dan 9 tot 15 jaar (met onderbrekingen), van Southend en Newlyn (in de Engelse zuidwesthoek) werden van lange reeksen de waarnemingen van 19 jaar gebruikt. Per station stelde men met behulp van harmonische analyse de astronomische uurstanden vast, waaruit de uurlijkse afwijkingen

(opzetten en „afwaalingen“) volgden - uiteraard met een grote autocorrelatie en een gemiddelde gelijk aan nul. Voor de astronomische standen werd voor 19 jaar de verdeling vastgesteld, voor de gevonden afwijkingen een numerieke verdeling. Voor alle stations behalve Southend waren beide verdelingen onafhankelijk zodat hieruit de verdeling der uurlijkse waterstanden afgeleid kon worden. Voor Newlyn bleken de zo gevonden HW's met overschrijdingsfrequenties $P = 2 \cdot 10^{-2}$ p.j. en $P = 10^{-2}$ p.j. een 25 cm hoger dan volgde uit verwerking van 62 jaarmaxima, wat althans voor een deel verklaarbaar zou zijn uit de verwaarlozing der genoemde autocorrelatie.

Graff (1981) onderzocht van 67 getijstations (t.w. 65 Engelse, Welse en Schotse en 2 Ierse, gelegen aan de Ierse Zee), de jaarmaxima, per reeks in aantal uiteenlopend van 10 tot 133 (gemiddeld 45); de meeste dezer reeksen vertonen „gaten“. Schrijver paste de EV-verdeling toe (waar nodig onder aanpassing der waarden in verband met trends) en bepaalde daarmee overschrijdingslijnen ten opzichte van de *herhalings*tijd $T (= 1 : P)$. Opvallend is dat de overschrijdingslijnen voor bijna alle stations aan de Zuidelijke Noordzee (waaronder Sheerness met 133 jaarmaxima en King's Lynn aan de Wash met 119 jaarmaxima) vergeleken met die volgens de Gumbel-verdeling opwaarts gekromd zijn. Daarentegen zijn deze lijnen voor de andere stations (waaronder die aan de Engelse zuidkust) merendeels neerwaarts gekromd (naar verhouding het sterkst voor $T < 1$ jaar) en overigens bijna alle recht, overeenkomende met de Gumbel-verdeling.

Tawn (1988) verruimde de schatting volgens de grootste aannemelijkheid van een algemene extreme-waardenverdeling voor de r hoogste waarnemingen per kalenderjaar onder gelijktijdige verrekening van een lineaire of kwadratische trend; desgewenst kan r per jaar variëren, is een ondergrens mogelijk en/of worden alleen het aantal overschrijdingen van een bovengrens meegenomen. Het model is zonder trendverrekening gebruikt voor 12 jaar aan uurwaarnemingen te Lowestoft, waarna 5 stormduren (t.w. 1 u; 15 u; 30 u; 66 u en 174 uur) zijn geselecteerd en, aan de hand van een grafische aanpassing aan de geschatte verdeling, vergelijkingen zijn gemaakt voor $r = 5$ resp. 10. Door transformatie van een reeks jaarmaxima van het naburige Great Yarmouth kon gebruik worden gemaakt van 85 jaar waarnemingen onder verrekening van een kwadratische trend, terwijl $r = 5$ voor het jongste gedeelte der waarnemingen. Bij vergelijking met verwerking van jaarmaxima bleek de standaardafwijking in de geschatte overschrijdingswaarden tot een derde teruggebracht te zijn. De mogelijke afhankelijkheid van deze standaardafwijking van de geschatte (negatieve) vormparameter is overigens niet onderzocht.

Führböter (1979) ging in op de mogelijke tijdsafhankelijkheid van de verdeling der jaarmaxima van de HW's voor de stations Cuxhaven, Wilhelmshaven en Husum aan de Duitse Bocht (met - tot aan 1977 - 164, resp. 123 en 110 jaarmaxima) en wel voornamelijk met de exponentiële verdeling. Om de invloed van opeenhoping van hoge waarden uit te schakelen acht hij toepassing van vijfjaarlijkse maxima (per opeenvolgende 5 jaar één maximum) nodig. Zijns inziens is er sprake van tijdsafhankelijkheid, maar toetsing van de significantie ervan blijft achterwege. De invloed der relatieve zeespiegelrijzing is in de studie verwerkt.

Jensen (1985) gebruikte van 10 Duitse Noordzeestations de jaarmaxima (aantallen: van 50 tot 100) der hoogwaterstanden, die in

verband met hun trendmatige rijzing eerst zijn herleid naar het jaar 2000. Uit een tiental verdelingen en berekeningswijzen ervan is er via lineaire regressie (met de stand als afhankelijk veranderlijke) één gekozen, t.w. een Gumbel-verdeling. Daarbij werd groot gewicht toegekend aan de overeenstemming op het oog van de berekende frequentielijn met de hoogste waargenomen standen. De extrapolatie is bescheiden omdat volgens de schrijver deze slechts zinrijk is tot een herhalingstijd $T = 2 \times$ de lengte der waarnemingsreeks en scepsis met betrekking tot de homogeniteit over 100 jaar op zijn plaats is.

3.3 De huidige aanpak

3.3.1 De in beschouwing genomen locaties en hun waarnemingsperioden

Het onderzoek is verricht voor de (vanuit hydrologisch oogpunt) belangrijkste getijstations langs de Nederlandse kust, te weten:

- Vlissingen,
- Hoek van Holland,
- Den Helder,
- Harlingen en
- Delfzijl.

Deze aanpak verschilt dus van die der Deltacommissie, die voor wat het theoretische frequentie-onderzoek betreft zich alleen op Hoek van Holland heeft gericht. De reden voor de hier gekozen stationsuitbreiding ligt in het feit dat in de verschillende delen van de Noordzee en de estuaria zowel de getijregimes als de uitwerking van de meteorologische invloeden onderling afwijken. Het eerder geuite bezwaar met betrekking tot dit aspect van de statistische onderzoekingen in het „Deltarapport” - § 3.1.2 ad (l) - is hiermee ondervangen.

Gestreefd is naar verwerking van gegevens uit zo lang mogelijke waarnemingsperioden. Deze eindigen alle in het jaar 1985 en beginnen zo vroeg als praktisch mogelijk was, dat wil zeggen voor Den Helder en Harlingen vanwege de afsluiting der Zuiderzee in juni 1932 en voor de overige stations in de jaren tachtig der vorige eeuw.

Een nadere uiteenzetting met betrekking tot deze keuzen van stations en perioden treft men aan in § 4.1.1.

3.3.2 De te verwerken veranderlijke(n)

Vereisten voor het onderwerpen van een reeks waarden van een toevalsgrootheid aan de gekozen statistische analyses zijn dat

- 1e aan de generatie van die waarden eenzelfde proces ten grondslag ligt - anders gezegd: dat die waarden tot dezelfde statistische verdeling behoren („homogeniteitseis”);
- 2e wederzijdse beïnvloeding van opvolgende waarden verwaarloosbaar klein is. Dit betekent dat de waarden als onderling onafhankelijk, *autocorrelatievrij* mogen worden beschouwd.

Voldoet een gegeven reeks niet aan deze eisen dan kan het nog zeer wel mogelijk zijn door passende selecties de homogeniteit en onderlinge onafhankelijkheid op bevredigende wijze te verzekeren.

Overigens is men wel doende voor tijdreeksen met autocorrelatie theorieën op te stellen, maar deze zijn nog niet zover ontwikkeld dat zij in het onderhavige geval bruikbaar zijn. Toepassing van analysemethoden, opgesteld voor onafhankelijke waarnemingen, op een reeks met onderlinge afhankelijkheid van al dan niet opeenvolgende elementen leidt tot te lage extreme waarden en is dus niet toelaatbaar. De grootheden, die in het onderhavige geval voor statistische analyse in aanmerking lijken te komen zijn HW en/of een (rechte of scheve) opzet. Voorafgaand aan enige analyse dient te worden nagegaan of een of meer van deze grootheden - al dan niet na te zijn onderworpen aan selecties - aan beide genoemde voorwaarden kunnen voldoen. Opzetten aan onze kust worden vrijwel altijd door aanlandige wind en/of (relatief) lage luchtdruk veroorzaakt (§ 2.1.2). Waar de weersverwachting voor het Noordzeegebied en de oostelijk aangrenzende kustgebieden principiëel beperkt moet blijven tot een korte periode van zeg hooguit een week is een voorspelling van de opzet aan onze kusten 12 uur vooruit reeds met zeer forse onzekerheden behept. Verder is er ook geen enkel verband van zware stormvloed met bepaalde maansfasen of met de maansafstand (§ 2.2.3). Daar staat tegenover dat een weerkundige situatie, die belangrijke opzetten teweegbrengt, zich gedurende enkele getijcycli kan doen gelden (§§ 2.1.2 ... 2.2.2). Verder kunnen b.v. HW-opzetten beneden een bepaalde waarde wat hun statistische eigenschappen betreft slecht overeenkomen met die boven die waarde (nader hierover in § 4.3.2). Tegen deze versturende omstandigheden kan men met het oog op een statistische analyse echter afdoende maatregelen nemen in de vorm van geschikte selecties (zie § 4.3), waaronder een geëigende drempelwaarde. De slotsom luidt dan ook dat een (geschikt gekozen) opzetgrootte zich leent voor een statistische analyse.

Dat laatste lijkt niet zonder meer het geval voor HW. Deze grootte is immers per definitie de som van het bijbehorende astronomische hoogwater en HW-opzet en van deze twee is de eerste factor deterministisch bepaald door een ingewikkeld samenstel van cyclische schommelingen en daarmee lang tevoren voorspelbaar (zie § 2.1.1); dit schendt de eis van homogeniteit. Beperkt men echter de waarden van HW tot die behorende bij de (op de hiervoor aangeduide wijze) geselecteerde HW-opzetten dan is er sprake van een willekeurige toedeling van HW-opzetten aan elementen der reeks astronomische hoogwaterstanden.

Van de zo ontstane tijdreeks van HW is het astronomisch bestanddeel der afzonderlijke waarden niet voorspelbaar. Op grond hiervan mag de HW van deze tijdreeks als stochastische grootte worden beschouwd. Vanwege de geschetste willekeurige toedeling van HW-opzet aan astronomische hoogwaterstanden kan men ook het astronomische hoogwater, betrokken bij die toedeling, als stochastische grootte opvatten. De verdelingsfunctie van het astronomische hoogwater wordt dan een kansverdelingsfunctie. Wanneer men HW-opzet heeft gekozen voor statistische analyse is deze kansverdelingsfunctie van belang voor de daarna noodzakelijke koppeling van HW-opzet met het astronomische hoogwater.

Alvorens de mogelijke stochastische grootheden nader op hun voor en tegen te beoordelen dient eerst te worden vastgesteld wat de voor het gestelde doel meest gereede opzetgrootte is.

Op het eerste oog lijkt dit y_{\max} de grootste afwijking ten opzichte van de getijkrommen in een opwaaiingsperiode (zie fig. 2.2), te zijn. Haar verwerking leidt echter tot aanzienlijke problemen vanwege de mogelijke wisselwerking tussen de weersinvloeden op het waterniveau en het astronomisch getij (zie § 2.1.2). Doet die wisselwerking zich namelijk voor dan is y_{\max} mede afhankelijk van de fase van het getij (zij is zeer gevoelig voor het al dan niet optreden van vervroeging of verlaten van het getij). Hierdoor zou, zo men toch een statistische analyse van y_{\max} zou verrichten, de daarna noodzakelijke koppeling van haar uitkomsten met astronomische standen in hoge mate worden bemoeilijkt.

Overigens zou de vaststelling van y_{\max} het nodig maken voor alle beschouwde stations harmonische analyses uit te voeren voor een groot aantal deelperioden uit het tijdvak van zeg 1885 t/m 1985 resp. (voor Den Helder en Harlingen) 1932 t/m 1985. De daarmee gemoeide arbeid is evenwel dermate omvangrijk dat dit op grote praktische bezwaren stuitte. Op grond van beide overwegingen tezamen vervalt y_{\max} als te verwerken veranderlijke.

Daarmee blijven als mogelijkerwijs bruikbare opzetgrootheden over HW-opzet en LW-opzet. De eerste komt uiteraard voor nadere beschouwing in aanmerking; zou nu voor een willekeurig station LW-opzet tot dezelfde verzameling behoren als HW-opzet dan zou men door mede gebruik te maken van de LW-opzetten de representativiteit van de verzamelingen kunnen verbeteren (de omvang van het materiaal zou er overigens niet belangrijk door worden vergroot vanwege de eis van minimalisering der auto-correlatie). Om e.e.a. na te gaan zijn voor de beschouwde stations de overschrijdingslijnen bepaald zowel van (de ongeselecteerde waarden van) HW-opzet als van (die van) LW-opzet. Over de periode 1933 ... 1980 bleek voor Den Helder het verschil tussen beide lijnen niet significant en bleek voor Harlingen het beeld niet geheel duidelijk te zijn; daarentegen werd voor Vlissingen gevonden HW-opzet $\approx 0,75$ LW-opzet en voor Hoek van Holland LW-opzet \approx HW-opzet + 20 à 25 cm bij $P < 10^{-1}$ p.j. [De Ronde 1985]. Voor Delfzijl tenslotte overtrof over de perioden 1901...1950 plus 1960...1980 bij $P < 10^{-1}$ p.j. HW-opzet met ong. 20 cm LW-opzet, terwijl voor $P > 10^{-1}$ p.j. merkwaardigerwijs het omgekeerde het geval was [Doekes 1986, bl. 2 en bijlage 2]. Klaarblijkelijk zijn de waterdiepten voor de verhouding tussen beide opzetten niet alles-bepalend en spelen ook zaken als vervroeging van het getij en kombergingsverhoudingen in het voorliggende zeegebied een rol. Hoe dit ook zij, het is duidelijk dat HW-opzet en LW-opzet voor het onderhavige onderzoek niet tezamen mogen worden beschouwd.

Gelet op de weergegeven bevindingen en overwegingen is voor het in te stellen onderzoek klaarblijkelijk HW-opzet de enige relevante opzetgrootte.

3.3.3 Beoordeling van HW en van HW-opzet als stochastische grootheden

De keuze van HW als te verwerken veranderlijke, zoals gedaan door de Deltacommissie, biedt de volgende voordelen:

- a. de benadering is rechtstreeks: ieder gezocht peil is immers een HW;
- b. HW-waarnemingen zijn verhoudingsgewijze nauwkeurig (fouten-

bronnen: stationsonvolkomenheden, instrumentele fouten, bedieningsfouten alsmede de onnauwkeurigheden van de waarneming, in de nulpunthandhaving en bij de bewerking); als standaardmeetafwijking van de vastgestelde meetwaarde wordt 2,5 cm aangehouden [Van der Made 1987, hoofdstuk 3].

Bij deze keuze zijn verder nog de volgende aspecten van belang:

1e De noodzakelijke selecties ter verzekering van homogeniteit en ter wering van autocorrelatie betreffen deels de bestanddelen HW-opzet van de waarnemingen en deels de waarnemingen zelf. Ter toelichting het volgende:

– de selecties via HW-opzet zijn:

- de selectie tegen autocorrelatie (zie § 4.3.4);
- de invoering van een - lage - drempel van HW-opzet om oneigenlijke waarden uit te sluiten (zie § 4.3.2).

HW's behorende bij de hierbij vervallen HW-opzetten blijven verder buiten beschouwing;

– de verdere selectie van de waarnemingen zelf geschiedt door invoering van een drempel voor HW en is nodig omdat de verdeling van lage HW's afwijkt van die van hogere (zie § 4.3.2).

De invoering van twee drempels lijkt te leiden tot minder waarnemingen dan toepassing van één drempel (zoals bij verwerking van HW-opzet). Of zich dit per paar parallele gevallen ook werkelijk voordoet hangt af van de aangehouden hoogten van de drempels (van HW resp. HW-opzet); is dit inderdaad het geval dan zal uit de verdere verwerking blijken in hoeverre dit een bezwaar is.

Zoals reeds in § 3.1.1 vermeld heeft destijds het Mathematisch Centrum voor Hoek van Holland ook twee drempels aangehouden ($HW \geq NAP + 170$ cm; $HW\text{-opzet} \geq 50$ cm).

2e Het jaargemiddelde van HW vertoont in samenhang met de zeespiegelrijzing een rijzende trend, welke voor vele Nederlandse (en Duitse) getijstations sinds ten minste ettelijke decennia die van de gemiddelde zeestand ter plaatse overtreft; bovendien komen in de grafieken van HW en van andere zeeniveaus als gevolg van uitgevoerde werken ook discontinuïteiten voor (§ 2.3.1). Beide verschijnselen zijn manifestaties van veranderingen in het verticale getij; de noodzakelijke correcties hierop, die per station moeten worden vastgesteld, zijn met enige onzekerheid behept.

Bij de *statistische verwerking van HW-opzet* kunnen de volgende kanttekeningen worden geplaatst:

- Er wordt slechts één drempel toegepast, wat uitzicht biedt op meer beschikbare waarden dan voor HW.
- Gelet op het relevante generatieve mechanisme - de versturende weersomstandigheden - is HW-opzet statistisch gesproken representatiever dan HW.
- De invloed van de rijzing van de jaargemiddelden van HW (zie § 2.3.1) is bij de hier gevolgde berekeningswijze geneutraliseerd.

N.B.: dit laatste is alleen juist mits

- de astronomische hoogwaterstanden op consistente wijze zijn berekend (om aan deze voorwaarde te voldoen zijn zij herberekend - zie § 4.2.2);
- HW-opzet onafhankelijk is van het astronomisch getij (dit bleek het geval te zijn - zie § 4.4).

- De astronomische hoog- en laagwaterstanden, berekend met behulp van de culminatiemethode, hebben betrekking op een onbekende „toestand van middelbaar weer” (§ 2.1.1). Daardoor zijn zij behept met onbekende onnauwkeurigheden, die tezamen met de onnauwkeurigheid van HW de onnauwkeurigheid van HW-opzet bepalen; bij gevolg zal de onnauwkeurigheid van HW-opzet aanmerkelijk groter zijn dan die van HW.
- Om het gevraagde antwoord te verkrijgen moet(en) de vast te stellen verdelingsfunctie(s) voor HW-opzet gekoppeld worden aan de (deterministisch bepaalde) verdelingsfunctie van het astronomische hoogwater, wat verre van eenvoudig is - zie § 5.2 (slot), § 5.3 („Methode 4”) en § 6.6.

Samenvattend kan gesteld worden dat HW en HW-opzet in aanmerking komen als statistisch te verwerken variabele en dat aan de keuze van beide voor- en nadelen kleven. Intuïtief zou men aan HW-opzet de voorkeur geven vanwege zijn bij uitstek stochastische karakter. Er zijn echter geen theoretische bezwaren om de HW's als uitgangspunt te nemen voor statistische analyse. In het onderhavige onderzoek zijn dan ook beide benaderingen toegepast. Vanwege de moeilijkheden met de koppeling van de verdelingsfunctie van extreme HW-opzetten met die van de astronomische hoogwaterstand is het grootste deel van het onderzoek verricht met HW als statistisch te verwerken veranderlijke.

4. De voorbereidingen van het waarnemingsmateriaal

In dit hoofdstuk komen achtereenvolgens ter sprake:

- de gemaakte keuzen van stations en waarnemingsperiodes;
- de vaststelling van de bestanden van de „uitgangsgegevens” (HW, HW-opzet, astronomische hoogwaterstanden);
- de uitgevoerde selecties voor het verkrijgen van homogene autocorrelatievrije tijdreeksen (zie § 3.3.2);
- enige aanvullende verwerkingen en opmerkingen.

De uitgevoerde selecties worden besproken in een volgorde van toenemende verbijzondering.

In § 3.3.2 is uiteengezet dat alleen de grootheden HW en HW-opzet in aanmerking komen voor statistische analyses. Bij de meeste toegepaste methoden stuit echter de nodige koppeling van hiermee berekende extreme waarden van HW-opzet aan de verdeling van dit astronomisch hoogwater op grote theoretische moeilijkheden en praktische bezwaren.

Als gevolg hiervan zijn de meeste statistische verwerkingen (hoofdstukken 5 en 6) alleen verricht met HW als veranderlijke.

Daarentegen hebben de uitgevoerde selecties vrijwel alle betrekking op HW-opzet; de reden daarvan is dat HW-opzet niet alleen het voornaamste bestanddeel vormt van hoge tot zeer hoge waterstanden, maar in die hoedanigheid ook behept kan zijn met inhomogeniteiten en belangrijke autocorrelatie, waarmee hij dan HW besmet (zie verder §§ 3.3.2 en 3.3.3). Tezamen brengen de uitgevoerde selecties voor elk der beide grootheden een ingrijpende reductie van het eerder vastgestelde gegevensmateriaal teweeg.

De aanvullende verwerkingen betreffen achtereenvolgens HW-opzet (toetsing van zijn onafhankelijkheid van het astronomische hoogwater), HW (zijn herleiding naar de toestand 1985) en het astronomische hoogwater (vaststelling kansdichtheidsgrafiek per station).

4.1 Keuze stations en waarnemingsperiodes

Maatstaf voor een stormvloed aan onze kust is het bereiken of overschrijden van het grenspeil aan tenminste een van de zes (tot 1 november 1964 zeven) basisstations tijdens de betrokken hoogwaterperiode (§ 2.1.2). Bij de keuze van deze basisstations is uiteraard een verdeling langs de kust nagestreefd, die past bij de wisselende ligging en grootten van door stormvloeden getroffen gebieden. Daarnaast speelden bij die keuze ook een rol de verschillen in het karakteristieke getijverloop langs onze kust, de daarmee gegeven tijdsverschillen tussen de astronomische hoogwaterstanden aan de getijstations en de uiteenlopende aard van de opwaaiingsverschijnselen in de onderscheidene kustgedeelten. Gezien deze overwegingen lag het op zich in de rede de huidige basisstations bij het onderhavige onderzoek te betrekken. Voor IJmuiden is dat echter niet gedaan omdat het aannemelijk leek dat voor de keuzen van verdelingen, selectieprocedures enz. met 5 stations kon worden volstaan. Een bijkomende overweging was dat de onnauwkeurigheid van HW's ter plaatse als gevolg van aldaar veelvuldig optredende slingeringsver-

schijnselen aanmerkelijk groter kan zijn dan die voor de andere stations. In § 7.10 zal op dit punt nog worden teruggekomen. De betrouwbaarheid van de uitkomsten van een statistische analyse neemt toe met het aantal waarnemingen, dat erbij gebruikt is. Daarom is voor elk der vijf resterende basisstations een zo lang mogelijke waarnemingsperiode beschouwd. Aangezien te Den Helder en Harlingen zowel het astronomische getij als het karakter van de opwaaiing belangrijk is veranderd door de afsluiting der Zuiderzee (28 mei 1932) kon voor deze stations deze waarnemingsperiode eerst op 1 juni 1932 beginnen. Voor de andere basisstations valt het begin van de bedoelde waarnemingsperiode niet samen met de aanvang der waarnemingen aan het betrokken stations maar met het tijdstip van waaraf leemten als gevolg van ontbrekende waarnemingen gevuld konden worden door met voldoende betrouwbaarheid gegiste waarden (zie § 4.2.1). Voor alle stations eindigt de beschouwde periode met het kalenderjaar 1985. De bedoelde perioden zijn voor:

Vlissingen	1 juli 1881 1985
Hoek van Holland	1 aug. 1887 1985
Den Helder	1 juni 1932 1985
Harlingen	idem
Delfzijl	1 mrt. 1881 ... 1985.

Enkele in een vroeg stadium van deze studie verrichte deelonderzoekingen voor een of meer der stations Vlissingen, Hoek van Holland en Delfzijl strekten zich overigens uit over kortere dan de zojuist vermelde perioden.

4.2 Samenstelling bestanden

4.2.1 Het HW-bestand

Over de gekozen perioden staan de standen (in cm t.o.v. NAP) en tijdstippen van hoog- en laagwaters aan getijstations afgedrukt in de „Jaarboeken der Waterhoogten enz.” van de Rijkswaterstaat. Deze gegevens zijn voor de basisstations (zoals voor alle belangrijke getijstations) de uitkomsten van bewerkingen van peilschaalregistraties en waar deze laatste onderbrekingen vertonen vrijwel altijd van de mede beschikbare plaatselijke peilschaalaflezingen. Wanneer ook deze aflezingen ontbreken worden de ontbrekende waarden in het algemeen bijgegist met behulp van de gegevens van omliggende stations. Geschiedde dit vroeger door middel van grafische vergelijking en slechts voor korte hiaten, sinds midden jaren tachtig past men hiervoor toe de methode der meervoudige lineaire regressie, waarmee onder bepaalde voorwaarden ook langduriger onderbrekingen kunnen worden overbrugd. Wat de basisstations betreft hebben zich enkele belangrijke leemten voorgedaan tijdens de Tweede Wereldoorlog (specificatie: [TO 1961-1970, bl. XVI Tabel 2]), alsmede twee daarbuiten. Al deze ontbrekende waarden zijn bijgegist; voor het onderhavige onderzoek betreft dit vanwege het gekozen lange stormseizoen (1 oktober t/m 15 maart, zie § 4.3.3) voor:

Vlissingen de periode	14 okt. - 24 nov. 1944;
Hoek van Holland	1 okt. - 8 dec. 1945;
Den Helder	1 jan. - 16 mrt. 1972;
Delfzijl	1 nov. 1902 - 1 feb. 1903 en 1 - 16 mrt. 1945.

De hoog- en laagwatergegevens van de gekozen vijf basisstations uit de in § 4.1 vermelde perioden (en van IJmuiden) zijn alvorens nader te worden verwerkt nog aan de volgende bewerkingen onderworpen:

- a. zij zijn nog eens op onjuistheden nagelopen en waar nodig gecorrigeerd;
- b. alle stormvloedhoogwaters zijn nader gezien teneinde verzekerd te zijn van een consistente gedragslijn bij de behandeling van secundaire schommelingen. Bij een aangetroffen afwijking ≥ 5 cm volgde een correctie.

Alle zo vastgestelde gegevens werden vervolgens opgeslagen in een elektronisch bestand; voor het verdere onderzoek is daarvan voornamelijk het gedeelte met de tijdreeksen van de hoogwaterstanden der vijf gekozen stations (het „HW-bestand”) gebruikt.

Ter toelichting op de bewerking ad b diene het volgende:

De genoemde secundaire schommelingen kunnen worden onderscheiden [Wemelsfelder 1960^b] in:

- *resonantieslingeringen*, die zich voornamelijk voordoen in bepaalde havens en havenkommen (krachtig in de voorhaven van IJmuiden en bij de Rozenburgsluis). Zij spelen voor de vijf gekozen basisstations geen rol van betekenis;
- *bui-oscillaties*, in series van 2 tot 20 stuks optredende lange golven met golfhoogten van 50 en meer cm en met perioden van enkele minuten tot meer dan een uur. Opgewekt op zee verplaatsen zij zich in het algemeen in een richting min of meer haaks op de kust, waarbij zij zich aan een niet zeer groot kustvak manifesteren;
- *buistoten*: eenmalige verheffingen (a.h.w. halve golven) met perioden en hoogten als de bui-oscillaties; zij verplaatsen zich min of meer evenwijdig aan de kust.

De inzichten met betrekking tot de weerkundige omstandigheden, waaronder buioscillaties en buistoten worden opgewekt [o.c.] doen met wat nadere specificaties nog steeds opgeld, die met betrekking tot de hydraulische generatiemechanismen zijn jonger en deels recent [Vogel 1992, hoofdstuk 3].

Bui-oscillaties worden toegeschreven aan secundaire stormverschijnselen, macroscopische turbulenties - waaronder de neerwaartse beweging van koude luchtbellens [a.w., bl. 18] - e.d., *buistoten* aan verschijnselen als het passeren van een zware bui, van een front - in het bijzonder een scherp koufront, veelal gepaard met hevige windvlagen [a.w., bl. 17/18] - of van een zware plaatselijke bui, aan het losbarsten van onweer of aan hoosachtige luchtbewegingen. Bij de bewerking van peilregistraties ter vaststelling van de hoog- en laagwatergegevens zijn in voorkomende gevallen de bui-oscillaties geëlimineerd, wat als regel zonder veel willekeur mogelijk is. Voor optredende buistoten is zo'n gedragslijn niet zonder meer gevolgd (voor de reeksen gebruikt voor het „Deltarapport” zou dit wel het geval geweest moeten zijn [Wemelsfelder 1960^b, par. 4.2 en 6.0], maar bij het nalopen van de HW's bleek dit slechts voor een deel ervan op te gaan). Eliminering van buistoten is namelijk in bepaalde gevallen verre van eenvoudig en veelal ook niet reëel. Door de bij het nalopen gevolgde consistente gedragslijn is van de HW's per station de onderlinge vergelijkbaarheid gewaarborgd.

N.B: bij het dimensioneren van zeeweringen moet men wel degelijk rekening houden met het optreden van *bui-oscillaties* en *buistoten* en wel door een toeslag op het ontwerp. Hiertoe dient de „toeslagregel van Wemelsfelder” [o.c., par. 5.3 bl. 128, voorts: Van Maide 1990, par. 4.2 en 7.2].

4.2.2 Bestanden van astronomische hoogwaterstanden en HW-opzetten

De astronomische hoog- en laagwaters aan de vijf stations (standen en tijdstippen) zijn vanaf 1898 (Den Helder vanaf 1900) gepubliceerd in de (tijdens de Tweede Wereldoorlog onderbroken) reeks jaaruitgaven „Getijtafels voor Nederland”. Zij zijn voor de jaren tot en met 1985 (de oorlogsjaren inbegrepen) berekend met de culminatiemethode (§ 2.1.1). Omdat die methode in de loop der jaren enige malen is herzien dienden, teneinde homogeen materiaal te verkrijgen, de astronomische hoogwaterstanden opnieuw te worden berekend. De toepassing van de harmonische analyse hiervoor stuitte op bezwaren gezien de buitensporige hoeveelheid ermee gemoeide arbeid (§ 3.3.2); daarom zijn die berekeningen verricht volgens de voor de getijtafels 1985 gebruikte versie der culminatiemethode [De Ronde 1984^b].

De bij deze herberekening gebezigde analyse-perioden zijn 9 jaren lang, dat is ongeveer de halve lengte der nodale periode (§ 2.1.1 ad e). Zij zijn wanneer mogelijk zo gekozen dat halverwege zo'n periode het maansbaanvlak de middenstand innam; in een enkel geval werd deze gedragslijn doorkruist door discontinuïteiten ten gevolge van uitgevoerde werken. De gebruikte bestaande programmatuur voor analyse en voorspelling met behulp van de culminatiemethode voorzag verder niet in toepassing op de jaren vóór 1901. Dit is gevolg van het feit dat het jaar 1900 geen schrikkeljaar is en daarmee de vierjarige reeks schrikkeljaren, die onderdeel uitmaakt van de programmatuur, doorbreekt. Met enige aanpassingen kon de programmatuur toch geschikt gemaakt worden voor de jaren voor 1900. Voor december 1899 en het jaar 1900 was dit echter niet mogelijk dan met onevenredig veel inspanning, waarvan moest worden afgezien. Daarmee verviel voor de culminatiemethode ook één analyseperiode (die het eeuwjaar zou bevatten), zodat niet alleen voor de aangeduide periode van 13 maanden, maar ook voor het daaraan voorafgaande en het erop aansluitende tijdvak andere oplossingen moesten worden gevonden. Een overzicht van de toegepaste berekeningsmethoden met hun analyseperioden biedt bijlage 3.

Alle zo berekende gegevens (standen en tijdstippen der astronomische hoog- en laagwaters) zijn eveneens opgeslagen in een electronisch bestand; daarvan is in het bijzonder het gedeelte met de tijdreeksen der astronomische hoogwaterstanden aan de vijf stations (het „bestand astronomische hoogwaterstanden”) voor het verdere onderzoek van belang. Door nu per hoog- c.q. laagwater de astronomische van de opgetreden hoog- c.q. laagwaterwaterstand af te trekken ontstonden tenslotte (zie § 2.1.2) nog twee series tijdreeksen, het „HW-opzetbestand” en het „LW-opzet-bestand”. Het laatstgenoemde bestand is alleen gebruikt om na te gaan of LW-opzet al dan niet bij de statistische analyses kon worden betrokken; naar in § 3.3.2 uiteengezet was de uitslag hiervan negatief.

De verschillen tussen de herberekende en de vroeger aangehouden astronomische hoogwaterstanden zijn niet verwaarloosbaar. Van de 332 HW's te Hoek van Holland, die het Mathematisch Centrum aanvankelijk had geselecteerd in het raam van zijn onderzoek ten behoeve van het „Deltarapport” (zie § 3.1.1) traden er 276 op vanaf het aangehouden stormseizoen 1900/1901. De herberekening bleek in 210 gevallen hiervan tot een hoger en in 57 gevallen tot een lager

astronomisch hoogwater te leiden; de gemiddelde verhoging beliep 5 cm [Samenwerkingsverband Basispeilen 1986, par. 2.4.1].

Voor wat het verdere onderzoek betreft diende de in deze onderparagraaf beschreven exercitie twee oogmerken:

- a. het per station (in de vorm van het HW-opzet-bestand) verkrijgen van het uitgangsmateriaal voor de selecties, die vooraf dienen te gaan aan een statistische analyse van hetzij HW hetzij HW-opzet;
- b. het per station kunnen verwerken van HW-opzet in een statistische analyse, waarvan de verdelingsfunctie, zoals reeds eerder ter sprake kwam, dan weer gekoppeld moet worden aan die van de astronomische hoogwaterstand ter plaatse. (De kansdichtheidgrafieken van de astronomische hoogwaterstanden, die ten grondslag liggen aan hun verdelingsfuncties, worden behandeld in § 4.6).

4.3 De uitgevoerde selecties

4.3.1 Probleemstelling

Reeds enkele malen is ter sprake gekomen (o.a. in § 3.3.2) dat voor de statistische analyse van HW of van HW-opzet het ten zeerste gewenst is dat de gegevensreeks bestaat uit elementen, die

- onderling onafhankelijk zijn en
- tot éénzelfde statistische verdeling behoren („homogeen” zijn).

De per station samengestelde tijdreeksen van HW en HW-opzet (§ 4.2) voldoen op zich niet aan de beide genoemde eisen. De oorzaken hiervan liggen voor wat betreft HW-opzet in meteorologische en hydraulische factoren en voor wat betreft HW in HW-opzet en het deterministische karakter van het astronomische getij (§§ 2.1.2 en 2.2).

Door middel van geëigende selectiemethoden kunnen de reeksen geschikt worden gemaakt voor statistische analyse; dit is goed mogelijk voor HW-opzet, echter - naar reeds in § 3.3.2 betoogd - minder goed voor HW, althans wanneer men die grootheid op zichzelf beschouwt, zodat zijn selectie goeddeels op die van HW-opzet teruggaat. Iedere selectie vermindert het aantal elementen; de uitdaging is uiteraard om alleen de versturende elementen te verwijderen en niet meer dan die omdat anders relevante informatie verloren gaat, wat de analyse zou kunnen beïnvloeden.

Om per station de homogeniteit van *HW-opzet* zo goed mogelijk te verzekeren is nadere beschouwing nodig zowel van zijn consistentie over het volle gebied van zijn waarden als van de aard van zijn verdelingen over verschillende jaargedeelten. Dit brengt dan mee de invoering van achtereenvolgens een drempel en een stormseizoen (beide ook toegepast in het „Deltarapport” en wel voor Hoek van Holland - zie § 3.1.1). Aansluitend kan onderlinge afhankelijkheid (autocorrelatie) worden onderdrukt door toepassing van een geschikte selectiemethode naar het tijdsaspect. Voor HW bestaan de selecties allereerst uit de beperking van zijn waarden tot die, welke behoren bij de geselecteerde HW-opzetten. Een aanvullende selectie door invoering van een drempel vloeit weer voort uit het streven naar consistentie over het gehele gebied van de waarden van HW.

In de hierna volgende onderparagrafen wordt verslag uitgebracht van de verrichte selecties. Daarop aansluitend is aan de hand van de na de selecties ontstane tijdreeksen van HW-opzet nagegaan of het stormklimaat zelf geacht mag worden homogeen te zijn geweest in de beschouwde periode.

4.3.2 Eerste homogeniseringsselecties: invoering drempels

De tijdreeksen van *HW-opzet* zijn behept met een zekere ruis. De standaardafwijking hiervan kan worden geschat uit de standaardafwijkingen van de twee bepalende grootheden, HW en het astronomisch hoogwater:

- a Voor de onnauwkeurigheid van de HW-waarden zelf mag als standaardafwijking $\sigma_{HW} = 2,5$ cm worden aangehouden (§ 3.3.1 ad b). Afgezien daarvan heeft de luchtdruk invloed op HW (1 hectoPascal = 1 mbar = 1 cm waterkolom); zonder enige bijkomende windopzet kan die luchtdrukinvloed oplopen tot 20 à 30 cm. Stel op grond hiervan de standaardafwijking door die luchtdrukvariaties $\sigma_l = 10$ cm.
- b Onvermijdelijkerwijze kleven ook aan de berekende astronomische hoogwater standen onnauwkeurigheden (zie § 2.1.1). Volgens een voorzichtige (waar schijnlijk wat ruime) schatting zou de standaardafwijking hiervan 10 cm kunnen belopen, dus $\sigma_a = 10$ cm.

Veronderstelt men nu dat de drie genoemde ruisfactoren onderling onafhankelijk zijn (wat bij benadering juist lijkt) dan geldt voor σ_0 , de ruis voor HW-opzet:

$$\sigma_0 = \sqrt{(\sigma_{HW}^2 + \sigma_l^2 + \sigma_a^2)} = \sqrt{206,25} = 14,4 \text{ cm.}$$

Houdt men nu een drempel aan van $2\sigma_0$ - afgerond: van 30 cm - dan heeft men een vrij grote zekerheid dat de tijdreeks van HW-opzet geen gevallen zonder opwaaiing meer zal bevatten. Aangezien windopzet nu juist een zeer wezenlijk element vormt van alle belangrijke waarden van HW-opzet is met die drempel van 30 cm de betrokken tijdreeks van HW-opzet wat dit aspect betreft gehomogeniseerd. Deze drempel moet als een minimumwaarde worden gezien. Welke drempelwaarde in concreto moet worden aangehouden zal veelal per verwerkingsgeval moeten worden bepaald.

Beschouwt men *HW* als stochastisch veranderlijke dan dient men te bedenken dat in zijn waarden bij stormvloeden het aandeel van HW-opzet, hoezeer ook variabel, van doorslaggevend belang is (anders was er namelijk geen sprake van een stormvloed). Dat brengt met zich mee dat, zoals reeds in § 3.3.2 uiteengezet, bij de samenstelling van de tijdreeksen van HW alleen die waarden van HW beschouwd moeten worden, waarvan de HW-opzetten zijn overgebleven na de uitvoering van alle selectieprocedures.

De zo gevormde tijdreeksen zullen echter nog niet zonder meer voldoen aan de eis van homogeniteit; immers voor elk van die reeksen geldt dat gemiddeld het astronomisch hoogwater een verhoudingsgewijs groter bestanddeel van HW uitmaakt naarmate HW kleiner is. Op grond hiervan valt te verwachten dat de verdeling van de laagste HW's duidelijk anders is dan die van de hogere HW's. De invloed hiervan kan op bevredigende wijze worden teruggedrongen door ook voor HW een geschikte drempel in te voeren.

Anders dan voor HW-opzet is het echter niet mogelijk op voorhand een schatting te maken van de hoogte van die drempel voor HW. Wel kan men aannemelijk maken (zie § 6.2) dat deze drempel niet lager is dan de som van

- het hoogste astronomische hoogwater, dat in de reeks geselecteerde HW's van het betrokken station is vervat en
- de zojuist genoemde drempel van HW-opzet, groot 30 cm.

Met dit als uitgangspunt zijn nu bij de aanpassingen van bepaalde verdelingsfuncties bij de waarden van de betrokken tijdreeksen van HW verschillende drempelwaarden aangehouden; op deze wijze is het mogelijk gebleken per geval een aannemelijke drempelwaarde vast te stellen.

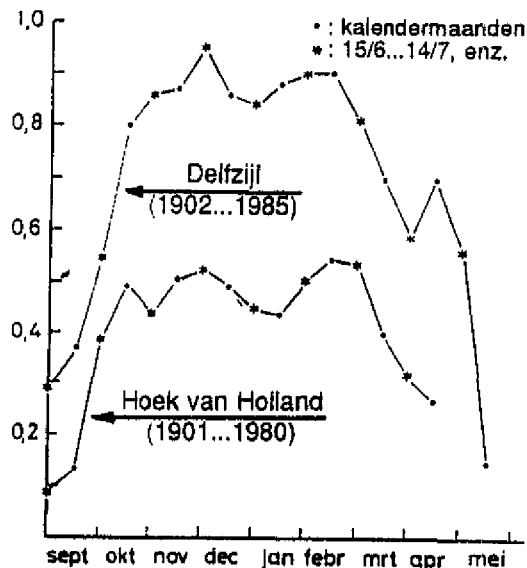
4.3.3 Tweede homogeniseringsselectie: vaststelling stormseizoenen

De eis dat het te analyseren waarnemingsmateriaal - althans bij goede benadering - homogeen is houdt in dat getracht moet worden alleen die waarnemingen „mee te nemen”, waarvan mag worden aangenomen dat ze afkomstig zijn uit eenzelfde gemeenschappelijke verdeling. Ten einde het in dit opzicht versturende gedeelte van de jaarcyclus te elimineren werd reeds in het „Deltarapport” onderscheid gemaakt tussen een „zomerperiode” en een „winterperiode”. Enigszins arbitrair werden als winterperiode gekozen de maanden november, december en januari, waarin ook de meeste jaarmaxima vallen; de waarnemingen uit de „zomerperiode” bleven bij de verrichte uitgebreide analyse buiten beschouwing (zie verder § 3.1.1). In de loop van de onderhavige studie bleek een nader onderzoek nodig ter vaststelling van het voor de statistische analyse meest geëigende jaargedeelte (het zgn. **stormseizoen**). Het onderzoek is gericht geweest op HW-opzet als zijnde de voor het optreden van stormvloed bepalende grootheid. Verder is nog onderscheid gemaakt tussen een lang en een kort stormseizoen.

Allereerst het *lange stormseizoen*, dat is het jaargedeelte dat voor de statistische analyse wordt beschouwd. De hiervoor geformuleerde eis luidt nu dat HW-opzetten, die optreden tijdens het lange stormseizoen, voor verschillende kalenderperioden binnen dit seizoen bij goede benadering dezelfde zgn. voorwaardelijke verdelingsfuncties bezitten (voor de theorie hierover zie men de beschouwing met betrekking tot (5.30) in § 5.3). Zij $P(Y_0 \leq y | Y_0 \geq u)$ de kans dat een HW-opzet Y_0 , die uitsteekt boven een drempel u , ten hoogste de waarde y aanneemt. De genoemde eis betekent nu dat de voorwaardelijke verdelingsfuncties P voor een geschikt gekozen drempel bij goede benadering gelijk moeten zijn voor verschillende kalenderperioden in het lange stormseizoen; het verdient daarbij de aandacht dat de gemiddelde aantallen van deze HW-opzetten in die verschillende perioden in het algemeen ongelijk zullen zijn.

Om te komen tot een keuze van het lange stormseizoen is gebruik gemaakt van zgn. **boxplots** [Chambers et al. 1983; Mc Gill et al. 1978]. Als eerste stap zijn voor de stations Hoek van Holland en Delfzijl uit de HW-opzet-bestanden de tijdreeksen gelicht met de betrekkelijk willekeurig gekozen drempel $u = 70$ cm. Deze reeksen omvatten achtereenvolgens 1157 HW-opzetten ($\approx 2\%$ van alle waarden) en 2294 HW-opzetten ($\approx 3\%$ van alle waarden).

Figuur 4.1:
Verhouding aantallen HW-opzetten
> 99 cm en (70,99) cm per
deelreeks



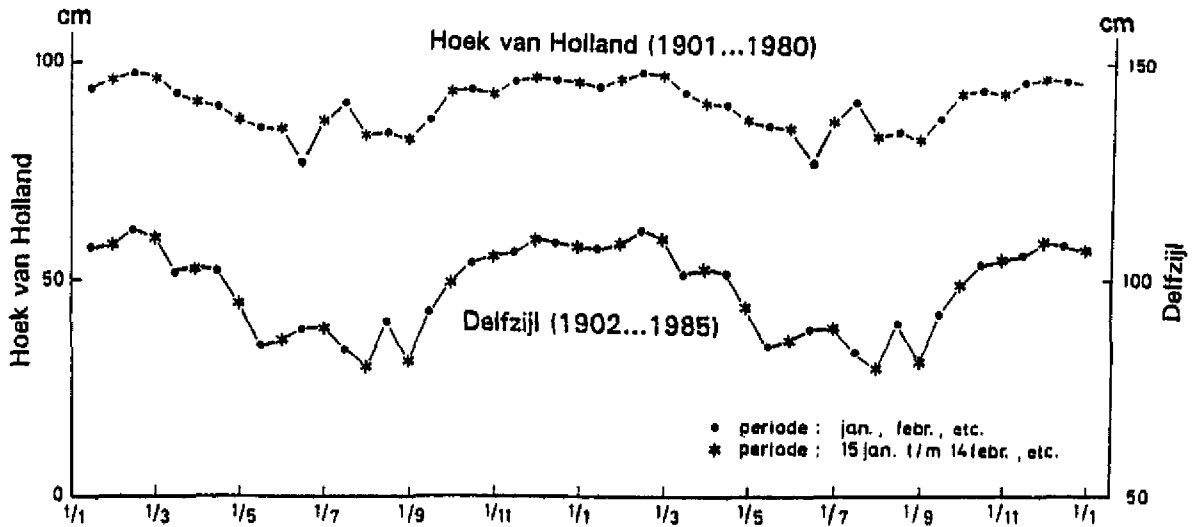
Per station zijn daarna de HW-opzetten naar kalenderperioden ingedeeld in deelreeksen en wel eerst volgens de kalendermaanden en vervolgens opnieuw volgens de perioden lopend van de 15e van elke kalendermaand tot en met de 14e van de volgende maand. Voor elk van de zo gevormde 2 x 12 deelreeksen is ten slotte door rechtstreekse afleiding uit het materiaal de voorwaardelijke verdelingsfunctie $P(Y_0 \leq y | Y_0 \geq u)$ geschat en grafisch weergegeven in een zogenaamde boxplot, waarna van de twee series kalenderperioden de boxplotten gezamenlijk in tijdsvolgorde naast elkaar zijn geplaatst (bovenste 2 afbeeldingen op de bijlagen 4 en 5). Daarbij is een box een zwart blokje met een horizontale uitsparing. De bovenkant van de box, de uitsparing en de onderkant van de box geven aan de ligging van het 25%-kwantiel h_1 (dat is de waarde overtroffen door 25% van het materiaal der deelreeks), de mediaan en het 75%-kwantiel h_2 . Uit de box steken twee verticale streepjeslijnen, onder de box over het gebied tot het „bakje”, dat is de kleinste waarde $\geq h_2 - 1,5(h_1 - h_2)$, boven de box over dat tot het „kapje”, d.i. de grootste waarde $\leq h_1 + 1,5(h_1 - h_2)$. Waarden, die niet door de box met streepjeslijnen worden gerepresenteerd staan elk afzonderlijk met een horizontaal lijntje aangegeven (dat op de bijlagen 4 en 5 geen horizontale lijntjes onder de boxen voorkomen is gevolg van de drempelkeuze $u = 70$ cm). Om een goed inzicht te krijgen in het verloop over het jaar zijn in de figuren steeds twee (identieke) jaarcyclus naast elkaar getekend. Ten behoeve van de keuze van het lange stormseizoen is nog aanvullende informatie in beschouwing genomen, ontleend aan de destijds beschikbare HW-opzetten-bestanden. Deze informatie betreft:

- de verhouding per deelreeks van de aantallen HW-opzetten ≥ 100 cm t.o.v. die in het interval $[70, 99]$ cm. Voor beide stations geeft figuur 4.1 hiervan het verloop over het jaar weer;
- het gemiddelde per deelreeks van HW-opzet. Figuur 4.2 toont voor beide stations het verloop over het jaar.

De keuze van het lange stormseizoen is gedaan op grond van de visuele indruk, die de gereproduceerde grafieken geven. Daarbij viel aan enige subjectiviteit niet te ontkomen; daardoor was het goed

Figuur 4.2:

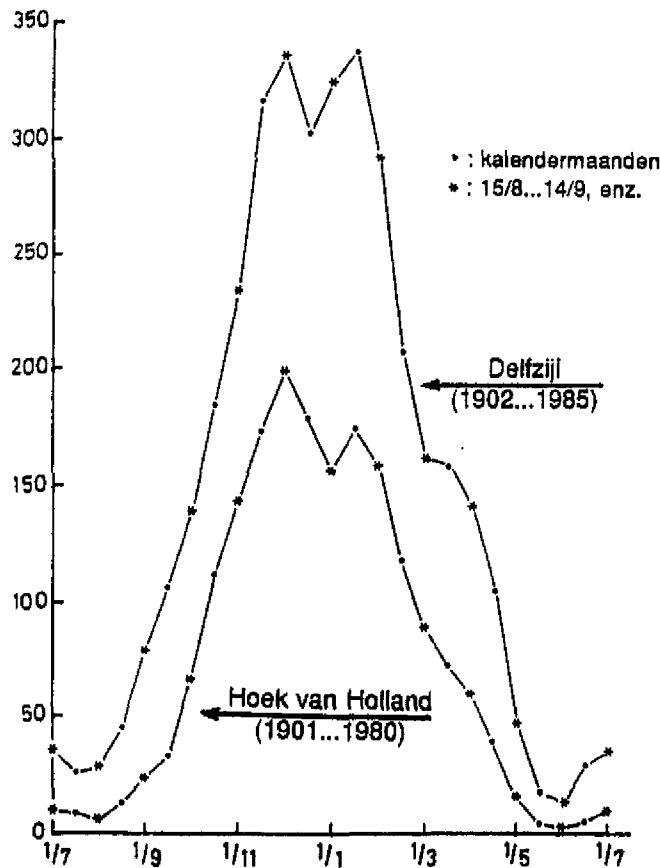
Maandgemiddelden van de HW-opzetten ≥ 70 cm



mogelijk voor beide stations hetzelfde lange stormseizoen te kiezen; dit samentreffen is weliswaar niet noodzakelijk, maar had de voorkeur aangezien het dan geoorloofd mag worden geacht ditzelfde seizoen ook voor de andere stations aan te houden. Gekozen is het seizoen 1 oktober...15 maart.

Figuur 4.3:

Aantal HW-opzetten ≥ 70 cm per deelreeks



Intussen is deze keuze nog niet definitief. Het voor statistische analyse te gebruiken materiaal dient namelijk niet alleen homogeen maar ook autocorrelatievrij te zijn en aan die tweede voorwaarde voldoen de beschouwde deelreeksen niet zonder meer. Nadat zij aan een selectie tegen autocorrelatie onderworpen zijn geweest dient de gemaakte keuze van het lange stormseizoen dan ook getoetst te worden (zie § 4.3.4). Voor de vaststelling van een selectieprocedure tegen autocorrelatie naar geavanceerde inzichten is het niet alleen van belang dat de HW-opzetten gelijk verdeeld zijn maar ook dat de reeks tijdstippen, waarop die HW-opzetten optreden welke de gekozen drempel te boven gaan, **stationair** is, dat wil zeggen dat het gemiddelde aantal HW-opzetten per tijdseenheid constant is. Bij die selectieprocedure speelt namelijk de tijdsduur tussen opeenvolgende HW-opzetten een belangrijke rol.

Voert men nu de procedure uit terwijl in een deel van de beschouwde kalenderperiode die tijdsduur gemiddeld markant afwijkt van die in het overige deel dan leidt dit tot vertroebelde uitkomsten. Figuur 4.3 geeft voor de twee betrokken stations weer de totale aantallen HW-opzetten ≥ 70 cm, die zich gedurende de desbetreffende 80 resp. 84 volle jaren hebben voorgedaan in elk der (per station) 24 onderscheiden deelreeksen. Op grond van figuur 4.3 vormen de HW-opzetten uit de maanden november...januari (het korte stormseizoen) het aangewezen materiaal voor het onderzoek ter vaststelling van een selectie tegen autocorrelatie volgens de aangeduide procedure.

4.3.4 Selectie tegen auto-correlatie van HW-opzet

De D-S-i-selectie

De onderlinge afhankelijkheid van HW-opzetten binnen een reeks (autocorrelatie) vormt een probleem voor de statistische analyse, immers toepassing van de gangbare theorie vereist onafhankelijke waarnemingen. Door selectie van HW-opzetten wordt nu geprobeerd de situatie van onafhankelijkheid te benaderen.

Het idee daarbij is dat als HW-opzetten in de tijd ver uit elkaar liggen ze ongeveer onafhankelijk zijn. Als nu in z'n algemeenheid wordt gesteld dat elke geselecteerde HW-opzet groter moet zijn dan de i direct voorafgaande en i onmiddellijk volgende HW-opzetten, dan gaat het er om een zodanige waarde voor i vast te stellen dat de autocorrelatie in voldoende mate is verwijderd en dus de HW-opzetten bij goede benadering onderling onafhankelijk zijn. Deze wijze van selecteren wordt verder aangeduid als de D-S-i-selectie (dunningsselectie). Bij de vaststelling van i zal een afweging tussen enerzijds het accepteren van enige afhankelijkheid en anderzijds het niet onnodig beperken van het aantal geselecteerde waarnemingen onvermijdelijk blijken te zijn. Zoals in par. 4.3.3 reeds is aangegeven hebben de onderstaande beschouwingen betrekking op het korte stormseizoen.

De autocorrelatiefunctie

De eerste gedachte voor het vaststellen van i gaat uit naar het berekenen van de autocorrelatiefunctie R_{xx} van iedere reeks HW-opzetten. Voor de reeksen van HW-opzet geldt [Van der Grinten en Lenoir 1973, par. 3.1.1, i.h.b. bl. 87]:

$$R_{xx}(t_1, t_2) = E\{X(t_1) \cdot X(t_2)\} \quad (4.1)$$

Hierin is

E = de verwachtingswaarde van

$X(t_1)$ = HW-opzet op tijdstip t_1

Deze functie hangt in het algemeen af van de tijdstippen t_1 en t_2 . Voor het geval $E(X(t)) = 0$ voor elke t is deze autocorrelatiefunctie gelijk aan de covariantie van de waarden op de tijdstippen t_1 en t_2 , welke een maatstaf is voor de samenhang.

Voor stationaire reeksen geldt dat de statistische eigenschappen onafhankelijk zijn van de tijd. In dat geval is de autocorrelatiefunctie alleen afhankelijk van het tijdsverschil tussen t_1 en t_2 . Dit verschil kan hier dus worden uitgedrukt in i . De autocorrelatiefunctie wordt dan:

$$R_{xx}(i) = E\{X(t) \cdot X(t+i)\} \quad \text{voor elke } t. \quad (4.2)$$

Deze functie, die dus een maatstaf vormt voor de samenhang tussen twee HW-opzetten die i tijdstappen in de tijdreeks van elkaar verwijderd zijn, zou per station berekend kunnen worden voor elk kort stormseizoen en vervolgens worden gemiddeld over de beschikbare stormseizoenen, dan wel kunnen worden berekend voor alle korte stormseizoenen tezamen mits daarbij alleen paren HW-opzetten worden beschouwd die in hetzelfde stormseizoen vallen.

Toch is deze weg niet gevolgd. Voor het onderzoek zijn namelijk slechts specifieke gebeurtenissen in de tijdreeksen van HW-opzet van belang. Het gaat daarbij met name om perioden rond potentieel gevaarlijke HW-opzetten. De circulatietypen die dergelijke HW-opzetten genereren vertonen een persistentie die verschilt van die van andere circulatietypen (zie § 2.2). Bij het berekenen van de autocorrelatiefunctie $R_{xx}(i)$ volgens formule (4.2) voor de volledige reeksen van HW-opzet in het korte stormseizoen wordt een gemiddeld beeld van de samenhang verkregen voor alle aanwezige weertypen, hetgeen dus niet de gewenste informatie oplevert.

Autocorrelaties van hoge HW-opzetten

Een methode om bovengenoemd bezwaar van vermenging van persistenties te ondervangen is het waarnemingsmateriaal te beperken tot die gebeurtenissen, die met het oog op het doel van het onderzoek van belang zijn, te weten hoge HW-opzetten. Dat kan bijvoorbeeld door alleen die HW-opzetten te beschouwen die boven een geschikt gekozen drempel u uitsteken. Noem deze HW-opzetten Y_t , dan wordt de autocorrelatiecoëfficiënt ρ_i bij een tijdsverschil van i hoogwaters gegeven door

$$\rho_i = \frac{\text{cov}(Y_t, Y_{t+i})}{\sigma^2(Y_t)} \quad (4.3)$$

waarbij „cov” staat voor covariantie en σ^2 voor variantie. Omdat bij goede benadering de Y_t 's in het korte stormseizoen een stationaire rij vormen is de autocorrelatiecoëfficiënt ρ_i in (4.3) onafhankelijk van tijdstip t .

De ρ_i 's zijn geschat m.b.v. de formule

$$\hat{\rho}_i = \frac{N_2^{-1} \sum_{t=1}^{N_2} (Y_t - \bar{Y})(Y_{t+i} - \bar{Y})}{N_0^{-1} \sum_{t=1}^{N_0} (Y_t - \bar{Y})^2} \quad (4.4)$$

met

$$\bar{Y} = N_0^{-1} \sum_{t=1}^{N_0} Y_t$$

waarin

N_0 = aantal HW-opzetten boven drempelwaarde u ;

N_2 = aantal paren (Y_t, Y_{t+i}) van HW-opzetten binnen een willekeurig stormseizoen met een tijdsverschil van i hoogwaters, waarvoor geldt dat zowel Y_t als Y_{t+i} boven de drempelwaarde u uitsteken.

De resultaten van de schattingen volgens formule (4.4) zijn voor de vijf stations gegeven op bijlage 6 voor de drempelwaarden $u = 30, 40, \dots, 70$ cm en voor $i = 1, 2, \dots, 10$.

Voor $i = 1$ is nog een duidelijke correlatie aanwezig; voor hogere waarden van i wordt de autocorrelatie snel klein. Op grond van deze resultaten lijkt een D-S- i selectie met $i = 3$ à 5 geschikt om de situatie van onafhankelijkheid voldoende te benaderen.

Het gedrag van HW-opzet rond piekwaarden

In de voorgaande beschouwing waren steeds alle HW-opzetten boven de gekozen drempel bij de verwerkingen betrokken. Uitgaande van de voorkennis dat alleen de hoogste HW-opzet van een stormperiode, die als één gebeurtenis mag worden opgevat, zal worden geselecteerd, kan het correlatie-onderzoek zich speciaal richten op de samenhang van HW-opzetten in de buurt van dergelijke piekwaarden.

Een *piekwaarde* is hoger dan de i voorafgaande en i volgende HW-opzetten. Piekwaarden kunnen dus worden verkregen door een D-S- i selectie; voor het onderzoek naar de bedoelde samenhang is het gewenst dat de waarde van deze i niet te klein en niet te groot wordt gekozen. Gezien de zojuist gevonden uitkomst is ter selectie van piekwaarden voor D-S-4 gekozen.

Om na te gaan of er sprake is van samenhang is zowel de gemiddelde HW-opzet op afstand i van een piekwaarde als de correlatiecoëfficiënt tussen de piekwaarden en de waarden op afstand i beschouwd. Voor de piekwaarden werd een ondergrens u aangehouden, voor de waarden op afstand i gold deze beperking niet.

De bijlagen 7 en 8 geven voor achtereenvolgens $u = 30$ cm en $u = 70$ cm het verloop van het gemiddelde van de HW-opzetten op i tijdstappen na een piekwaarde (volgens D-S-4 geselecteerd) en dat van de correlatiecoëfficiënt als functie van i voor de vijf beschouwde stations.

De correlatiecoëfficiënten zijn berekend met de formule

$$\hat{\rho}_i = \frac{\sum_{t=1}^N (X_t - \bar{X}_t)(X_{t+i} - \bar{X}_{t+i})}{\sqrt{\sum_{t=1}^N (X_t - \bar{X}_t)^2 \sum_{t=1}^N (X_{t+i} - \bar{X}_{t+i})^2}} \quad (4.5)$$

waarin: N = aantal piekwaarden
 X_t = piekwaarde van HW-opzet op tijdstip t
 X_{t+i} = HW-opzet op afstand i van X_t

Voor $i = 0$ wordt de gemiddelde piekwaarde verkregen; de correlatiecoëfficiënt is dan uiteraard 1.

De gemiddelde HW-opzet op afstand $i > 1$ en $i < -1$ van de piekwaarden blijkt vrijwel meteen terug te vallen tot het niveau van het gemiddelde van alle HW-opzetten. De correlatiecoëfficiënt toont wat meer samenhang in de omgeving van de piekwaarden. Een duidelijke, ondubbelzinnige keuze is op grond hiervan moeilijk te maken. Voor $i = 1$ lijkt de meeste afhankelijkheid verdwenen te zijn.

Afhankelijkheid en clustering van hoge HW-opzetten

Bij onafhankelijkheid van HW-opzetten kan het proces, dat de stochastische tijdstippen genereert waarop HW-opzetten optreden, benaderd worden door een zogenaamd *Poisson-proces*. Dit houdt in het bijzonder in dat geen clustering optreedt van hoge HW-opzetten. Er bestaat dan dus geen bijzondere neiging tot het vóórkomen van HW-opzetten in groepen, dat wil zeggen dat zij in de tijd dicht bij elkaar liggen; voor zover toch groepen optreden berust dit op puur toeval. Afhankelijkheid van HW-opzetten voor een gegeven station zal daarentegen in de regel wel leiden tot clustering van HW-opzetten. Stelt men de rij HW-opzetten in een willekeurig kort stormseizoen voor door $Y_1, Y_2, \dots, Y_i, \dots, Y_n$ (N.B.: i heeft hier dus de betekenis van rangnummer), dan wordt een intuïtief plausibele maat voor het optreden van clustering van hoge HW-opzetten gegeven door [Diggle 1983]

$$V_r(u) = \frac{\#\{(i, j) : i \neq j, |i-j| \leq r, Z_i = Z_j = 1\}}{\#\{i : Z_i = 1\}} \quad (4.6)$$

Hierin is: $Z_i = 1$ voor $Y_i \geq u$
 $= 0$ voor $Y_i < u$
 u = geschikt gekozen drempel
 r = afstand tussen twee HW-opzetten in tijdseenheden
 $\#$ = aantal

In woorden: de grootte $V_r(u)$ is het quotiënt van het aantal paren hoge HW-opzetten ($\geq u$), welke hoogstens op een afstand van r tijdseenheden (getijcycli) van elkaar optreden, en het totale aantal hoge HW-opzetten.

Beschouw verder de „clustergrootte” $V_r^*(u) = nV_r(u)$, dan kan worden afgeleid dat haar verwachtingswaarde in het geval van *onafhankelijkheid* evenredig is met r . Deze afleiding wordt in Aanhangsel 1 behandeld. De werkwijze is nu dat $V_r^*(u)$ per station wordt berekend voor elk beschikbaar kort stormseizoen, waarna wordt gemiddeld over de stormseizoenen.

De bijlagen 9 A t/m 9 E geven voor de vijf stations het verloop van $V_r^*(u)$ als functie van r bij de selecties D-S-0 t/m D-S-4 en voor achtereenvolgens $u = 30$ cm en $u = 70$ cm. Naarmate een daarop uitgezette kromme meer op een rechte lijkt, benadert de corresponderende selectie de situatie van onafhankelijkheid beter. Uit de figuren komen de volgende twee punten duidelijk naar voren:

- de clustering is duidelijker bij hogere drempels;
- de krommen benaderen meer een rechte lijn naarmate de i van de selectie D-S- i groter is.

Op grond van deze resultaten lijkt een selectie van HW-opzetten volgens D-S-4 verantwoord en is een keuze voor D-S-1 of D-S-2 niet goed mogelijk. Gekozen is voor een selectie van HW-opzetten volgens D-S-4 voor alle stations. Deze keuze bevat een zekere subjectiviteit (zie § 7.5).

Meteorologische argumenten voor D-S-4

De *gemiddelde* persistentie van 2,41 dagen (ca. 4,6 getijcycli) van de weertypen met een verhoogde kans op een hoge HW-opzet (zie § 2.2.2) geeft een indicatie van de tijdschaal van de autocorrelatie door meteorologische oorzaak.

Deze tijdschaal omvat de lengte van alle in § 2.1.2 genoemde mechanismen. Door de keuze voor D-S-4 (2 x 4 getijcycli) is de verwijdering van deze autocorrelatie van HW-opzet in goede mate verzekerd.

De correlatie tussen circulatietypen zal hiermee niet geheel zijn verwijderd. Dit zou een dermate „brede” selectie vergen dat het waarnemingsmateriaal te zeer zou worden uitgedund, wat bezwaarlijker wordt geacht dan de *geringe* overblijvende autocorrelatie in HW-opzet.

Conclusies en slotopmerkingen

In het voorgaande is het vraagstuk welke waarde van i moet worden gekozen voor de selectie D-S- i tegen autocorrelatie op vier aspecten onderzocht. De uitkomsten voeren tot de slotsom dat voor $i = 4$ gekozen dient te worden.

Rest vervolgens de vraag of de gemaakte keuze van D-S-4 herziening van de vaststelling van het lange stormseizoen nodig maakt. Om deze vraag te beantwoorden zijn voor de beide stations Hoek van Holland en Delfzijl de in § 4.3.3 beschouwde HW-opzetten onderworpen aan de selectie D-S-4, waarvan de overblijvende waarden eerst zijn ondergebracht in de deelreeksen 15 januari...14 februari enz. en vervolgens uitgezet in boxplots (bijlagen 4 en 5, onderste afbeeldingen). Ten opzichte van de overeenkomstige grafieken voor HW-opzet zonder de selectie D-S-4 zijn er wel verschillen, maar de keuze van het lange stormseizoen blijkt door de toepassing van de selectie D-S-4 niet te worden beïnvloed. De overeenkomstige grafieken voor de volle kalendermaanden tonen een geheel vergelijkbaar beeld. In een vrij vroeg stadium van het onderzoek is nog nagegaan in hoeverre dunningsselecties voor Hoek van Holland leiden tot andere tijdreeksen van HW dan de selectie „Van der Ham” (§ 3.1.1). Uit de resultaten van dit onderzoek [Samenwerkingsverband Basispeilen 1986, par. 5.2.4] kan worden afgeleid dat zonder invoering van drempels bij toepassing van alleen D-S-4 in een 70% van de gevallen identieke HW's zouden voorkomen, maar dat bij die dunningsselectie het totale aantal waarden 20% groter zou zijn. Invoering van drempels verhoogt echter het eerste percentage en verlaagt het tweede.

4.3.5 De homogeniteit van het stormklimaat in de beschouwde waarnemingsperiode

Bij het vaststellen van het stormseizoen en de breedte van de D-S- i -selectie tegen autocorrelatie zijn alle HW-opzetten beschouwd als trekkingen uit één populatie met een bepaalde kansverdeling.

Daarbij is dus stilzwijgend aangenomen dat zich geen veranderingen hebben voorgedaan in het klimaat van de HW-opzetten in de beschouwde waarnemingsperioden. Deze veronderstelling zal moeten worden geverifieerd.

De bijlagen 10 en 11 geven voor achtereenvolgens Hoek van Holland en Delfzijl het verloop van het aantal (ν_u) HW-opzetten per lang stormseizoen onder toepassing van de selectie D-S-4 en van drempelwaarden u . Voor de drempels zijn de waarden 30,50,...,130 cm genomen. Als zogenaamde nulhypothese H_0 wordt nu gesteld dat ν_u niet onderhevig is aan een trend of, anders gezegd, dat ν_u gerekend over een voldoende lang tijdvak onafhankelijk is van de tijd zonder dat er binnen dat tijdvak sprake is van een duidelijk niet-lineair (b.v. sinusoidaal) verband tussen ν_u en de tijd. Uit de bijlagen 10 en 11 blijkt dat zo'n niet-lineair verband zich niet voordoet. Zou bij toetsing voor een station blijken dat H_0 voor een scala van drempelwaarden u niet verworpen behoeft te worden dan mag worden geconcludeerd dat daar het klimaat van HW-opzet in het lange stormseizoen homogeen is geweest gedurende de beschouwde periode.

Een toets die zich goed leent voor het toetsen op onafhankelijkheid is de Spearman-toets. Een korte beschrijving van deze toets zal met de formule voor de toetsingsgrootheid \underline{t} worden gegeven in § 4.4. Bij de toepassing van deze toets op de reeksen aantallen zijn de gevallen met $u = 30,50,...,110$ cm in beschouwing genomen en is rekening gehouden met groepen van gelijke aantallen.

In tabel 4.1 zijn de uitkomsten van de toetsing weergegeven voor de stations Hoek van Holland en Delfzijl.

De kritieke gebieden voor \underline{t} bij tweezijdige toetsing en onbetrouwbaarheidsdrempel $\alpha = 5\%$ zijn: $\underline{t} \geq 1,96$ en $\underline{t} \leq -1,96$.

Geconcludeerd wordt dat in geen van de gevallen de nulhypothese wordt verworpen. Er mag derhalve aangenomen worden dat het klimaat van de HW-opzetten niet veranderd is in de beschouwde waarnemingsperioden. Deze conclusie wordt bevestigd door de figuren op de bijlagen 10 en 11; daarin is het aantal HW-opzetten per lang stormseizoen uitgezet tegen het jaartal, waarin de maand januari van het betrokken stormseizoen valt.

Tabel 4.1:
Toetsingsgrootheid \underline{t} van de Spearman-toets voor de toetsing van de homogeniteit van het klimaat van HW-opzet

drempel u (cm)	Hoek van Holland 1887/88...1984/85		Delfzijl 1881/82...1984/85	
	$\Sigma \nu_u$	\underline{t}	$\Sigma \nu_u$	\underline{t}
30	1576	+0,25	1828	-0,66
60	977	-0,93	1281	+0,41
70	562	-0,72	857	-0,41
90	315	-0,49	601	+0,47
110	162	+0,91	400	+0,58

4.4 Toetsing van de onafhankelijkheid van de astronomische hoogwaterstanden en HW-opzetten

Bij het gebruik van HW-opzet als te verwerken statistisch veranderlijke zullen per station de uitkomsten van de statistische analyse gekoppeld moeten worden aan de verdelingsfunctie van het astronomische

hoogwater ten einde de gezochte HW's van gekozen overschrijdingskansen te berekenen. Voorwaarde voor die koppeling is dat HW-opzet onafhankelijk is van de astronomische hoogwaterstanden; of daaraan wordt voldaan wordt in deze paragraaf onderzocht.

4.4.1 De Spearman-toets

Met de klassieke toets om te onderzoeken of twee reeksen waarnemingen, gegeven in de vorm van N getallenparen, onafhankelijk van elkaar zijn, gaat men na of de „steekproef-correlatiecoëfficiënt“ al dan niet significant is. Daarbij wordt aangenomen dat elk van beide reeksen een steekproef is uit de eigen normale verdeling. Deze beperkende vooronderstelling mag in het onderhavige geval zeker niet gemaakt worden. Een toets, die zich wel voor het bedoelde onderzoek leent, is de Spearman-toets [Hájek 1969]; dit is een zogenaamde verdelingsvrije of niet-parametrische toets, dat wil zeggen dat hij van toepassing is ongeacht uit welke twee verdelingen de beide reeksen als steekproef afkomstig zijn.

Gegeven zijn de getallenparen $(X_1, Y_1), \dots, (X_n, Y_n)$, waarvan de waarden X_i onderling onafhankelijk zijn en „gelijkverdeeld zijn“ (dat wil zeggen behoren tot dezelfde statistische verdeling F_X); voor de waarden Y_i geldt het overeenkomstige. De nulhypothese H_0 luidt nu dat X en Y onafhankelijk van elkaar zijn.

De procedure is als volgt: Ken aan elk der waarden X_i toe het rangnummer naar grootte R_g en insgelijks aan elke waarde Y_i het rangnummer Q_h . Vervang vervolgens de getallenparen (X_i, Y_i) door de corresponderende rangnummerparen (R_g, Q_h) . De grootte voor het toetsen van de nulhypothese H_0 volgens Spearman is

$$S = \sum_{i=1}^N R_g \cdot Q_h \quad (4.7)$$

„Onder H_0 “ (dat wil zeggen dat H_0 inderdaad van toepassing is) is de verdeling van S symmetrisch en voor grote N bij benadering normaal. Men kan verder bewijzen dat onder H_0 voor de verwachtingswaarde van S en voor haar variantie achtereenvolgens gelden:

$$E(S) = \frac{1}{4} N(N+1)^2 \quad (4.8)$$

$$\text{Var}(S) = \frac{1}{144} N^2 (N+1)^2 (N-1) \quad (4.9)$$

De toetsingsgrootte t is nu

$$t = \frac{S - E(S)}{\sqrt{\text{Var}(S)}} \quad (4.10)$$

Onder H_0 en voor grote N behoort t bij de gestandaardiseerde normale verdeling (d.i. die met gemiddelde waarde = 0 en standaardafwijking = 1). Zoals gebruikelijk bij een nulhypothese wordt H_0 niet verworpen zolang de berekende toetsingsgrootte behoort tot het centrale domein der verdeling met in totaal 95% waarschijnlijkheid, dat wil zeggen zolang $|t| < 1,96$. Anders geformuleerd: de kritieke gebieden voor t bij tweezijdige toetsing en onbetrouwbaarheidsdrempel $\alpha = 5\%$ zijn $t \geq 1,96$ en $t \leq -1,96$.

4.4.2 Gelijke waarnemingen

Een complicatie in de bovenbeschreven procedure wordt gevormd door het voorkomen van groepen van gelijke waarnemingen, de zgn. „ties“. Het is dan niet zonder meer duidelijk hoe de rangnummers moeten worden toegekend binnen die ties.

Een willekeurige toedeling heeft als nadeel dat de toets niet exact gelijk verloopt bij herhaling of uitvoering door een ander. Daarom wordt meestal, ook hier, aan elke waarneming in een tie het gemiddelde rangnummer toegekend van alle waarnemingen in die tie.

Als $\tau_{x,j}$ en $\tau_{y,j}$ de omvang van de ties voorstellen in (X_1, \dots, X_n) resp. (Y_1, \dots, Y_n) en μ_x en μ_y het aantal ties in de respectievelijke waarnemingsreeksen, dan geldt onder H_0 :

$$E(S | \tau_x, \tau_y) = \frac{1}{N} \left(\sum_{j=1}^N j \right)^2 = \frac{1}{4} N(N+1)^2 \quad (4.8a)$$

$$\text{Var}(S | \tau_x, \tau_y) = \frac{1}{144(N-1)} (A - T_x)(A - T_y) \quad (4.9a)$$

waarin

$$A = N(N+1)(N-1)$$

$$T_x = \sum_{j=1}^{\mu_x} \tau_{x,j} (\tau_{x,j} + 1) (\tau_{x,j} - 1)$$

$$T_y = \sum_{j=1}^{\mu_y} \tau_{y,j} (\tau_{y,j} + 1) (\tau_{y,j} - 1)$$

De verwachtingswaarde is dus gelijk aan het geval zonder ties, de variantie wordt gecorrigeerd met een waarde afhankelijk van het aantal en de omvang van de ties.

De maat voor de onafhankelijkheid blijft $S = \sum R_g \cdot Q_h$, d.i. (4.7), waarin bij ties voor R_g en Q_h de gemiddelde rangnummers moeten worden ingevuld.

De verdeling van S is weer symmetrisch en voor grote N bij benadering normaal.

De formule voor de toetsingsgrootte luidt nu:

$$\underline{t} = \frac{S - E(S | \tau_x, \tau_y)}{\sqrt{\text{Var}(S | \tau_x, \tau_y)}} \quad (4.10a)$$

Onder H_0 heeft ook \underline{t} voor grote N de gestandaardiseerde normale verdeling. De kritieke gebieden zijn weer $\underline{t} \geq 1,96$ en $\underline{t} \leq -1,96$.

4.4.3 Toepassing op de astronomische hoogwaterstanden en HW-opzetten

De toetsing van de onafhankelijkheid van astronomische hoogwaterstanden en HW-opzetten is voor de 5 stations gedaan op de wijze als beschreven in § 4.4.2.

De getallenparen bestaan voor elk station uit de geselecteerde HW-opzetten en de bijbehorende astronomische hoogwaterstanden. De toegepaste selecties betreffen de keuzen van een drempel u (zie § 4.3.2), van het lange stormseizoen (1 okt...15 mrt; zie § 4.3.3) en van

de selectie D-S-4 (zie § 4.3.4), waarmee voor de HW-opzetten is voldaan aan de eisen van onafhankelijkheid en gelijke verdeling. Aangenomen mag worden dat stormen geen voorkeur hebben voor een bepaalde getijconstellatie (§ 2.2.3). De bij de geselecteerde HW-opzetten behorende astronomische hoogwaterstanden kunnen dus als willekeurige trekkingen uit de populatie van astronomische hoogwaterstanden worden beschouwd. Voor de als stochastische grootheid opgevatte astronomische hoogwaterstanden mag er eveneens van worden uitgegaan dat ze onafhankelijk en gelijk verdeeld zijn. De toetsing is verricht voor een vijftal drempels voor HW-opzet met behulp van (4.10^a). De uitkomsten van de berekeningen zijn verzameld in de hiernavolgende tabel:

Tabel 4.2:
Spearman's grootheid \underline{t} voor HW-opzet en astronomische hoogwaterstanden

drempel HW-opzet = u (cm)	Vlissingen 1882...1985		Hoek v.Holl. 1888...1985		Den Helder 1933...1985		Harlingen 1933...1985		Delfzijl 1878...1985	
	N	\underline{t}	N	\underline{t}	N	\underline{t}	N	\underline{t}	N	\underline{t}
30	1660	+0,14	1576	+0,22	866	+0,06	936	+1,20	1828	+1,56
50	921	-0,98	977	+0,32	569	+0,75	665	+0,53	1281	+1,08
70	447	-1,21	562	+1,14	353	-0,62	468	+1,30	857	+0,68
90	223	-1,50	315	-0,52	204	-0,32	324	+0,09	601	+0,09
110	99	-0,54	162	+0,19	132	-0,29	205	-1,21	400	-0,42

NB: elke tijdsbepaling doelt op de in die genoemde periode voorgekomen lange stormseizoenen, waarbij een stormseizoen is aangeduid met het jaar waarin zijn maand januari valt.

Slotsom: Aangezien in geen van de beschouwde gevallen \underline{t} in een der kritieke gebieden valt en dus de nulhypothese niet wordt verworpen, mag voor de verdere statistische bewerkingen worden uitgegaan van de onafhankelijkheid van astronomische hoogwaterstanden en HW-opzetten. Bij gevolg behoeven berekende HW-opzetten (§ 4.2.2) geen correcties in verband met de opgetreden veranderingen van de gemiddelde hoogwaterstanden. Ter illustratie zijn in de figuren 1, 2 en 3 van bijlage 12 de getallenparen voor drie der beschouwde stations uitgezet bij een drempelniveau van 30 cm.

4.5 Herleiding van HW's naar de toestand 1985

In § 2.3.1 kwam reeds ter sprake dat de veranderingen, die de gemiddelde hoogwaterstanden aan onze kusten in de loop der jaren hebben ondergaan, niet zonder meer gelijk zijn aan die, welke de gemiddelde zeespiegel daar vertoont.

Voor zover zij van die relatieve zeespiegelrijzing afwijken is dit, naar men moet aannemen, gevolg van menselijk ingrijpen (grootschalig baggeren, grote werken) en/of van morfologische veranderingen. Ten einde per station voor de statistische analyses van HW's over zo homogeen mogelijk waarnemingsmateriaal te beschikken dienen de HW's dus gecorrigeerd te worden op de ter plaatse opgetreden wijzigingen van de gemiddelde hoogwaterstanden, die zich in het algemeen geleidelijk, maar soms abrupt hebben voltrokken. De grootheid, die zich goed leent voor de beoogde correcties is het jaargemiddelde van de plaatselijke hoogwaterstand; de toestand, waarnaar gecorrigeerd wordt, is die van het jaar 1985.

Bijlage 13 geeft voor de vijf betrokken stations weer het verloop van het jaargemiddelde van de plaatselijke hoogwaterstand. Dit veeljarig verloop is over zo lang mogelijke perioden benaderd door een regressielijn, samengesteld uit een rechte (de trendlijn) plus een sinusoïde met een periode van 18,6 jaar, welke sinusoïde de invloed weergeeft van het nodale tij. De formule van de regressielijn luidt:

$$Y = A \cdot J + B \sin \frac{J+K}{18,6} \cdot 2\pi + C \quad (4.11)$$

- met
- Y = jaargemiddelde van HW aan het betrokken peilmeetstation in cm (t.o.v. NAP);
 - J = jaartal;
 - K = faseverschuiving = -2 jaar;
 - A = lineaire trend in cm/jaar;
 - B = amplitude der sinusoïde in cm;
 - C = constante in cm.

A, B en C zijn bepaald volgens de methode der kleinste kwadraten. De toe te passen correcties volgen in beginsel uit de trendlijnen. De afsluiting der Zuiderzee (voltooid 28 mei 1932) heeft bij *Den Helder* en *Harlingen* zowel het verticale getij als de opwaaiingsverschijnselen ingrijpend gewijzigd [Doekes 1985]. Hierdoor is het voor de HW's noch voor de HW-opzetten mogelijk de waarden aan deze stations van voor 1933 te homogeniseren naar de omstandigheden van na de afsluiting. Voor deze stations zijn voor het onderhavige onderzoek dus alleen bruikbaar de HW-waarnemingen vanaf 1933; in beide gevallen kan met één regressielijn volgens (4.11) worden volstaan. Dit laatste is voor geen der drie andere stations mogelijk. Te *Vlissingen* en *Hoek van Holland* is sprake van een min of meer abrupte sprong rond achtereenvolgens 1887 (14 cm; oorzaak niet duidelijk) en 1965 (7,6 cm; oorzaak waarschijnlijk de aanleg van de Maasvlakte [TO 1971-1980, bl. XLV]). Voor elk van deze stations zijn nu twee trendlijnen bepaald, t.w. voor de periode voor en die na de betrokken trendbreuk; voor de twee genoemde jaren is als correctie aangehouden het gemiddelde van de correcties, die zouden volgen uit de twee desbetreffende trendlijnen. Te *Delfzijl* hebben zich twee trendbreuken voorgedaan; de eerste (omstreeks 1960) was gevolg van uitbaggeren van de vaargeul naar Emden, de tweede (in het najaar van 1978) van werken in en bij de haven van Delfzijl. In dit tweede geval was er sprake van een min of meer abrupte rijzing van de hoogwaterstanden van een 6 cm [De Ronde 1984^a]. De geselecteerde HW's zijn aan de hand van deze gegevens gecorrigeerd naar de toestand 1985; daarbij is per stormseizoen aangehouden één correctie in hele cm en wel die van het jaar J, waarin de maand januari van dat seizoen valt. In tabel 4.3 zijn de toegepaste correctieformules en de twee gemiddelde correctiewaarden verzameld.

4.6 De kansdichtheidsfuncties der astronomische hoogwaterstanden in het stormseizoen

Bij de keuze van HW-opzet als statistisch veranderlijke dienen per station de uitkomsten van de statistische analyse gekoppeld te worden aan de relevante verdelingsfunctie van de astronomische hoogwaterstanden.

Tabel 4.3:

Correcties op de geselecteerde
HW's (in cm)

TIJDVAK OF JAAR	FORMULE	STATION
1882 ... 1886 1887	0,34 (1985-J) - 14,0 26	Vlissingen
1888 ... 1985	0,34 (1985-J)	
1888 ... 1964 1965	0,12 (1985-J) + 15,8 14	Hoek van Holl.
1966 ... 1985	0,53 (1985-J)	
1933 ... 1985	0,18 (1985-J)	Den Helder
1933 ... 1985	0,30 (1985-J)	Harlingen
1882 ... 1959 1960 ... 1978	0,18 (1985-J) + 11,4 0,40 (1985-J) + 6,0	Delfzijl
1978 ... 1985	0,40 (1985-J)	

De relevantie van zo'n functie houdt in dat ze betrekking moet hebben op:

- dezelfde hydrologische enz. toestand als waarop de gezochte HW's van gekozen overschrijdingskansen betrekking hebben, dat wil zeggen op de toestand 1985 (§ 4.5);
- alleen het lange stormseizoen, zijnde het jaargedeelte, waarin de verwerkte HW-opzetten zich voordeden (§ 4.3.3).

Om aan eis a te voldoen is uitgegaan van de verrichte culminatie-analyses 1978...1986 en - voor Delfzijl - 1979...1986 (§ 4.2.2); deze analyseperioden worden geacht met voldoende nauwkeurigheid de toestand 1985 te representeren.

Met behulp van deze analyses zijn pseudo astronomische hoogwaterstanden berekend over de periode 1901...1994, die 93 lange stormseizoenen (1 oktober...15 maart) bevat, dat wil zeggen alle lange seizoenen uit 5 volledige perioden van het nodale tij (93 = 5 x 18,6).

Voor deze 93 lange stormseizoenen is nu per station de kansdichtheidsgrafiek samengesteld met klassebreedten van 1 cm; deze grafieken zijn zodanig genormeerd dat per grafiek de gesommeerde kansdichtheid juist gelijk is aan 1.

De kansdichtheidsgrafieken van de astronomische hoogwaterstanden aan de vijf betrokken basisstations figureren op bijlage 14; de hoogste en laagste van deze standen treft men aan in tabel 4.4.

Tabel 4.4:

Hoogste en laagste astronomische
hoogwaterstanden in 93 lange
stormseizoenen (toestand 1985; in
cm t.o.v. NAP)

Station	Hoogste	Laagste
Vlissingen	+ 279	+ 113
Hoek van Holland	+ 152	+ 57
Den Helder	+ 97	+ 22
Harlingen	+ 139	+ 40
Delfzijl	+ 183	+ 83

Bij een en ander passen de volgende kanttekeningen:

- De periode van 93 jaar is mede ingegeven door het feit dat ze van dezelfde orde van grootte is als de langste waarnemingsperioden. Echter, een reeks van 19 stormseizoenen zal kansdichtheidfuncties opleveren, die niet veel afwijken van die van bijlage 14.

- De berekende grafieken verschillen belangrijk van die van alle astronomische hoogwaterstanden over de volle periode van 93 jaar. Zo zijn b.v. de laagste astronomische hoogwaterstanden in de gepresenteerde grafieken aanmerkelijk hoger dan de jaarminima, die buiten het stormseizoen (t.w. in april-mei) plegen te vallen (vergelijk tabel 2.1).
- Opvallend is de duidelijke tweetoppigheid van 4 der 5 diagrammen. De belang rijkste oorzaak hiervan is voor Vlissingen en Hoek van Holland de spring tij/doodtij-cyclus en voor Den Helder en Harlingen de dagelijkse ongelijkheid.

4.7 Histogrammen van het bewerkte waarnemingsmateriaal

De HW-opzetten en de HW's, die verkregen zijn na alle in dit hoofdstuk beschreven bewerkingen op de bestanden van deze grootheden (zie §§ 4.1 en 4.2), vormen het uitgangsmateriaal voor de verdere statistische analyses. Die bewerkingen behelsden met name

- a voor de HW-opzetten:
 1. de invoering van een drempel van 30 cm (§ 4.3.2);
 2. de beperking tot het lange stormseizoen 1 oktober...15 maart (§ 4.3.3);
 3. selectie volgens D-S-4 (§ 4.3.4);
- b voor de HW's:
 4. beperking tot de HW's, die behoren bij de geselecteerde HW-opzetten, verkregen na de selecties genoemd onder a;
 5. herleiding naar de toestand 1985 (§ 4.5).

Ten einde een indruk te geven van de aard van de zo verkregen reeksen van HW-opzet resp. HW zijn van alle reeksen histogrammen gemaakt. Bijlage 15 geeft hiervan de histogrammen van HW-opzet voor de vijf stations, bijlage 16 die van HW (n.b.: zonder invoering van een drempel voor HWI). Vanuit een oogpunt van uniformiteit is voor alle histogrammen als klassebreedte 10 cm aangehouden. De verhoudingsgewijze lage aantallen voor Den Helder en Harlingen zijn uiteraard gevolg van het feit dat voor deze stations de beschikbare waarnemingsperiode aanzienlijk korter is dan voor de andere drie stations.

5. Toegepaste Methoden

5.1 Inleiding

In dit hoofdstuk beschrijven we de verschillende statistische methoden welke gebruikt zijn bij het beantwoorden van de centrale vraag van het onderzoek *een zodanig hoogwaterniveau aan te geven dat de kans op overschrijding van dat niveau in een willekeurig gegeven jaar¹ gelijk is aan p_0* . Ook de *nauwkeurigheid* waarmee dit hoogwaterniveau kan worden bepaald - uitgedrukt in een *betrouwbaarheidsinterval* - zal hierbij aan de orde komen. In dit hoofdstuk wordt verondersteld dat p_0 een van te voren vastgesteld getal is. Vooral p_0 -waarden tussen 10^{-3} en 10^{-5} zijn van belang. Voor ons doel is het nodig de kansverdeling van het *jaarmaximum* van de hoogwaterstanden (HW) zo nauwkeurig mogelijk te schatten. We zullen nu veronderstellen dat de in § 4.3 beschreven selecties van waarnemingen reeds heeft plaatsgevonden en dat we de beschikking hebben over onafhankelijke waarnemingen X_1, \dots, X_N - dit zijn de geselecteerde in het verleden waargenomen hoogwaterstanden bij een gegeven station - alle afkomstig uit een in principe onbekende verdeling met verdelingsfunctie F . M.a.w. F is de gemeenschappelijke verdelingsfunctie van de geselecteerde hoogwaterstanden bij een gegeven station.

We moeten nu een hoogwaterniveau x_{p_0} bepalen zodanig dat

$$P(\{\text{jaarmaximum}^1 \text{ van geselecteerde hoogwaterstanden in een willekeurig (toekomstig) jaar aan een gegeven station}\} \leq x_{p_0}) = 1 - p_0 \quad (5.1)$$

M.a.w. we willen voor een gegeven station het niveau x_{p_0} bepalen dat de hoogste hoogwaterstand in een (toekomstig) willekeurig jaar, onder ongewijzigde klimatologische, hydrologische en hydraulische omstandigheden, met kans p_0 overschrijdt.

We merken allereerst op dat de geselecteerde HW's X_1, \dots, X_N ruwweg de hoogste waarnemingen zijn gemeten tijdens stormen. Men kan staande houden dat de aantallen stormen die voorkomen gedurende disjunctie (d.w.z. niet-overlappende) willekeurig gekozen tijdsintervallen, min of meer onafhankelijk zijn en dat het verwachte aantal stormen (in het 'korte' stormseizoen) min of meer evenredig is met de lengte van het beschouwde tijdsinterval. Deze twee veronderstellingen (samen met de stationariteit) leiden tot de Poisson-verdeling als kansverdeling voor N , het totaal aantal geselecteerde waarnemingen gedurende stormseizoenen.

¹ In wat volgt zullen we i.p.v. jaar en jaarmaximum vrijwel steeds spreken over stormseizoen en stormseizoenmaximum, omdat alle in dit hoofdstuk te beschrijven methoden toegepast worden op geselecteerde HW's (of geselecteerde HW-opzetten) in het 'lange' stormseizoen.

Ook voor het 'lange' stormseizoen kan men eenvoudig nagaan dat N - het totaal aantal geselecteerde hoogwaterstanden gedurende s jaren - een Poisson-verdeling heeft.

Deze aanpak gaat dus uit van een stochastisch aantal N. We kunnen schrijven $N = \sum_{i=1}^s N_i$, waarbij N_i het aantal geselecteerde waarnemingen in het i-de stormseizoen voorstelt. De N_i 's zijn onafhankelijk en voldoen - op grond van de hiervoor gegeven argumentatie - aan een Poisson-verdeling. We noemen de parameter van de Poisson-verdeling q. M.a.w.

$$P(N_i = n) = \frac{q^n e^{-q}}{n!}, n = 0, 1, \dots, i = 1, \dots, s.$$

De parameter q van de Poisson-verdeling is gelijk aan de verwachte waarde van N_i ($i = 1, \dots, s$); het totaal aantal waarnemingen N heeft nu een Poisson-verdeling met parameter sq.

Zij nu L het aantal geselecteerde hoogwaterstanden in een willekeurig (toekomstig) stormseizoen en laten Y_1, \dots, Y_L de (fictieve) geselecteerde hoogwaterstanden in dat stormseizoen voorstellen. L is een stochastisch aantal - met Poisson-verdeling $P(L=l) = \frac{q^l e^{-q}}{l!}, l = 0, 1, 2, \dots$ - en onafhankelijk van de Y_i 's. De verwachte waarde van L is gelijk aan q; de parameter q is in principe onbekend, maar kan (zeer nauwkeurig) geschat worden door $\frac{N}{s}$, het gemiddeld aantal waarnemingen per stormseizoen, onder de X_i 's ($i = 1, \dots, N$), de vanuit het verleden beschikbare geselecteerde hoogwaterstanden. Merk op dat de verwachtingswaarde van $\frac{N}{s}$ gelijk is aan q en dat de standaardafwijking $\sigma(\frac{N}{s}) = \sqrt{\frac{q}{s}}$.

We concluderen dat $\sigma(\frac{N}{s})$ bij benadering gelijk is aan $\sqrt{N/s}$, een voor ons doel kleine waarde.

We bekijken nu (5.1) opnieuw en herschrijven deze relatie als volgt:

$$P(\max(Y_1, \dots, Y_L) \leq x_{p_0}) = 1 - p_0 \tag{5.2}$$

Omdat het linkerlid in (5.2) gelijk is aan

$$\begin{aligned} & \sum_{l=0}^{\infty} P(\max(Y_1, \dots, Y_l) \leq x_{p_0} | L=l) P(L=l) = \\ & = \sum_{l=0}^{\infty} F^l(x_{p_0}) \frac{q^l e^{-q}}{l!} = \exp\{-q(1 - F(x_{p_0}))\} \end{aligned} \tag{5.3}$$

vinden we dat (5.2) equivalent is met

$$\exp\{-q(1 - F(x_{p_0}))\} = 1 - p_0$$

ofwel met,

$$1 - F(x_{p_0}) = \frac{-\ln(1 - p_0)}{q} \tag{5.4}$$

Omdat $-\ln(1 - p_0) \approx p_0$, een nauwkeurige benadering voor kleine waarden van p_0 is, geldt ook

$$1 - F(x_{p_0}) \approx \frac{p_0}{q} \tag{5.5}$$

M.a.w. het gevraagde niveau x_{p_0} is bij goede benadering gelijk aan het niveau dat één geselecteerde hoogwaterstand met kans p overschrijdt, waarbij

$$p = \frac{p_0}{q} \approx \frac{p_0^s}{N}, \quad (5.6)$$

d.w.z., x_{p_0} is bij goede benadering gelijk aan het $(1-p)$ -de kwantiel² van de verdeling F . We kunnen dus i.p.v. (5.1) ook vergelijking (5.4) of, bij benadering, (5.5) oplossen.

De grote moeilijkheid is nu natuurlijk dat F - de verdeling van de waargenomen geselecteerde hoogwaterstanden bij een gegeven station - in principe volledig onbekend is. Daar komt nog bij dat omdat p (vgl. (5.6)) veel kleiner is dan $\frac{1}{N}$, het zonder enige kennis over F onmogelijk is om op basis van de waarnemingen X_1, \dots, X_N alleen een bruikbare schatter van het onbekende kwantiel x_{p_0} (vgl. (5.5) en (5.6)) te vinden. Immers, indien F volledig onbekend is, zal men F schatten m.b.v. de zgn. empirische verdelingsfunctie \hat{F}_N , welke gegeven is door $\hat{F}_N(x) = \frac{1}{N} \times$ [aantal X_i 's $\leq x$, $i = 1, \dots, N$], voor alle x . Dit betekent dat we i.p.v. (5.5) (vgl. ook (5.6)) de relatie $1 - \hat{F}_N(x_{p_0}) = \frac{p_0^s}{N}$ bekijken en concluderen dat een schatter voor x_{p_0} gegeven wordt door $\hat{x}_{p_0} = \hat{F}_N^{-1}(1 - \frac{p_0^s}{N}) = \max_{1 \leq i \leq N} X_i$, het maximum van de geselecteerde hoogwaterstanden X_1, \dots, X_N . Merk op dat deze schatter van x_{p_0} onafhankelijk van de exacte waarde van p_0 is! M.a.w: verschillende p_0 -waarden, tussen bijvoorbeeld 10^{-3} en 10^{-5} , leiden tot precies dezelfde schatter voor x_{p_0} . Dit nu is natuurlijk niet acceptabel en de hier beschreven 'naïeve' benadering van ons probleem is dan ook niet bruikbaar.

We concluderen dat het noodzakelijk is om een of andere veronderstelling omtrent F te maken.

In het navolgende komen een aantal verschillende mogelijkheden voor modellering van F aan de orde. Ieder van de te bespreken modellen correspondeert met een of andere veronderstelling omtrent F , de verdelingsfunctie van een geselecteerde hoogwaterstand. Afhankelijk van het gekozen model worden verschillende statistische schattingsmethoden voor het bepalen van het gevraagde hoogwaterniveau x_{p_0} gebruikt. In het bijzonder komen aan de orde:

- (1) de verdelingsvrije methode - c (afgekort: VVM-c)
 - (2) de verdelingsvrije methode - 0 (afgekort: VVM-0)
- en parametrische methoden:
- (3) de methode gebaseerd op de gegeneraliseerde Pareto-verdeling (GPV)
 - (4) de methode gebaseerd op een convolutiemodel (CON)
 - (5) de methode gebaseerd op stormseizoenmaxima (GEV-s)

5.2 Verdelingsvrije benadering

In het onderzoek ligt de nadruk op de zgn. *verdelingsvrije methoden*. Er wordt niet een of ander parametrisch model gepostuleerd voor F , maar - afgezien van enige aanvullende veronderstellingen, welke hier buiten

² het $(1-p)$ -de kwantiel van een verdeling F wordt (ruwweg) gedefinieerd als het niveau dat met kans p overschreden wordt.

beschouwing kunnen blijven - slechts verondersteld dat aan de volgende aanname is voldaan:

Laat X_1, X_2, \dots een rij onafhankelijke stochastische grootheden zijn, alle met dezelfde verdelingsfunctie F . Laat $M_n = \max_{1 \leq i \leq n} X_i$ het maximum zijn van de eerste n grootheden X_1, \dots, X_n in de rij X_1, X_2, \dots . We veronderstellen nu dat, voor grote waarden van n ,

$$P\left(\frac{\max(X_1, \dots, X_n) - b_n}{a_n} \leq x\right) \approx G_\gamma(x) \quad (5.7)$$

voor alle x waarvoor $1 + \gamma x > 0$. Hierbij stellen a_n en b_n geschikt gekozen constanten voor. De extreme-waardentheorie leert dat de klasse van alle mogelijke limietverdelingen ($n \rightarrow \infty$) voor genormeerde maxima $a_n^{-1} (\max(X_1, \dots, X_n) - b_n)$ precies gelijk is aan de klasse van de zgn. extreme-waardenverdelingen $G_\gamma, -\infty < \gamma < \infty$, waarbij

$$G_\gamma(x) = \exp\{-(1 + \gamma x)^{-1/\gamma}\} \quad (5.8)$$

voor alle x waarvoor $1 + \gamma x > 0$. M.a.w. (5.7) is geldig, met G_γ gegeven door (5.8). In het speciale geval $\gamma = 0$ moet $(1 + \gamma x)^{-1/\gamma}$ vervangen worden door e^{-x} . M.a.w. $G_0(x) = \exp(-e^{-x})$: de Gumbel-verdeling. De parameter γ , welke de extreme-waardenverdeling G_γ vastlegt, noemt men wel de extreme-waardenindex.

Bij de verdelingsvrije methoden (modellen (1) en (2) corresponderend met de hierna te beschrijven schattingsmethoden VVM-c en VVM-0) baseren we ons op de aanname (5.7). Deze eis kan ook als volgt geformuleerd worden: voor grote waarden van n geldt,

$$F^n(a_n x + b_n) \approx G_\gamma(x) \quad (5.9)$$

voor alle x waarvoor $1 + \gamma x > 0$. Merk op dat de voorwaarde (5.7) (zie ook (5.9)) uitgaat van een deterministisch aantal n . In dit onderzoek hebben we echter te maken met een stochastisch aantal N . Men kan laten zien dat de eis (5.7) impliceert dat, voor grote waarden van s ,

$$P\left(\frac{\max(X_1, \dots, X_N) - b(sq)}{a(sq)} \leq x\right) \approx G_\gamma(x) \quad (5.10)$$

voor alle x waarvoor $1 + \gamma x > 0$. Hierbij zijn a en b functies die zo gekozen zijn dat $a(n) = a_n$ en $b(n) = b_n$; merk op dat N een Poisson-verdeling met parameter sq heeft.

Hoe kunnen we nu het voorafgaande gebruiken om het kwantiel x_{p_0} m.b.v. relatie (5.7) te schatten? De details van de argumentatie welke leidt tot een schatter voor x_{p_0} - gebaseerd op X_1, \dots, X_N , de in het verleden waargenomen geselecteerde hoogwaterstanden bij een gegeven station - achterwege latend, geven we hier in het kort de essentie van de redenering weer. Voor een uitvoerige beschrijving verwijzen we naar [Dekkers et al., 1989].

Zij $k = k(n)$ een aantal dat begrensd blijft als n onbeperkt toeneemt, of een aantal dat onbeperkt toeneemt als n onbeperkt toeneemt. In beide situaties moet gelden dat $r = r(n) \rightarrow \infty$ als $n \rightarrow \infty$, met $r = \frac{n}{k}$. We vervangen n door r in (5.9) en krijgen, voor voldoende grote r ,

$$F^r(a_r x + b_r) \approx G_\gamma(x) \quad (5.11)$$

voor alle x met $1 + \gamma x > 0$. Schrijven we nu p voor $\frac{p_0}{q}$ (vgl. (5.6)) en x_p voor x_{p_0} dan leidt (5.11) ons (vgl. (5.5) en (5.6)) tot:

$$p \approx 1 - F(x_p) \approx 1 - G_\gamma^{1/\tau} \left(\frac{(x_p - b_r)}{a_r} \right) \quad (5.12)$$

Hierbij zijn a_r en b_r constanten gegeven door $a_r = a(r)$ en $b_r = b(r)$. Combinatie van (5.8) met (5.12) en toepassing van de elementaire benadering $1 - e^{-x} \approx x$ (voor kleine x) levert nu

$$p \approx 1 - \exp\left\{-\frac{1}{r} \left(1 + \gamma \frac{x_p - b_r}{a_r}\right)^{-\frac{1}{\gamma}}\right\} \approx \frac{1}{r} \left(1 + \gamma \frac{x_p - b_r}{a_r}\right)^{-\frac{1}{\gamma}} \quad (5.13)$$

zodat

$$x_p \approx \frac{\left(\frac{k}{np}\right)^\gamma - 1}{\gamma} a\left(\frac{n}{k}\right) + b\left(\frac{n}{k}\right) \quad (5.14)$$

met $k = \frac{n}{p}$. Benadering (5.14) vormt nu de basis voor de twee verdelingsvrije schattingsmethoden VVM-c (methode (1)) en VVM-0 (methode (2)). In feite wordt de schatter \hat{x}_p van het gevraagde kwantiel x_p gegeven door het rechterlid van (5.14), waarbij de extreme-waardenindex γ en de normeringsconstanten $a\left(\frac{n}{k}\right)$ en $b\left(\frac{n}{k}\right)$ natuurlijk geschat moeten worden.

In beide verdelingsvrije schattingsmethoden is gebruik gemaakt van de zgn. *momentenschatter* $\hat{\gamma}_n$ voor de parameter γ , welke gegeven wordt door [vgl. Dekkers et.al. 1989]

$$\hat{\gamma}_n = M_n^{(1)} + 1 - \frac{1}{2} \left(1 - \frac{(M_n^{(1)})^2}{M_n^{(2)}}\right)^{-1} \quad (5.15)$$

met

$$M_n^{(1)} = \frac{1}{m} \sum_{i=0}^{m-1} \ln X_{n-i:n} - \ln X_{n-m:n} \quad (5.16)$$

en

$$M_n^{(2)} = \frac{1}{m} \sum_{i=0}^{m-1} (\ln X_{n-i:n} - \ln X_{n-m:n})^2 \quad (5.17)$$

waarbij $X_{n-i:n}$ de $(n-i)$ -de in grootte is onder de eerste n waarnemingen X_1, \dots, X_n in de rij X_1, X_2, \dots . Verder zal het aantal gebruikte grootste waarnemingen $m + 1$ geschikt gekozen moeten worden. Merk op dat we stilzwijgend hebben verondersteld dat $X_{n-m:n}$ altijd positief is. In ons onderzoek is hieraan altijd voldaan.

De normeringsconstanten $a\left(\frac{n}{k}\right)$ en $b\left(\frac{n}{k}\right)$ in (5.14) kunnen ook geschat worden. We gebruiken hiervoor de schatters $\hat{a}\left(\frac{n}{k}\right)$ en $\hat{b}\left(\frac{n}{k}\right)$ gegeven door

$$\hat{a}\left(\frac{n}{k}\right) = \frac{X_{n-k:n}}{k} \left(\sum_{i=0}^{k-1} \ln X_{n-i:n} - \ln X_{n-k:n} \right) (1 - \min(0, \hat{\gamma}_n)) \quad (5.18)$$

en $\hat{b}(\frac{n}{k}) = X_{n-k:n}$. Combineren we nu (5.18) met (5.14) t/m (5.17) dan vinden we de schatter

$$\hat{x}_p = \hat{\gamma}_n^{-1} \left\{ \left(\frac{k}{np} \right)^{\hat{\gamma}_n} - 1 \right\} \frac{X_{n-k:n}}{k} \left(\sum_{i=0}^{k-1} \ln X_{n-i:n} - \ln X_{n-k:n} \right) (1 - \min(0, \hat{\gamma}_n)) + X_{n-k:n} \quad (5.19)$$

voor het gevraagde kwantiel x_p . Hierbij zal niet alleen m (vgl. (5.16) en (5.17)) maar ook k geschikt gekozen moeten worden; $k + 1$ en $m + 1$ nu zijn het aantal grootste waarnemingen dat gebruikt wordt bij de schatting van γ respectievelijk van x_p . Verder vervangen we n door N en $p = \frac{p_0}{q}$ (vgl. (5.6)) door $\frac{p_0^5}{N}$. De fout die bij deze laatste substitutie gemaakt wordt is klein en wordt verwaarloosd.

We bespreken nu achtereenvolgens de verdelingsvrije schattingsmethoden VVM-c en VVM-0. Beide methoden zijn gebaseerd op dezelfde formule voor de schatting \hat{x}_p (vgl. (5.19)) voor het gevraagde kwantiel x_p , maar leiden tot verschillende betrouwbaarheidsintervallen (d.w.z. intervallschattingen) voor x_p . We merken hierbij nog op dat in de praktijk de schattingen \hat{x}_p verkregen m.b.v. de VVM-c- en de VVM-0-methode i.h.a. zullen verschillen. Dit is een gevolg van het feit dat bij toepassingen van beide methoden in de regel verschillende waarden voor zowel k als m gebruikt zullen worden. In het onderzoek is eerst de schattingsmethode VVM-c ontwikkeld en toegepast, pas in een later stadium werd ook de schattingsmethode VVM-0 ontwikkeld en toegepast. Hoewel zal blijken dat de methode VVM-0 voor ons doel het meest adequaat lijkt, beginnen we - in overeenstemming met de tijdens dit onderzoek gevolgde werkwijze - met een bespreking van VVM-c. De VVM-c-methode geeft een betrouwbaarheidsinterval voor het gewenste kwantiel x_p , gebaseerd op \hat{x}_p (vgl. (5.19)). We maken gebruik van het feit dat, voor grote waarden van n , de verdeling van

$$\frac{x_p - \hat{x}_p}{X_{n-k:n} M_n^{(1)}} \quad (5.20)$$

bij benadering gelijk is aan een volledig gespecificeerde (gecompliceerde) limiet-verdeling, mits,

$$\begin{aligned} & \text{als } n \text{ onbeperkt toeneemt, } m \text{ onbeperkt toeneemt, terwijl } k \\ & \text{begrensd blijft, en } np \rightarrow c, \text{ voor een zekere constante } c \\ & \text{waarvoor geldt } 0 < c < \infty \end{aligned} \quad (5.21)$$

De grootheid $M_n^{(1)}$ in (5.20) is gelijk aan (5.16) met m vervangen door k . Merk op dat het noodzakelijk is k veel kleiner dan m te kiezen. Verder betekent (5.21) dat $p = p(n) \rightarrow 0$ als $n \rightarrow \infty$ in een tempo dat bij benadering gelijk is aan cn^{-1} . M.a.w. $p = p(n) \approx cn^{-1}$ voor een zekere positieve en eindige constante c .

Met behulp van het resultaat (5.20) kan een benaderend betrouwbaarheidsinterval voor x_p geconstrueerd worden: bepaal het $(\frac{\alpha}{2})$ -de en $(1 - \frac{\alpha}{2})$ -de kwantiel van de bovengenoemde limietverdeling - deze noemen we $l_{\alpha/2}$ en $l_{1-\alpha/2}$ - dan wordt het betrouwbaarheidsinterval voor x_p gegeven door

$$(\hat{x}_p - l_{1-\alpha/2} X_{n-k:n} M_n^{(1)}, \hat{x}_p - l_{\alpha/2} X_{n-k:n} M_n^{(1)}) \quad (5.22)$$

De kans dat het interval (5.22) de onbekende parameter x_p bevat is bij benadering $1 - \alpha$. Evenals in (5.20) is $M_n^{(1)}$ in (5.22) gelijk aan (5.16) met m vervangen door k . We merken op dat het interval (5.22) niet symmetrisch is om de (punt)schatting \hat{x}_p . Voor meer details over de VVM-c methode verwijzen we naar [Dekkers et. al. 1989].

We merken op dat de aanname $np \rightarrow c$, als $n \rightarrow \infty$ (vgl. (5.21)), welke noodzakelijk is voor de geldigheid van de VVM-c methode, voor ons probleem weinig realistisch is. Immers, we kunnen de grootheid np interpreteren als het verwachte aantal overschrijdingen van niveau x_p . M.a.w. de eis $0 < c < \infty$ betekent dat we, bij benadering, c overschrijdingen van het niveau x_p verwachten als n voldoende groot is. We weten echter dat het niveau x_p dat correspondeert met bijvoorbeeld $p_0 = 10^{-4}$ in feite niet is overschreden. Dit, en het feit dat in ons onderzoek n vervangen wordt door N , $p \approx \frac{p_0 s}{N}$, zodat $np \approx p_0 s$, en $c \approx 0$, suggereert dat de keuze $c = 0$ in de relatie $np \rightarrow c$ (vgl. (5.21)) realistisch voor ons probleem is. Zoals we nog zullen zien ligt juist deze laatste veronderstelling ten grondslag aan de methode VVM-0. Een tweede nadeel van de VVM-c methode is dat de variabiliteit van de schatter $\hat{\gamma}_n$ van γ een veel kleinere rol speelt dan de variabiliteit die optreedt in de schatting \hat{x}_p (vgl. (5.19)) van x_p , bij gegeven γ .

De VVM-0-methode is een belangrijke variant op de hiervoor beschreven VVM-c methode. De VVM-0-methode is uitvoerig beschreven in [de Haan en Rootzen, 1989]. Hier volstaan we met de volgende korte beschrijving. In de eerste plaats wordt bij de VVM-0 methode de aanname (5.21) vervangen door:

$$\text{als } n \text{ onbeperkt toeneemt, } m \text{ onbeperkt toeneemt, terwijl } k \\ \text{gelijk aan } m \text{ gekozen wordt, en } np \rightarrow 0 \quad (5.23)$$

M.a.w. het aantal gebruikte grootste waarnemingen voor de schatting $\hat{\gamma}_n$ van γ wordt nu ook gebruikt bij de schatting \hat{x}_p van x_p . Verder wordt de aanname $np \rightarrow c$, voor zekere $0 < c < \infty$, welke nodig was voor de VVM-c-methode, nu vervangen door de eis $np \rightarrow 0$, als $n \rightarrow \infty$. We veronderstellen dus weer dat $p = p(n) \rightarrow 0$, maar nu in een sneller tempo. In zekere zin is de veronderstelling $np \rightarrow 0$ voor ons doel realistischer dan de aanname $np \rightarrow c$, voor zekere $0 < c < \infty$: indien $np \rightarrow 0$, als $n \rightarrow \infty$ dan is het verwachte aantal overschrijdingen van het niveau x_p , bij benadering, gelijk aan nul. Dit is in overeenstemming met de feitelijke situatie in dit onderzoek. Indien veronderstelling (5.23) geldig is, dan kunnen we opnieuw gebruik maken van formule (5.19) voor de puntschatting \hat{x}_p van het gezochte kwantiel x_p . Wel moeten nu de waarden van k en m aan elkaar gelijk gekozen worden. In plaats van het door de limietverdeling van (5.20) bepaalde betrouwbaarheidsinterval (5.22) (dat nu niet meer correct is) geldt nu dat $\hat{x}_p - x_p$ bij benadering normaal verdeeld is. Precieser gezegd: voor grote waarden van n , heeft

$$\left\{ \hat{\alpha} \left(\frac{n}{k} \right) \int_0^{k/np} s^{\gamma-1} \ln s ds \right\}^{-1} k^{+\frac{1}{2}} (\hat{x}_p - x_p) \quad (5.24)$$

bij benadering, een *normale verdeling* met gemiddelde waarde nul en variantie σ^2 , welke gegeven wordt door

$$\sigma^2 = \sigma^2(\gamma) \quad (5.25)$$

waarbij

$$\sigma^2(\gamma) = \begin{cases} 1 + \gamma^2, & \gamma \geq 0 \\ \frac{(2-3\gamma)^2}{1-2\gamma} + (5-11\gamma)(1-3\gamma)(1-4\gamma)(1-2\gamma)^2 - \\ -4 \frac{(2-3\gamma)(1-2\gamma)}{1-3\gamma}, & \gamma < 0 \end{cases} \quad (5.26)$$

M.a.w. het benaderend betrouwbaarheidsinterval voor x_p , gegeven door de VVM-0 methode, wordt gegeven door

$$\left(\hat{x}_p - u_{1-\alpha/2} \hat{a}(\frac{n}{k}) \int_0^{k/np} s^{\hat{\gamma}_n-1} \ln s ds \frac{\hat{\sigma}_n}{\sqrt{k}}, \hat{x}_p + u_{1-\alpha/2} \hat{a}(\frac{n}{k}) \int_0^{k/np} s^{\hat{\gamma}_n-1} \ln s ds \frac{\hat{\sigma}_n}{\sqrt{k}} \right) \quad (5.27)$$

waarbij $\hat{\sigma}_n$ gelijk is aan $\sigma(\hat{\gamma}_n)$; de normeringsconstante $\hat{a}(\frac{n}{k})$ wordt weer gegeven door (5.18), terwijl $u_{1-\alpha/2}$ het $(1 - \frac{\alpha}{2})$ -de kwantiel van de standaard normale verdeling voorstelt.

In het onderzoek is naast de VVM-c en VVM-0 methode ook nog aandacht besteed aan een verdelingsvrije methode gebaseerd op koppeling van HW-opzetten en astronomische standen. Uitgangspunt bij deze methode is de veronderstelling dat geselecteerde HW's beschouwd kunnen worden als de som van een geselecteerde hoogwateropzet en bijbehorende astronomische hoogwaterstand, waarbij aangenomen wordt dat de HW-opzetten en de astronomische hoogwaterstanden stochastisch onafhankelijk zijn. Indien we verder aannemen dat zowel de HW-opzetten als de astronomische hoogwaterstanden voldoen aan de extreme waarden voorwaarden (5.7), d.w.z. (5.7) geldt met de X_i 's vervangen door opzetten respectievelijk astronomische hoogwaterstanden), met extreme waardenindex $\gamma_1 < 0$ respectievelijk $\gamma_2 < 0$, dan kan men laten zien [de Haan 1986] dat dan ook de HW's aan de extreme-waarden voorwaarde (5.7) voldoen, met extreme-waardenindex (vormparameter)

$$\gamma = (\gamma_1^{-1} + \gamma_2^{-1})^{-1}.$$

Het idee is nu γ_1 te schatten met behulp van de geselecteerde HW-opzetten. De vormparameter γ_2 van de astronomische hoogwaterstanden kan in principe exact bepaald worden uit de volledig bekende verdeling van de astronomische hoogwaterstanden. Met behulp van de 'koppelingsformule' $\gamma = (\gamma_1^{-1} + \gamma_2^{-1})^{-1}$ kan nu de vormparameter van de extreme-waardenverdeling van de HW's geschat worden. Toepassing van deze methode leidde echter tot sterk afwijkende resultaten (zie § 7.6). De reden hiervan lijkt te liggen in het feit dat de 'koppelingsformule' niet exact geldig is, maar een benadering is, welke in ons onderzoek klaarblijkelijk geen acceptabele antwoorden oplevert.

Tenslotte vermelden we nog dat in het onderzoek naast de momenten schatter $\hat{\gamma}_n$ aanvankelijk ook nog gewerkt is met een schatter voor de parameter $\hat{\gamma}$ welke is voorgesteld door Pickands (1975) - zie ook [Dekkers en De Haan 1989]. De Pickands-schatter $\hat{\gamma}_n^p$ wordt gegeven

door

$$\hat{\gamma}_n^P = \log \frac{X_{n-m+1:n} - X_{n-2m+1:n}}{X_{n-2m+1:n} - X_{n-4m+1:n}}$$

Toepassing van $\hat{\gamma}_n^P$ i.p.v. $\hat{\gamma}_n$ leidde echter tot minder bevredigende resultaten (zie § 7.6).

5.3 Parametrische aanpak

We beschrijven nu achtereenvolgens de methoden (3), (4), en (5). Deze statistische schattingsmethoden zijn - in tegenstelling tot de hiervoor besproken methoden (1) en (2) - gebaseerd op een parametrisch model. Evenals in het voorafgaande zullen we weer veronderstellen dat we de beschikking hebben over onafhankelijke waarnemingen X_1, \dots, X_n - dit zijn de in het verleden verkregen geselecteerde hoogwaterstanden aan een gegeven station - alle afkomstig uit een in principe onbekende verdeling met verdelingsfunctie F . M.a.w. F is de verdelingsfunctie van een geselecteerde hoogwaterstand bij een gegeven station.

We beginnen met methode (3): de methode gebaseerd op de *gegeneraliseerde Pareto-verdeling (GPV)*. Laat X_1, X_2, \dots een rij onafhankelijke stochastische grootheden zijn, alle met dezelfde verdelingsfunctie F . Veronderstel dat (evenals bij de methoden (1) en (2)) aan de klassieke extreme-waardenvoorwaarde (5.7) is voldaan: Voor grote waarden van n geldt (vgl. (5.7)):

$$P\left(\frac{\max(X_1, \dots, X_n) - b_n}{a_n} \leq x\right) \approx G_\gamma(x) \quad (5.28)$$

voor alle x waarvoor $1 + \gamma x > 0$. Hierbij stellen a_n en b_n geschikt gekozen constanten voor; de extreme-waardenverdelingen G_γ , $-\infty < \gamma < \infty$, zijn gegeven in (5.8).

Bepaal nu, onder de waarnemingen X_1, \dots, X_n diegenen die boven een bepaalde goed gekozen drempel u uitsteken. Het aantal van deze geselecteerde waarnemingen, hierna uitstekende waarnemingen genoemd, is stochastisch. Als dit aantal boven de drempel u uitstekende waarnemingen nu gefixeerd gedacht wordt - d.w.z. voorwaardelijk gegeven dit aantal - dan zijn de uitstekende waarnemingen weer onafhankelijk en hebben alle dezelfde kansverdeling, die rechtstreeks uit de oorspronkelijke kansverdeling F volgt: de voorwaardelijke verdelingsfunctie van de uitstekende waarnemingen minus de drempel u , wordt gegeven door

$$\begin{aligned} F_u(y) &= P(X_1 - u \leq y | X_1 > u) = \\ &= \frac{P(u < X_1 \leq y + u)}{P(X_1 > u)} = \frac{F(y + u) - F(u)}{1 - F(u)} \end{aligned} \quad (5.29)$$

voor alle $y > 0$ en gegeven drempel u .

Men kan nu aantonen [Pickands 1975] dat (5.7) - vgl. (5.28) - equivalent is aan de bewering dat voor voldoende hoge drempels u , de GPV-verdeling een goede benadering is voor F_u . M.a.w. voor grote waarden van u geldt

$$F_u(y) \approx G(y; \sigma(u), \gamma) \quad (5.30)$$

waarbij

$$G(y; \sigma, \gamma) = 1 - (1 + \gamma \frac{y}{\sigma})^{-\frac{1}{\gamma}} \quad (5.31)$$

voor alle γ waarvoor $1 + \gamma \frac{y}{\sigma} > 0$. De verdelingsfunctie $G(y; \sigma, \gamma)$ noemt men de gegeneraliseerde Pareto-verdeling, de GPV (met vormparameter γ en schaalparameter σ). In het speciale geval $\gamma = 0$ moet $(1 + \gamma \frac{y}{\sigma})^{-1/\sigma}$ in (5.31) vervangen worden door $\exp(-\frac{y}{\sigma})$. M.a.w. $G(y; \sigma, 0) = 1 - \exp(-\frac{y}{\sigma})$, $\gamma > 0$, de exponentiële verdeling. De parameter γ in (5.31) is dezelfde als de extreme-waardenindex in (5.28); de schaalparameter $\sigma = \sigma(u)$ in (5.30) is afhankelijk van de drempelwaarde u . De drempel u moet geschikt (d.w.z. afhankelijk van de oorspronkelijke kansverdeling F) gekozen worden. Om in te zien dat (5.30) inderdaad volgt uit (5.28) merken we allereerst op dat (5.28) equivalent is met

$$F^n(a_n x + b_n) \approx G_\gamma(x) \quad (5.32)$$

voor alle x waarvoor $1 + \gamma x > 0$ en grote waarden van n . Nemen we nu logaritmen in (5.32) dan zien we dat geldt

$$n\{-\ln F(a_n x + b_n)\} \approx -\ln G_\gamma(x) \quad (5.33)$$

Omdat $-\ln(1+x) \approx x$ voor kleine x , volgt nu uit (5.33)³ dat

$$n\{1 - F(a_n x + b_n)\} \approx -\ln G_\gamma(x) \quad (5.34)$$

en, in het bijzonder, met de keuze $x = 0$,

$$n\{1 - F(b_n)\} \approx -\ln G_\gamma(0) = 1 \quad (5.35)$$

zodat we kunnen concluderen dat

$$\frac{1 - F(a_n x + b_n)}{1 - F(b_n)} \approx -\ln G_\gamma(x) \quad (5.36)$$

Men kan aantonen dat a_n in feite een functie is van b_n ; d.w.z. $a_n = f(b_n)$. Vervangen we nu b_n door t dan vinden we dat

$$\frac{1 - F(f(t)x + t)}{1 - F(t)} \approx -\ln G_\gamma(x) \quad (5.37)$$

voor alle t in een omgeving van het rechter-eindpunt x^* van de verdeling van F . (Merk op dat x^* eindig of oneindig kan zijn). Nu is (vgl. (5.8))

$-\ln G_\gamma(x) = (1 + \gamma x)^{-1/\gamma}$ voor alle x waarvoor $1 + \gamma x > 0$. Verder geldt

$$\begin{aligned} \frac{1 - F(f(t)x + t)}{1 - F(t)} &= \frac{P(X_1 > f(t)x + t)}{P(X_1 > t)} = \\ &= P(X_1 > f(t)x + t | X_1 > t) = \\ &= P\left(\frac{X_1 - t}{f(t)} > x | X_1 > t\right) \end{aligned} \quad (5.38)$$

³ merk op dat $F(a_n x + b_n) - 1$ klein is voor grote x .

Vervangen we nu t door u dan vinden we met $y = xf(t)$ m.b.v. (5.37) dat

$$P(X_1 - u > y | X_1 > u) \approx (1 + \gamma \frac{y}{\sigma})^{-\frac{x}{\gamma}} \quad (5.39)$$

waaruit (5.30) onmiddellijk volgt.

Het uitgangspunt bij de schattingsmethode (3) is nu dat F_u , voor grote waarden van n , exact gelijk is aan een GPV-verdeling. M.a.w. de benadering (5.30) wordt vervangen door de gelijkheid:

$$F_u(y) = G(y; \sigma(u), \gamma) \quad (5.40)$$

voor alle y met $1 + \gamma \frac{y}{\sigma} > 0$. We veronderstellen nu dus dat de uitstekende waarnemingen minus de drempel u - voorwaardelijk gegeven hun aantal - een kansverdeling hebben uit de klasse van de gegeneraliseerde Pareto-verdelingen. We hebben gezien dat de benadering (5.30) waar is, als aan de klassieke extreme-waardenaanname (5.28) voldaan is. We vervangen nu (5.30) door de meer strikte veronderstelling dat het parametrische model (5.40) exact geldig is.

Op basis van het parametrische GPV-model (5.32) kunnen nu - gebruik makend van de uitstekende waarnemingen - de parameters van de GPV-verdeling geschat worden m.b.v. de methode van de grootste aannemelijkheid. De uitstekende waarnemingen minus de drempel u noemen we verder de overschrijdingen.

Deze overschrijdingen worden aangeduid met $Y_1, \dots, Y_{n(u)}$, waarbij $n(u)$ het feitelijk aantal opgetreden uitstekende waarnemingen voorstelt. Voorwaardelijk gegeven het aantal $n(u)$ overschrijdingen kunnen we de kans bepalen dat het maximum van $n(u)$ overschrijdingen een hoogte y overschrijdt:

$$\begin{aligned} P(\max_{1 \leq i \leq n(u)} Y_i \geq y) &= 1 - P(\max_{1 \leq i \leq n(u)} Y_i < y) \\ &= 1 - \prod_{i=1}^{n(u)} P(Y_i < y) = 1 - F_u^{n(u)}(y) = \\ &= 1 - G^{n(u)}(y; \sigma(u), \gamma) = 1 - \{1 - (1 + \gamma \frac{y}{\sigma})^{-\frac{1}{\gamma}}\}^{n(u)} \approx \\ &\approx 1 - \exp\{-n(u)(1 + \gamma \frac{y}{\sigma})^{-\frac{1}{\gamma}}\} \end{aligned} \quad (5.41)$$

waarbij we gebruik gemaakt hebben van (5.31), (5.40) en het bekende feit dat, voor iedere waarde van x , $(1 - \frac{x}{n})^n \approx e^{-x}$ een goede benadering is voor grote waarden van n . Met behulp van de benadering (5.41) voor de kans dat het maximum van $n(u)$ overschrijdingen van een drempel u een hoogte y overschrijdt (voorwaardelijk gegeven het aantal $n(u)$), kan men nu een formule afleiden voor het kwantiel x_{p_0} (vgl. (5.1)) welke het stormseizoen-maximum van de uitstekende waarnemingen met kans p_0 overschrijdt. Met $n(u)$ overschrijdingen van drempel u , verdeeld over s stormseizoenen vinden we de benadering

$$x_{p_0} \approx -\frac{\sigma}{\gamma} \{1 - (\frac{-s \ln(1 - p_0)}{n(u)})^{-\gamma}\} + u \quad (5.42)$$

Om in te zien dat (5.42) correct is, merken we op dat - nu uitgaande van (5.41) - de argumentatie leidend van (5.2) naar (5.4) herhaald kan worden, met L vervangen door $\frac{n(u)}{s}$, het geschatte gemiddelde aantal overschrijdingen per stormseizoen en de Y_i 's in (5.2) nu opgevat als overschrijdingen. Wegens (5.41) kan nu een benadering voor x_{p_0} - u gevonden worden door de vergelijking

$$1 - \exp\left\{-\frac{n(u)}{s}\left(1 + \gamma\frac{y}{\sigma}\right)^{-\frac{1}{\gamma}}\right\} = p_0 \quad (5.43)$$

op te lossen. Nemen we nu logaritmen in (5.43) dan zien we dat (5.42) inderdaad geldt. M.a.w. het gevraagde kwantiel x_{p_0} kan worden bepaald m.b.v. het rechterlid van (5.42), waarbij de vormparameter en de schaalparameter natuurlijk geschat moeten worden. Bij gegeven drempel u en overschrijdingen $Y_1, \dots, Y_{n(u)}$ worden de parameters γ en σ geschat. [Smith 1984, 1987] geeft meest aannemelijke schatters $\hat{\gamma}_{n(u)}$ en $\hat{\sigma}_{n(u)}$ voor de parameters γ en σ , alsmede formules voor $Var(\hat{\gamma}_{n(u)})$, $Var(\hat{\sigma}_{n(u)})$ en $Cov(\hat{\sigma}_{n(u)}, \hat{\gamma}_{n(u)})$. Vervangen we nu in het rechterlid van (5.42) de parameters en door schatters $\hat{\gamma}_{n(u)}$ en $\hat{\sigma}_{n(u)}$ dan krijgen we

$$\hat{x}_{p_0} = \frac{-\hat{\sigma}_{n(u)}}{\hat{\gamma}_{n(u)}} \left\{ 1 - \left(\frac{-s \ln(1 - p_0)}{n(u)} \right)^{-\hat{\gamma}_{n(u)}} \right\} + u \quad (5.44)$$

de schatter \hat{x}_{p_0} voor het kwantiel x_{p_0} van de stormseizoenmaxima-verdeling dat met kans p_0 overschreden wordt. Men bepaalt eenvoudig een schatting voor de variantie van \hat{x}_{p_0} met behulp van de formule

$$\begin{aligned} Var(\hat{x}_{p_0}) \approx & \left(\frac{\partial h}{\partial \gamma}\right)^2 Var(\hat{\gamma}_{n(u)}) + 2 \frac{\partial h}{\partial \gamma} \frac{\partial h}{\partial \sigma} Cov(\hat{\gamma}_{n(u)}, \hat{\sigma}_{n(u)}) + \\ & + \left(\frac{\partial h}{\partial \sigma}\right)^2 Var(\hat{\sigma}_{n(u)}) \end{aligned} \quad (5.45)$$

waarbij de functie h gegeven wordt door

$$h(\gamma, \sigma) = -\frac{\sigma}{\gamma} \left\{ 1 - \left(\frac{-s \ln(1 - p_0)}{n(u)} \right)^{-\gamma} \right\}. \quad (5.46)$$

Hiertoe moeten de partiële afgeleiden $\frac{\partial h}{\partial \sigma}$ en $\frac{\partial h}{\partial \gamma}$ bepaald worden en vervolgens geëvalueerd in de punten $\hat{\gamma}_{n(u)}$ en $\hat{\sigma}_{n(u)}$. Ook de formules voor $Var(\hat{\gamma}_{n(u)})$, $Cov(\hat{\gamma}_{n(u)}, \hat{\sigma}_{n(u)})$ en $Var(\hat{\sigma}_{n(u)})$, welke gegeven worden door

$$Var(\hat{\gamma}_{n(u)}) = \frac{2}{n(u)(1 + \gamma)(1 + 2\gamma)}$$

$$Var(\hat{\sigma}_{n(u)}) = \frac{2}{n(u)\sigma^2(1 + 2\gamma)}$$

en

$$Cov(\hat{\gamma}_{n(u)}, \hat{\sigma}_{n(u)}) = \frac{-1}{n(u)\sigma(1 + \gamma)(1 + 2\gamma)} \quad (5.47)$$

moeten worden geëvalueerd in de punten $\hat{\gamma}_{n(u)}$ en $\hat{\sigma}_{n(u)}$.

Zoals de formules in (5.47) al suggereren, is de hier gegeven bepaling van de variabiliteit in \hat{x}_{p_0} - via (5.45), (5.46) en (5.47) - alleen geldig indien $\gamma > -\frac{1}{2}$. M.a.w. bij toepassing van de hiervoor beschreven analyse, gebaseerd op het parametrische GPV-model (5.40), moet de a priori aanname $\gamma > -\frac{1}{2}$ gemaakt worden. In dit onderzoek lijkt hieraan inderdaad te zijn voldaan. Op basis van (5.45) t/m (5.47) kunnen nu benaderende betrouwbaarheids intervallen

$$(\hat{x}_{p_0} - u_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{x}_{p_0})}, \hat{x}_{p_0} + u_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{x}_{p_0})}) \quad (5.48)$$

voor het gewenste kwantiel x_{p_0} bepaald worden.

In ons onderzoek wordt nu het deterministische aantal n door een stochastisch aantal N vervangen en dus ook $n(u)$ door $N(u)$, het aantal geselecteerde hoogwaterstanden die boven een geschikt gekozen drempel u uitsteken. In [Smith, 1987] worden de eigenschappen van de hier beschreven methode (3) onderzocht, indien het exacte parametrische model (5.40) vervangen wordt door de meer realistische veronderstelling (5.30). Onder een extra veronderstelling welke betrekking heeft op de keuze van de drempel in relatie tot de steekproefomvang, blijken de hier beschreven resultaten ongewijzigd van kracht te blijven. In meer algemene situaties blijkt dat een correctie voor onzuiverheid in schattingen voor γ en σ moet worden toegepast.

Methode (4) - de schattingsmethode gebaseerd op het *convolutiemodel CON* is een belangrijke variant op de GPV-methode. Uitgangspunt is nu dat de onafhankelijke waarnemingen $X_i, i = 1, \dots, N$ - dit zijn de in het verleden verkregen geselecteerde hoogwaterstanden van een gegeven station - nu beschouwd worden als de som van een geselecteerde hoogwateropzet $Y_{i0}, i = 1, \dots, N$ en een bijbehorende astronomische hoogwaterstand $Z_i, i = 1, \dots, N$. De cruciale aanname is nu dat Y_{i0} en Z_i voor iedere $i = 1, \dots, N$, stochastisch onafhankelijk verondersteld worden (zie § 4.5). M.a.w.

$$\begin{aligned} F(x) &= P(X_i \leq x) = \\ &= \int_0^\infty F_0(x-z) dF_a(z) \end{aligned} \quad (5.49)$$

waarbij F_0 de verdelingsfunctie van de hoogwateropzetten, de Y_{i0} 's, en F_a de verdelingsfunctie van de astronomische hoogwaterstanden, de Z_i 's, voorstellen. De integraal in (5.49) staat bekend als de convolutie-integraal. De verdeling van de astronomische hoogwaterstanden is in principe volledig bekend (zie § 4.6). Voor wat betreft F_0 , de verdeling van de geselecteerde hoogwateropzetten Y_{i0}, \dots, Y_{N0} bij een gegeven station, zullen we veronderstellen dat aan de klassieke extreme-waardenvoorwaarde voldaan is. D.w.z. de argumentatie leidend van (5.28) naar (5.40) kan nu herhaald worden voor hoogwateropzetten (de Y_{i0} 's) in plaats van hoogwaterstanden (de X_i 's). We zullen nu veronderstellen dat F_0 de eigenschap heeft dat de bijbehorende voorwaardelijke verdeling van de uitstekende geselecteerde hoogwateropzetten, minus een geschikt gekozen drempel u_0 , namelijk

$$F_{0u_0}(y) = \frac{F_0(y + u_0) - F_0(u_0)}{1 - F_0(u_0)} \quad (5.50)$$

exact gelijk is aan een gegeneraliseerde Pareto-verdeling. M.a.w.

$$F_{0u_0}(y) = G(y; \sigma_0(u_0), \gamma_0) \quad (5.51)$$

voor alle y met $1 + \gamma_0 \frac{y}{\sigma_0} > 0$. Op basis van dit parametrische GPV-model (5.51) voor de geselecteerde hoogwateropzetten kunnen nu - gebruikmakend van de geselecteerde uitstekende hoogwateropzetten bij een gegeven station de parameters σ_0 en γ_0 van de GPV-verdeling geschat worden m.b.v. de methode van de grootste aannemelijkheid. Analoot aan (5.49) vinden we nu een benadering \tilde{F} voor F :

$$\begin{aligned} \tilde{F}(x) &= \int_0^\infty F_{0u_0}(x - z) dF_a(z), \\ &= \int_0^\infty G(x - z; \sigma_0(u_0), \gamma_0) dF_a(z), \end{aligned} \quad (5.52)$$

onder gebruikmaking van (5.51). Benadering \tilde{F} (vgl. (5.52)) voor F , de verdeling van de geselecteerde hoogwaterstanden X_i , $i = 1, \dots, N$, vormt nu de basis van de CON-methode. In feite wordt de schatter x_{p_0} van het gevraagde kwantiel x_{p_0} (vgl. (5.1)) van de stormseizoenmaximaverdeling van de geselecteerde hoogwaterstanden gevonden door de vergelijking

$$\tilde{F}(x) = (1 - p_0)^{\frac{x}{\sigma_0}} \quad (5.53)$$

op te lossen, met $N(u_0) =$ aantal geselecteerde hoogwateropzetten Y_{i_0} ($1 \leq i \leq N$) die boven een goed gekozen drempel u_0 uitsteken, en met F als in (5.52). Vervolgens moeten natuurlijk σ_0 en γ_0 vervangen worden door $\hat{\sigma}_{0N}$ en $\hat{\gamma}_{0N}$, de meest aannemelijke schattingen van σ_0 en σ_0 gebaseerd op de $N(u_0)$ uitstekende geselecteerde hoogwateropzetten. Om in te zien dat (5.53) correct is, merken we op dat - nu uitgaande van (5.52) - de argumentatie leidend van (5.2) naar (5.5) en (5.6) herhaald kan worden, met L vervangen door $\frac{N(u_0)}{s}$ en F in (5.5) vervangen door \tilde{F} . Benaderende betrouwbaarheidsintervallen gebaseerd op de CON-schatter \hat{x}_{p_0} van het gevraagde kwantiel x_{p_0} zijn van de vorm (5.48), waarbij $\text{Var}(\hat{x}_{p_0})$ nu uiteraard niet meer m.b.v. (5.45) bepaald kan worden. In ons onderzoek is $\text{Var}(x_{p_0})$ geschat m.b.v. Monte-Carlo-simulaties (zie § 6.6).

We merken nog op dat de benadering \tilde{F} voor ons doel voldoende dicht bij F zal liggen indien aan het analogon van (5.31) voor hoogwateropzetten is voldaan. Verder moet het rechter-eindpunt van de verdeling F_a van de astronomische hoogwaterstanden voldoende klein zijn in vergelijking met het gezochte kwantiel x_{p_0} van F , minus de drempel u_0 van de hoogwateropzetten. M.a.w. we veronderstellen dat de mogelijkheid dat 'kleine' hoogwateropzetten in combinatie met 'grote' astronomische hoogwaterstanden tot 'hoge' hoogwaterstanden leiden verwaarloosd kan worden. In ons onderzoek lijkt dit inderdaad het geval.

Methode (5) - de traditionele schattingsmethode gebaseerd op stormseizoenmaxima (GEV-s) - is ongetwijfeld conceptueel de eenvoudigste van de hier besproken technieken. Uitgaande van de N geselecteerde HW's X_1, \dots, X_N , bepalen we de maximale hoogwaterstand voor ieder van de s stormseizoenen. (Het aantal verschillende stormseizoenen waar de X_i 's over verdeeld zijn is precies gelijk aan het aantal jaren). M.a.w. we beschikken over s stormseizoenmaxima $X_{\max}^1, \dots, X_{\max}^s$, hoogste waarnemingen gemeten tijdens s stormseizoenen. We veronderstellen nu dat de X_{\max}^i 's, $i = 1, \dots, s$, stochastisch onafhankelijk zijn en alle dezelfde verdeling bezitten, welke exact gelijk is aan een extreme-waardenverdeling, met (onbekende) locatie en schaalparameter σ . M.a.w., voor iedere $i = 1, \dots, s$, geldt

$$P(X_{\max}^i \leq x) = G_\gamma\left(\frac{x-\mu}{\sigma}\right) \quad (5.54)$$

waarbij G_γ gegeven wordt door (5.8). Met behulp van (5.35) vinden we eenvoudig het kwantiel dat het stormseizoenmaximum van geselecteerde HW's met kans p_0 overschrijdt. We moeten de vergelijking $G_\gamma\left(\frac{x-\mu}{\sigma}\right) = 1 - p_0$ oplossen. Wegens (5.8) vinden we voor het gewenste kwantiel de formule

$$x_{p_0} = \sigma\gamma^{-1}\{-1 + (-\ln(1 - p_0))^{-\gamma}\} + \mu \quad (5.55)$$

(vgl. (5.1)), Hierbij is stilzwijgend aangenomen dat de verdeling van de (gewone) jaarmaxima benaderd mag worden door die van de stormseizoenmaxima. De fout die bij deze substitutie gemaakt wordt is verwaarloosd (vgl. § 7.4).

We concluderen dat een schatting \hat{x}_{p_0} voor het gevraagde kwantiel x_{p_0} (vgl. (5.1)) bepaald kan worden m.b.v. formule (5.55), waarbij de vormparameter γ , de locatieparameter μ en de schaalparameter σ natuurlijk geschat moeten worden. M.b.v. de methode van de grootste aannemelijkheid kunnen schatters $\hat{\gamma}_s, \hat{\mu}_s$ en $\hat{\sigma}_s$ voor de parameters γ, μ en σ , gebaseerd op de s stormseizoenmaxima, bepaald worden. Ook vindt men m.b.v. een procedure, analoog aan die beschreven in (5.44) - (5.47), een schatting voor de variantie van x_{p_0} . Op basis hiervan kan een benaderend betrouwbaarheidsinterval van de vorm (5.45) worden bepaald.

5.4 Welke statistische methode verdient de voorkeur?

We besluiten dit hoofdstuk met een korte afweging van de pro's en contra's van de beschreven methoden voor ons onderzoek. De verdelingsvrije methoden (1) en (2) hebben als belangrijk voordeel dat geen specifiek parametrisch model voor de geselecteerde HW's wordt gepostuleerd. Voor ons onderzoek is van deze twee methode (2) te verkiezen omdat veronderstelling (5.23) voor ons onderzoek realistischer is dan aanname (5.21). De methoden (3), (4) en (5) zijn parametrisch: ze zijn alle gebaseerd op de veronderstelling van een exact parametrisch model; bij methode (3) voor de HW's, bij methode (4) voor de hoogwateropzetten, en bij methode (5) voor de stormjaarmaxima. Dit is een belangrijk nadeel van deze methoden: een parametrisch model is in de regel alleen bij benadering correct.

Bovendien moeten we bij de methoden 3 en 4 veronderstellen dat $\gamma > -\frac{1}{2}$. Een voordeel van methode (4) is echter het expliciete gebruik van een belangrijk a priori gegeven: de verdeling van de astronomische hoogwaterstanden is in principe volledig bekend. Het aantal te gebruiken waarnemingen blijkt bij methode (4) aanzienlijk groter te zijn dan bij de andere methoden (zie hoofdstuk 6), zodat kleinere betrouwbaarheidsintervallen voor het gewenste kwantiel mogen worden verwacht, dan bij toepassing van de andere methoden. Methode 5 is gebaseerd op een analyse van de stormseizoenmaxima. Hoewel deze methode in zekere zin aantrekkelijk is vanwege haar eenvoud (geen drempelkeuze), wordt slechts een deel van de beschikbare informatie (alle waargenomen geselecteerde HW's) gebruikt. Samenvattend: Methode (2) lijkt het meest recht te doen aan de bijzondere kenmerken van het gestelde probleem.

6. Berekeningsresultaten

6.1 Inleiding

Na de beschrijving van de toegepaste methoden volgen in dit hoofdstuk de resultaten van de toepassingen van deze methoden op de ter beschikking staande tijdreeksen. Hierbij wordt dezelfde volgorde aangehouden als in hoofdstuk 5. Bij elke methode wordt in het kort haar essentie weergegeven. Dit heeft tot doel hoofdstuk 6 tot een zelfstandig leesbaar hoofdstuk te maken. Voor een goed begrip wordt daartoe hieronder enige algemene theorie kort en in zo eenvoudig mogelijke formulering gegeven.

Diegene die zich verder in de materie wil verdiepen wordt verwezen naar hoofdstuk 5 en de daar aangehaalde literatuur.

Het gaat er bij dit onderzoek om een zodanig hoogwaterniveau x_0 aan te geven dat de kans op bereiken of overschrijden van dat niveau in een willekeurig gegeven jaar gelijk is aan een van tevoren vastgesteld getal p_0 . Het gaat hierbij vooral om waarden voor p_0 tussen 10^{-3} en 10^{-5} .

Hiervoor kan beschikt worden over N onafhankelijke en gelijk verdeelde (d.i. tot dezelfde verdeling behorende) waarnemingen X_1, \dots, X_N : de geselecteerde en op relatieve zeeniveauveranderingen gecorrigeerde, in het verleden gedurende s jaren waargenomen hoogwaterstanden aan een gegeven station. Verondersteld wordt dat de kans op overschrijden van buitengewoon hoge stormvloedstanden buiten het lange stormseizoen verwaarloosbaar klein is t.o.v. deze kans binnen dit stormseizoen (zie ook § 7.4). In verband hiermee wordt in het vervolg gesproken over (het lange) „stormseizoen” i.p.v. „jaar”.

De waarnemingen X_1, \dots, X_N zijn alle afkomstig uit een in principe onbekende verdeling met verdelingsfunctie F .

Als het gemiddelde aantal waarnemingen per stormseizoen q wordt genoemd ($q = N/s$), dan geldt voor het stormseizoenmaximum M_a van een willekeurig stormseizoen

$$P(M_a \leq x_0) = \{F(x_0)\}^q = 1 - p_0 \quad (6.1)$$

Hieruit volgt

$$F(x_0) = (1 - p_0)^{\frac{1}{q}} \quad (6.2)$$

Voor de bij x_0 behorende overschrijdingskans p uit F geldt $F(x_0) = 1 - p$, zodat

$$p = 1 - (1 - p_0)^{\frac{1}{q}} \quad (6.3)$$

Voor kleine waarden van p_0 volgt uit de vergelijkingen (6.2) en (6.3)

$$p = \frac{p_0}{q}$$

en

$$F(x_0) = 1 - \frac{p_0}{q} \quad (6.4)$$

Zouden we F kennen, dan kan het gezochte peil x_0 worden berekend uit

$$x_0 = F^{-1} \left(1 - \frac{P_0}{q} \right) \quad (6.5)$$

F^{-1} is hierin de inverse functie van F . Omdat F niet bekend is, en niet direct geschat kan worden uit de waarnemingen voor de kleine overschrijdingskansen waarop het onderzoek is gericht, zal over F één of andere veronderstelling moeten worden gedaan. Het volgende uit de extreme-waardentheorie is daarbij van belang.

Verdelingsvrije methoden

Als X_1, X_2, \dots een rij onafhankelijke en gelijk verdeelde stochastische grootheden is met verdelingsfunctie F , dan geldt voor de kansverdeling van het maximum M_n van de eerste n grootheden X_1, \dots, X_n

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = F^n(x) \quad (6.6)$$

waarin x een willekeurige waarde is.

Hieruit volgt dat voor iedere x waarvoor $F(x) < 1$ geldt:

$$\lim_{n \rightarrow \infty} P(M_n \leq x) = 0.$$

Men zegt dan dat deze verdeling in de limiet „ontaardt“. Door het invoeren van geschikt gekozen constanten $a_n > 0$ en b_n convergeert bovenstaande uitdrukking echter voor de meeste bekende verdelingen F naar één van de zgn. extreme-waardenverdelingen G_γ ; voor grote waarden van n is dan

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \approx G_\gamma(x) \quad (6.7)$$

waarbij

$$G_\gamma(x) = \exp\left\{-\left(1 + \gamma x\right)^{\frac{-1}{\gamma}}\right\} \quad (6.8)$$

voor alle x waarvoor $1 + \gamma x > 0$.

De parameter γ noemt men de extreme-waardenindex. Voor $\gamma = 0$ ontstaat de zgn. Gumbel-verdeling: $G_0(x) = \exp(-e^{-x})$.

Men zegt van deze verdelingen F dat ze behoren tot het attractiedomein van G_γ .

Voor de toepassing van (6.7) voor onderhavige problematiek speelt het quotiënt n/k een belangrijke rol. Daarin is k een geschikt gekozen aantal grootste waarnemingen van de reeks van n waarnemingen. Er moet gelden dat $n/k \rightarrow \infty$ als $n \rightarrow \infty$, m.a.w. bij toenemende n neemt k langzamer toe dan n . Voor voldoende grote n/k geldt dan

$$F^{\frac{n}{k}}\left(a_{\frac{n}{k}} x + b_{\frac{n}{k}}\right) \approx G_\gamma(x) \quad (6.9)$$

en dus

$$P = \frac{P_0}{q} = 1 - F(x_0) \approx 1 - \left\{G_\gamma\left(-\frac{x_0 - b_{\frac{n}{k}}}{a_{\frac{n}{k}}}\right)\right\}^{\frac{1}{(n/k)}} \quad (6.10)$$

Substitutie van (6.8) in (6.10) en toepassing van de benadering $1 - e^{-x} \approx x$ voor kleine x levert

$$p \approx \frac{1}{\frac{n}{k}} \left(1 + \gamma \frac{x_0 - b_{\frac{n}{k}}}{a_{\frac{n}{k}}} \right)^{-\frac{1}{\gamma}} \quad (6.11)$$

zodat

$$x_0 \approx \frac{\left(\frac{k}{np}\right)^{\gamma} - 1}{\gamma} a_{\frac{n}{k}} + b_{\frac{n}{k}} \quad (6.12)$$

Deze benadering vormt de basis voor de in dit rapport genoemde verdelingsvrije methoden. Een schatting van x_0 wordt verkregen door het rechterlid van (6.12), waarbij de parameters γ , $a_{n/k}$ en $b_{n/k}$ uit de waarnemingen geschat moeten worden.

GPV-methode

Bij een andere benadering worden van de reeks onafhankelijke en gelijkverdeelde waarnemingen X_1, \dots, X_N met verdelingsfunctie F die waarnemingen geselecteerd, die boven een goed gekozen drempel u uitsteken, en wordt de verdeling F_u van de boven deze drempel uitstekende delen van de waarnemingen Y_1, \dots, Y_M beschouwd. Aangetoond kan worden, dat voor voldoende hoge drempel u de verdeling F_u goed benaderd wordt door de zgn. GPV-verdeling, de *gegeneraliseerde Pareto-verdeling*, hier aangeduid als $G(y; \sigma(u), \gamma)$, met vormparameter γ en schaalparameter $\sigma(u)$. Voorwaarde daarbij is dat F behoort tot het attractiedomein van de extreme-waardenverdelingen G_γ . Dus voor grote waarden van u geldt

$$F_u(y) \approx G(y; \sigma(u), \gamma) \quad \text{met} \quad G(y; \sigma(u), \gamma) = 1 - \left(1 + \gamma \frac{y}{\sigma} \right)^{-\frac{1}{\gamma}} \quad (6.13)$$

voor alle y waarvoor $1 + \gamma(y/\sigma) > 0$. De vormparameter γ is daarbij dezelfde als de extreme-waardenindex van de extreme-waardenverdelingen G_γ ; de schaalparameter $\sigma(u)$ is afhankelijk van de drempelwaarde u .

Voor $\gamma = 0$ ontstaat de exponentiële verdeling: $G(y; \sigma, 0) = 1 - \exp(-y/\sigma)$. Als nu de werkwijze gevolgd bij de afleiding van (6.5) wordt toegepast en $y_0 = x_0 - u$ en $q_u = M/s$ (= gemiddeld aantal uitstekende waarden boven drempel u per stormselzoen) worden gedefiniëerd, dan volgt

$$1 - \frac{P_0}{Q_0} = F(x_0) \approx G(y_0; \sigma_u, \gamma) \quad (6.14)$$

met

$$G(y_0; \sigma_u, \gamma) = 1 - \left(1 + \gamma \frac{y_0}{\sigma_u} \right)^{-\frac{1}{\gamma}}$$

En dus

$$x_0 = y_0 + u \approx -\frac{\sigma_u}{\gamma} \left(1 - \left(\frac{P_0}{Q_0} \right)^{-\gamma} \right) + u \quad (6.15)$$

Een schatting van x_0 wordt verkregen door de parameters σ en γ te schatten uit de waarnemingen bij geschikt gekozen drempel u .

GEV-methode

Er wordt weer uitgegaan van de reeks van N onafhankelijke en gelijk verdeelde waarnemingen X_1, \dots, X_N van een gegeven station. Bij toepassing van de GEV-methode (gegeneraliseerde extreme-waardenverdeling) wordt de reeks verdeeld in een aantal perioden van gelijke lengte en wordt per periode de grootste waarde geselecteerd. Gebruikelijk bij studies als deze is perioden met de duur van een jaar te nemen, waarbij de periode ten opzichte van het kalenderjaar zo gekozen wordt dat onafhankelijkheid van de maxima gegarandeerd is. Hier worden stormseizoenmaxima beschouwd.

Voor het maximum $X_{\max,i}$ van elk stormseizoen i ($i = 1, \dots, s$) geldt formule (6.7). Aangenomen wordt nu dat de verdeling van de $X_{\max,i}$'s, mits juist genormeerd, exact gelijk is aan de extreme-waardenverdeling G_γ . Worden de hiertoe in te voeren constanten μ (plaatsparameter) en σ (schaalparameter) genoemd, dan geldt dus

$$P\left(\frac{X_{\max,i} - \mu}{\sigma} \leq x\right) = G_\gamma(x) \quad (6.16)$$

waarbij $G_\gamma(x)$ gegeven wordt door (6.8).

Deze formule kan worden omgewerkt tot de gedaante

$$P(X_{\max,i} \leq x) = G_\gamma\left(\frac{x - \mu}{\sigma}\right) \quad (6.17)$$

Wegens formule (6.1) wordt dus een schatting van x_0 , het p_0 -kwantiel, verkregen door het oplossen van de vergelijking

$$1 - p_0 = G_\gamma\left(\frac{x_0 - \mu}{\sigma}\right) \quad (6.18)$$

Na enige herleiding wordt de volgende formule gevonden

$$x_0 = \frac{\sigma}{\gamma} [-1 + \{-\ln(1 - p_0)\}^{-\gamma}] + \mu \quad (6.19)$$

Een schatting van x_0 wordt verkregen door de parameters γ , σ en μ te schatten uit de waarnemingen en in te vullen in (6.19).

Voor de afleiding van de formules waarmee de parameters uit de waarnemingen kunnen worden geschat, alsmede die voor de bijbehorende betrouwbaarheidsintervallen, wordt verwezen naar hoofdstuk 5.

6.2 Gebruik van de empirische frequentieverdeling

In § 6.1 is al aangegeven dat het aantal te gebruiken grootste waarnemingen k , dan wel de drempelwaarde u , geschikt gekozen moet worden, met name voor de verdelingsvrije methoden, de GPV-methode en de nog te bespreken convolutiemethode (CON, § 6.6). Dit levert een spanningsveld tussen enerzijds zoveel mogelijk waarnemingen om zo betrouwbaar mogelijke resultaten te krijgen en anderzijds niet teveel waarnemingen om nog wel iets over extreme waarden te kunnen zeggen. Het gebruik van teveel grootste waarnemingen leidt n.l. tot een systematische afwijking (bias) in de schattingen van extreme

kwantielen (het α^e -kwantiel van een verdeling $F(x)$ is hier gedefinieerd als die waarde van x waarvoor geldt $1-F(x) = \alpha$); er wordt dan ook niet meer voldaan aan de gestelde voorwaarden voor n/k bij de afleiding van de verdelingsvrije methoden.

Om een geschikte keuze voor k te kunnen maken zijn berekeningen uitgevoerd voor olopende reeksen van k . Voor de GPV- en de CON-methode is een reeks van drempelwaarden toegepast. Die zijn echter eenvoudig te herleiden tot reeksen van k , omdat bij elke drempelwaarde een aantal overschrijdingen hoort, af te leiden uit de waarnemingen.

In eerste instantie is nu nagegaan of op grond van de empirische frequentieverdelingen van HW's c.q. HW-opzetten iets gezegd kan worden over het maximale aantal te gebruiken grootste waarnemingen. In de bijlagen 17 zijn voor de beschouwde stations de geselecteerde en gecorrigeerde HW's (verder kortweg aangeduid als HW's) in cm boven NAP uitgezet tegen de logaritme van het gemiddelde aantal keren per stormseizoen dat de betreffende peilen x werden bereikt of overschreden. Dit gemiddelde aantal is berekend als het product van het gemiddelde aantal waarnemingen per lang stormseizoen $q (= N/s)$ en de geschatte overschrijdingskans $1-F(x) \approx i/N$, met i het rangnummer van de naar hoogte gesorteerde waarnemingen; $i = 1$ is hierbij het rangnummer van de hoogste waarneming. Opmerking: de algemene formule voor het schatten van overschrijdingskansen direct uit de waarnemingen is $1-F(x) \approx (i-c_1)/(N+c_2)$, waarbij c_1 en c_2 constanten zijn met een waarde tussen 0 en 1. Over de toe te passen waarden voor deze constanten is veel gepubliceerd. De aanbevelingen zijn in het algemeen afhankelijk van het type verdeling en zijn niet altijd eensluidend. Veel gebruikte waarden zijn: $c_1 = 0,3$ en $c_2 = 0,4$. Aangezien bij dit onderzoek de getalswaarden alleen gebruikt worden voor de visuele presentatie (de zgn. „plot“-positie) en geen invloed hebben op de gevraagde kwantielschattingen is eenvoudigheidshalve gekozen voor $c_1 = c_2 = 0$.

Het verloop van de aldus uitgezette punten is gelijkvormig aan dat van de geschatte $1-F(x)$ zelf, behoudens een constante factor $^{10}\log(q)$. Bij de hoge frequenties (vanaf ca. 5 maal per stormseizoen) vertoont dit verband een vrij sterke kromming. De verdeling van de HW's wijkt in dit gebied kennelijk af van die voor de hoge HW's. Bij deze relatief lage HW's is geen sprake meer van willekeurige toedeling van HW-opzetten aan astronomische hoogwaters; naarmate de HW-stand lager is (lager dan de hoogst mogelijke astronomische standen plus de toegepaste drempel bij de selectie van HW-opzet: 30 cm) zijn per HW minder astronomische standen „beschikbaar“. Voor het bepalen van de overschrijdingskansen van zeer hoge waterstanden mag dit deel van de waarnemingen in ieder geval niet meegenomen worden. Een praktische maat voor de hoogst mogelijke astronomische stand is het hoogste voorkomende verschil tussen HW en HW-opzet in de tijdreeksen. Een essentiële veronderstelling is verder dat voor de zeer lage frequenties, buiten het waarnemingsgebied, zich niet een afwijkend gedrag voordoet.

Toepassing van bovenstaande redenering leidt tot een minimale drempel u_{min} zoals aangegeven in tabel 6.1 (naar beneden afgerond tot een veelvoud van 5 cm) en het bijbehorende maximale aantal te gebruiken grootste beschikbare waarden van HW, k_{max} (afgerond op een veelvoud van 10).

Deze waarden komen redelijk overeen met de positie waar de frequentieverdelingen beginnen af te wijken (zie bijlage 17). Aldus is hiervoor een objectieve maat gegeven.

Tabel 6.1

Enige algemene gegevens van de reeksen van HW

station	N	s	$q = \frac{N}{s}$	k_{max}	U_{min} (NAP + cm)
Vlissingen	1672	104	16,1	270	300
Hoek van Holland	1577	98	16,1	530	185
Den Helder	866	53	16,3	430	120
Harlingen	936	53	17,7	520	160
Delfzijl	1767	104	17,0	860	205

Opmerkelijk is de relatief lage k_{max} voor Vlissingen. Dit wordt veroorzaakt door het brede bereik van mogelijke astronomische hoogwaterstanden (bijlage 14). Relatief veel geselecteerde HW's blijven daardoor beneden het niveau van de hoogst mogelijke astronomische standen en vallen daardoor af.

Bij de convolutie-methode wordt gebruik gemaakt van de verdeling van de geselecteerde HW-opzetten. De empirische verdelingsfuncties van de HW-opzetten zijn weergegeven in bijlage 18. Bij de selectie van de HW-opzetten is een minimumdrempel van 30 cm toegepast. Uit deze bijlage blijkt dat er voor de geselecteerde HW-opzetten geen niveau is aan te wijzen waarbeneden de verdelingsfunctie zich duidelijk anders gedraagt. De grafieken op zich geven dus geen aanleiding een hogere drempel te kiezen en daarmee het aantal te gebruiken hoge HW-opzetten te beperken.

6.3 Verdelingsvrije methode VVM-c

Bij de verdelingsvrije methoden wordt van de feitelijke, maar onbekende, verdeling F van de waarnemingen aangenomen dat ze behoort tot het attractiedomein van de extreme-waardenverdelingen G_γ . Hoe F er zelf uitziet blijft dus buiten beschouwing. De VVM-c-methode gaat uit van (6.12) - zie ook (5.14). De (momenten)schaters voor de parameters in (6.12) zijn gegeven met (5.15) en (5.18). Voor de afleiding van het betrouwbaarheidsinterval voor deze methode gelden additionele voorwaarden (zie 5.21). Een essentiële veronderstelling daarbij is dat $pn \rightarrow c$ voor grote n ($0 < c < \infty$). Aan deze positieve c ontleent deze methode haar aanduiding. Bij de VVM-c methode wordt eerst γ geschat alvorens de kwantielen te schatten. Het aantal te gebruiken grootste waarnemingen k_γ (in hoofdstuk 5 met m aangeduid) dient daarbij veel groter te zijn dan dat voor de schatting van de kwantielen zelf (k_p).

De bijlagen 19A t/m 19E geven γ als functie van k . Op grond van k_{max} en het verloop van γ in deze figuren is gekozen voor de waarden van γ volgens tabel 6.2. Deze gekozen waarden van γ en het daarbij beschouwde k -interval zijn aangegeven in de bijlagen. Deze keuze vormt een subjectief element in de procedure. Gekeken is naar het verloop tot k_{max} ; daarbij is gezocht naar een k -interval waarvoor γ min of meer constant is. De schommelingen in het verloop voor kleine k blijven daarbij buiten beschouwing (te weinig waarnemingen).

Het neerwaartse verloop van γ bij hogere waarden van k wijst op een toenemende systematische afwijking als gevolg van het meenemen van teveel waarnemingen. Een gemiddelde waarde over het aangegeven interval is aangehouden.

Tabel 6.2

Gekozen waarden van γ en beschouwde k -intervallen bij de methode VVM-c

station	γ	k_γ	k_k
Vlissingen	0,0	40 - 270	40 - 100
Hoek van Holland	0,0	40 - 350	40 - 250
Den Helder	-0,10	40 - 300	40 - 170
Harlingen	-0,15	40 - 520	40 - 225
Delfzijl	-0,05	40 - 860	40 - 325

Vervolgens zijn met de gekozen γ -waarden kwantielschattingen met de bijbehorende 95%-betrouwbaarheidsintervallen berekend als functie van k . De bijlagen 19A t/m 19E geven het verloop van de 10^{-4} -kwantielen en hun betrouwbaarheidsintervallen.

De keuze van een geschikte waarde voor k , dan wel voor de kwantielschatting, wordt gedaan op grond van de volgende overwegingen:

- $k > \sim 40$; bij zeer kleine k verlopen de γ - en de kwantielschattingen over het algemeen zeer instabiel;
- k_k aanmerkelijk kleiner dan k_γ ;
- in het geval er duidelijk sprake is van een traject waarover het geschatte 10^{-4} -kwantiel stabiel (dus onafhankelijk van k) is, wordt de schatting bij een zo groot mogelijke waarde van k gekozen, omdat daarmee het kleinste betrouwbaarheidsinterval wordt verkregen;
- valt een dergelijk verloop niet te onderscheiden, dan wordt op het oog over het beschouwde k -interval een gemiddelde waarde bepaald, zowel van de kwantielschatting zelf als van het betrouwbaarheidsinterval;

De zo bepaalde waarden van de kwantielschattingen en bijbehorende 95%-betrouwbaarheidsintervallen staan in tabel 6.7 (§ 6.8). Daarbij zijn afgeronde getallen genomen; grotere nauwkeurigheid is niet zinvol; dat geldt overigens ook voor de andere methoden.

In de bijlagen 19A t/m 19E is het verloop van de betrouwbaarheidsintervallen minder grillig dan dat van de kwantielschattingen zelf. Dat komt omdat de betrouwbaarheidsintervallen zijn ingetekend voor k -waarden die een veelvoud van 20 zijn; de verkregen punten zijn door rechte lijnen verbonden. De kwantielschattingen zelf zijn getekend voor opeenvolgende waarden van k .

Zoals uit het vervolg zal blijken zijn de betrouwbaarheidsintervallen relatief klein. Dit komt omdat de variabiliteit van de γ bij deze methode niet in het betrouwbaarheidsinterval tot uitdrukking komt: er is een vaste waarde voor γ ingevoerd. Het betrouwbaarheidsinterval is daardoor niet erg realistisch.

Ter illustratie zijn in bijlage 19F overschrijdingslijnen voor de stations Hoek van Holland en Den Helder gegeven. Ze zijn geconstrueerd m.b.v. de kwantielschattingen en hun betrouwbaarheidsinterval voor de in de figuren aangegeven waarden van p .

De daarbij toegepaste k -waarden vallen buiten het in tabel 6.2 gegeven interval van k_i ; het zijn min of meer arbitrair gekozen waarden uit een eerdere fase van het onderzoek, die toch een goede indruk geven van het eigenlijke verloop der lijnen. Gezien de ermee gemoeide inspanning en de in § 7.2 gegeven overwegingen t.a.v. de methode VVM-c, is er van afgezien voor deze methode nieuwe overschrijdingslijnen te berekenen.

6.4 Verdelingsvrije methode VVM-0

Evenals de VVM-c-methode gaat de VVM-0-methode uit van (5.12) voor de kwantielschatting. De methoden verschillen met name van elkaar in de afleiding van het betrouwbaarheidsinterval. Bij de VVM-0-methode wordt het aantal te gebruiken grootste waarnemingen voor de γ - en de kwantielschatting gelijk genomen. Beide schattingen kunnen dus bij deze methode in één rekenslag als functie van k berekend worden. Daardoor werkt de variabiliteit van γ direct door in de kwantielschatting; dit is een voordeel van deze methode. Verder moet gelden dat $pn \rightarrow 0$ voor grote n ; met name hierin verschilt deze methode van de methode VVM-c. Voor de parameterschattingen wordt weer gebruik gemaakt van de formules (5.15) en (5.18).

Schattingen voor γ en het 10^{-4} -kwantiel als functie van k voor de vijf stations zijn weergegeven in de bijlagen 20A t/m 20E. Op basis van de volgende overwegingen is een keuze tot stand gekomen voor de kwantielschattingen:

- $k < k_{maxi}$
- $k > \sim 40$ (zie ook § 6.3.)
- het te beschouwen k -interval wordt mede bepaald door het gebied waarin de γ -waarde zich beweegt rond een redelijk stabiele waarde (bijv. de gekozen γ -schatting bij de methode VVM-c). Het gebied waar de γ -schattingen significant afwijken van deze waarde wordt buiten beschouwing gelaten;
- in het geval er duidelijk sprake is van een traject waarover het geschatte 10^{-4} -kwantiel stabiel (dus onafhankelijk van k) is, wordt de schatting bij een zo groot mogelijke waarde van k gekozen, omdat daarmee het kleinste betrouwbaarheidsinterval wordt verkregen;
- valt een dergelijk verloop niet te onderscheiden, dan wordt op het oog een gemiddelde waarde bepaald over het beschouwde k -interval, zowel van de kwantielschatting als van het betrouwbaarheidsinterval.

De resultaten van bovenstaande overwegingen staan in tabel 6.7 (§ 6.8.) en zijn aangegeven in de bijlagen 20A t/m 20E.

De bijlagen 20F, G en H geven de bijbehorende overschrijdingslijnen. De beschouwde k -intervallen, alsmede de grootste waarden van k waarvoor het gekozen 10^{-4} -kwantiel en het bijbehorende betrouwbaarheidsinterval kunnen worden berekend (k_{rek}), staan in tabel 6.3. Deze k -waarden zijn toegepast bij de berekening van de overige in de bijlagen 20F, G en H aangegeven kwantielschattingen teneinde de overschrijdingslijnen te kunnen construeren. Het verloop van deze kwantielschattingen als functie van k is gelijkvormig aan dat van het 10^{-4} -kwantiel.

In de tabel zijn tevens de schattingen van de parameters van het VVM-0-model opgenomen, zoals die berekend zijn voor k_{rek} .

Tabel 6.3

Beschouwde k -intervallen, k_{rek} en de bijbehorende schattingen van de parameters γ , $a_{n/k}$ en $b_{n/k}$ bij de methode VVM-0

station	k -interval	k_{rek}	γ	$a_{n/k}(cm)$	$b_{n/k}(cm)$
Vlissingen	40 - 270	268	+0,0131	22,28	300
Hoek van Holland	40 - 390	210	+0,0374	23,45	216
Den Helder	40 - 270	269	-0,0665	29,60	142
Harlingen	40 - 520	500	-0,1593	46,53	161
Delfzijl	40 - 470	458	-0,0360	41,05	239

6.5 GPV-model

Bij het GPV-model wordt aangenomen dat de waarnemingen die boven een geschikt gekozen drempel u uitsteken minus die drempel exact een GPV-verdeling bezitten. Dit is bij *benadering* het geval wanneer F behoort tot het attractiedomein van de GEV-verdeling. De parameters σ en γ worden geschat m.b.v. de methode van de grootste aannemelijkheid, waarbij de benadering wordt vervangen door exacte gelijkheid. De kwantielschatting wordt verkregen m.b.v. (6.15) of (5.42).

De bijlagen 21A t/m 21E geven de 10^{-4} -kwantielschattingen volgens het GPV-model als functie van k . In de berekeningen werd gewerkt met oplopende drempelwaarden; in de bijlagen zijn i.p.v. de toegepaste drempels de bijbehorende k -waarden uitgezet teneinde een overeenkomstige wijze van presentatie voor de verschillende methoden te bieden.

De keuze voor een geschikte drempelwaarde is gedaan op grond van de volgende overwegingen:

- groter dan u_{min} ofwel $k < k_{maxi}$
- niet te weinig waarnemingen. Bij zeer hoge drempels fluctueren de kwantielschattingen sterk als functie van de drempel. Grote pieken in het verloop worden veroorzaakt door tekenomslag van de γ -waarden;
- stabiele kwantielschatting als functie van de drempel.

De op grond van deze overwegingen gekozen drempelwaarden, alsmede de bijbehorende k -waarden en schattingen van de parameters staan in tabel 6.4. De daarmee berekende 10^{-4} -kwantielen met hun 95%-betrouwbaarheidsinterval zijn opgenomen in de tabel 6.7 (§ 6.8).

Tabel 6.4

Gekozen drempelwaarden u voor HW en de bijbehorende k -waarden en schattingen van de parameters γ en σ volgens het GPV-model

station	u (NAP+cm)	k	γ	$\sigma(cm)$
Vlissingen	300	269	+0,0119	23,58
Hoek van Holland	210	249	-0,0102	27,71
Den Helder	170	127	-0,0472	30,56
Harlingen	200	254	-0,1726	49,04
Delfzijl	255	325	-0,0674	48,62

De bijlagen 21A t/m 21E geven tevens overschrijdingslijnen, die geconstrueerd zijn m.b.v. kwantielschattingen voor de verschillende in de grafieken aangegeven waarden van p .

Opgemerkt moet worden dat voor Vlissingen en Delfzijl de overschrijdingslijnen van de bijlagen 21A en 21E zijn berekend voor drempelwaarden afwijkend van die van tabel 6.4. Zij werden berekend in een eerdere fase van het onderzoek, voordat de definitieve drempelkeuze gedaan werd.

Gezien de ermee gemoeide inspanning en in het licht van de evaluatie van § 7.2 is er van afgezien nieuwe overschrijdingslijnen voor deze twee stations volgens het GPV-model te berekenen (vergelijk het slot van § 6.3).

6.6 Convolutiemethode (CON)

De convolutie-methode is de enige overgebleven methode in dit onderzoek waarbij de geselecteerde en gecorrigeerde HW-standen beschouwd zijn als som van de HW-opzet en de bijbehorende astronomische hoogwaterstand, die als stochastisch onafhankelijk mogen worden verondersteld (§ 4.4.). Voor de HW-opzet wordt het GPV-model aangenomen. Verder worden de berekende kansverdelingen van de astronomische standen van de verschillende stations (§ 4.6.) als exact juist verondersteld. Koppeling tot een kansuitspraak over de HW-stand geschiedt d.m.v. de zgn. convolutie-integraal (5.52).

In de bijlagen 22A t/m 22E zijn weer de 10^{-4} -kwantielschattingen voor HW als functie van k weergegeven. Ze werden berekend voor oplopende drempelwaarden voor de HW-opzet. Uit een oogpunt van uniformiteit is echter de voorkeur gegeven aan een presentatie als functie van k , die immers gekoppeld is aan de drempelhoogte van HW-opzet via de empirische verdelingsfunctie van de HW-opzet van het betreffende station.

Op grond van de volgende overwegingen is een keuze gemaakt voor k :

- aan de empirische verdelingsfunctie van HW-opzet ≥ 30 cm is geen argument te ontleen voor een hogere drempel van HW-opzet, dus ook niet voor een duidelijke bovengrens voor k ;
- van het traject waarover het geschatte 10^{-4} -kwantiel stabiel (dus onafhankelijk van drempel of k) is, wordt de schatting bij een zo groot mogelijke waarde k gekozen, omdat daarmee het kleinste betrouwbaarheidsinterval wordt verkregen;
- trajecten met sterke fluctuaties (bij relatief lage k -waarden) blijven buiten beschouwing.

De gekozen drempelwaarden voor HW-opzet, alsmede de bijbehorende k -waarden en schattingen van γ en σ staan in tabel 6.5. De berekende 10^{-4} -kwantielen en de bijbehorende 95%-betrouwbaarheidsintervallen staan in tabel 6.7 (§ 6.8).

Voor de berekening van de betrouwbaarheidsintervallen was bij deze methode geen theoretisch model voorhanden. Ze zijn daarom bepaald volgens de zogenaamde parametrische-bootstraphmethode, waarbij gebruik gemaakt wordt van simulaties volgens een Monte-Carlo-procedure - een inleiding over deze methode treft men aan in [Helmers 1984]. De toegepaste procedure is als volgt:

- bij de gekozen waarden voor k corresponderen schattingen voor de parameters σ en γ van de GPV-verdeling (zie tabel 6.5). De aldus verkregen verdeling („moeder"-verdeling) wordt als uitgangspunt genomen;

- volgens een Monte-Carlo-procedure worden k onafhankelijke trekkingen genomen uit deze verdeling. Met deze steekproef, even groot dus als de oorspronkelijke, worden opnieuw σ en γ geschat en de gewenste kwantielen berekend;
- deze simulatie wordt 1000 maal uitgevoerd. Het resultaat is 1000 schattingen van de verschillende kwantielen. Hiermee kunnen de varianties van de kwantielschattingen worden berekend en het 95%-betrouwbaarheidsinterval.

De met deze resultaten verkregen overschrijdingslijnen met bijbehorende, enigzins asymmetrische, 95%-betrouwbaarheidsintervallen zijn weergegeven in de bijlagen 22A t/m 22E. In tabel 6.7 (§ 6.8) zijn de 10^{-4} -kwantielschattingen gegeven. De in deze tabel gegeven waarden van de 95%-betrouwbaarheidsintervallen zijn, ten behoeve van de onderlinge vergelijking van de uitkomsten van de verschillende methoden, echter berekend als $\pm 1,96$ maal de standaardafwijking van de duizend 10^{-4} -kwantielschattingen.

Tabel 6.5

Gekozen drempelwaarden u voor HW-opzet en de bijbehorende k waarden schattingen van de parameters σ en γ bij de CON-methode

station	$u(cm)$	k	γ	$\sigma(cm)$
Vlissingen	55	775	-0,0324	27,03
Hoek van Holland	65	662	-0,0441	31,40
Den Helder	50	569	-0,1112	41,83
Harlingen	75	665	-0,1984	55,99
Delfzijl	65	940	-0,0419	50,23

6.7 Het GEV-s-model

De resultaten van de berekeningen volgens het GEV-model zijn weergegeven in de bijlagen 23A t/m 23C. Het voordeel van de methode is dat de parameter k er geen rol in speelt, hetgeen de uitkomsten minder subjectief maakt. Omdat de methode uitgaat van gelijkverdeelde waarnemingen zijn hier stormseizoenmaxima gebruikt in plaats van jaarmaxima; om die reden is het model aangeduid als GEV-s. In § 7.4 wordt teruggekomen op het gebruik van jaarmaxima (GEV-j).

De uitkomsten voor de 10^{-4} -kwantielen en hun 95%-betrouwbaarheidsinterval zijn weergegeven in tabel 6.7 (§ 6.8). De parameters μ , γ en σ zijn geschat m.b.v. de methode van de grootste aannemelijkheid. Onderstaande tabel geeft de schattingen van deze parameters.

Tabel 6.6

Schattingen van de parameters van het GEV-s-model

station	$\mu(cm)$	γ	$\sigma(cm)$
Vlissingen	322,11	+0,0394	23,67
Hoek van Holland	234,83	+0,0215	27,27
Den Helder	195,54	-0,0959	32,02
Harlingen	265,13	-0,2451	42,03
Delfzijl	303,06	-0,0378	44,71

6.8 Samenvattende tabel van de resultaten

In onderstaande tabel zijn voor de vijf basisstations de resultaten van de vijf toegepaste methoden weergegeven voor de 10^{-4} -kwantilen en hun 95% betrouwbaarheidsinterval. Ter vergelijking zijn daaraan toegevoegd de basispeilen volgens het „Deltarapport”, de hoogste HW-standen opgetreden in de beschouwde perioden en de herleiding van deze laatste standen naar de toestand van 1985. Het zij hier reeds uitdrukkelijk gesteld dat een methode niet noodzakelijkerwijze betere resultaten oplevert naarmate de bijbehorende betrouwbaarheidsintervallen kleiner zijn. In hoofdstuk 7 worden deze uitkomsten nader geëvalueerd en voor zover nodig geïnterpreteerd.

Tabel 6.7		methode	Vlissingen	Hoek v. H.	Den Helder	Harlingen	Delfzijl
Schattingen van het 10^{-4} -kwantiel (cm boven NAP) en de bijbehorende 95%-betrouwbaarheidsintervallen (cm) voor de toestand 1985		VVM-c	540 ± 50	460 ± 40	360 ± 30	430 ± 30	600 ± 40
		VVM-0	540 ± 140	500 ± 190	370 ± 100	405 ± 80	600 ± 150
		GPV	555 ± 140	475 ± 125	415 ± 155	440 ± 75	615 ± 150
		CON	540 ± 65	460 ± 80	390 ± 70	445 ± 45	660 ± 120
		GEV-s	585 ± 185	510 ± 180	390 ± 110	420 ± 65	650 ± 195
		bp	565	500	505	580	640
		hs-herleid	466	402	331	379	480
		hs-gemeten	455 (1953)	385 (1953)	325 (1953)	369 (1954;1976)	453 (1901)

- bp = basispeil uit het „Deltarapport”
- hs-herleid = hoogste stand van het gebruikte, naar de toestand 1985 herleide, HW-bestand
- hs-gemeten = hoogste gemeten stand uit de beschouwde periode met het jaar van optreden (de hoogst bekende stand van Delfzijl, opgetreden in 1825, is NAP + 460 cm; herleid naar 1985 is dit - zie § 8.1 - : NAP + ~500 cm)

7. Evaluatie van de berekeningsresultaten

Na de presentatie van de schattingen volgens een vijftal statistische modellen, volgt logischerwijze een evaluatie daarvan, die moet leiden tot een eindresultaat. De mate waarin de verschillende modellen voldoen bij toepassing op de beschikbare waarnemingsreeksen en de vragen die daarbij zijn gerezen worden in dit hoofdstuk besproken. Eerst zal de aandacht worden gericht op de vorm van de empirische overschrijdingslijnen. Dan volgt een bespreking van de resultaten van hoofdstuk 6, gevolgd door een beschouwing over de betrouwbaarheidsintervallen. Vervolgens wordt stilgestaan bij kwantielschattingen van bijzondere selecties uit de beschikbare waarnemingsreeksen, wordt de gevoeligheid voor de breedte van de D-S-i-selectie en voor de hoogte van de ondergrens van HW-opzet onderzocht, worden twee in de loop van het onderzoek afgevallen methoden besproken en wordt een vergelijking gemaakt met het statistische model dat door de Deltacommissie werd toegepast voor Hoek van Holland.

Na het geven van een overzicht van de resulterende overschrijdingslijnen en betrouwbaarheidsintervallen eindigt dit hoofdstuk met een beschouwing over de uitkomsten in relatie tot de achterliggende fysische processen.

7.1 De empirische frequentieverdelingen

Alvorens de uitkomsten van de statistische modellen te bespreken, kunnen aan de empirische frequentieverdelingen van de geselecteerde en gecorrigeerde waarden (bijlagen 17 en 18) reeds enige opvallende kenmerken worden ontleend.

Het traject boven de u_{min} in deze bijlagen is voor de stations Vlissingen, Hoek van Holland en Delfzijl op het oog redelijk te benaderen door een rechte lijn. Er is echter geen fysische theorie die aannemelijk maakt dat het op voorhand een rechte (corresponderend met $\gamma = 0$) zou moeten zijn. Het is zelfs fysisch onmogelijk dat een rechtlijnige extrapolatie zich voortzet tot in het oneindige. Op grond van fysische beperkingen t.a.v. meteorologische omstandigheden en de te verplaatsen hoeveelheid water is het eerder te verwachten dat zich ergens een neerwaartse afbuiging manifesteert. Thans is echter niet aan te geven boven welk waterstandsniveau dat zichtbaar zou moeten worden of wat een fysisch mogelijke bovengrens zal zijn. Het is zelfs denkbaar dat een dergelijke afbuiging zich pas zal aftekenen na een voorafgaande opwaartse afbuiging; hierover valt met de huidige kennis niets te zeggen. In een zuiver statistisch onderzoek als het onderhavige kan niet anders gedaan worden dan uit te gaan van de beschikbare waarnemingen, en kan niet anders aangenomen worden dan dat de kansverdelingsfuncties zich in het extrapolatiegebied niet anders gedragen dan bij de gekozen k hoogste waarnemingen. Ook een geringe, op het oog niet duidelijk zichtbare, afwijking van een „rechte” lijn ($\gamma = 0$) heeft grote invloed op kwantielschattingen voor zeer lage overschrijdingskansen vanwege de verregaande extrapolatie vanuit de

beschikbare waarnemingsreeksen. De vraag kan gesteld worden of een dergelijke extrapolatie wel verantwoord is. De bedreiging van ons land door buitengewoon hoge stormvloed (aantal te verwachten doden, maatschappelijke ontwrichting en schade aan kapitaalgoederen bij inundatie) is echter van dien aard, dat het noodzakelijk is dergelijke kleine kansen te beschouwen.

Verder blijkt uit bijlage 17 dat de „kromme” voor Vlissingen boven u_{min} iets flauwer verloopt dan die voor Hoek van Holland, mede vanwege de invloed van het Kanaal („weglekken”), die voor Vlissingen groter is dan voor Hoek van Holland, en de ligging aan de brede Westerschelde. Bijlage 18 vertoont dit beeld nog duidelijker. De krommen van Delfzijl verlopen duidelijk veel steiler, hetgeen mede veroorzaakt wordt door de ligging van Delfzijl in het Eems-Dollard estuarium. Zowel het „HW-klimaat” als het „HW-opzet”-klimaat zijn kennelijk per station duidelijk verschillend. Dit ondersteunt de gekozen aanpak waarbij het statistisch onderzoek niet beperkt blijft tot één station. Er is geen reden aan te nemen dat voor lagere frequenties het verloop bij Delfzijl af zou moeten buigen tot evenwijdig aan dat van Hoek van Holland en Vlissingen of omgekeerd.

Bij Den Helder wijkt het verloop van de empirische frequentieverdeling van de HW-standen boven u_{min} al meer van een rechte af dan bij de voornoemde drie stations; bij Harlingen is dat nog meer het geval. Dit duidt op een negatieve waarde van γ . De enige mogelijke verklaring van dit afwijkende gedrag moet gelegen zijn in de invloed van de configuratie van de westelijke Waddenzee op het verloop van de waterstanden en het effect van de locale opwaaiing aldaar.

Opmerkelijk is verder dat de krommen voor Den Helder voor de hoge frequenties steiler zijn dan voor Hoek van Holland; voor de lage frequenties is dat andersom. Mogelijke effecten die hierbij een rol spelen zijn het „brede” verloop van de astronomische getijkrommen rond hoogwater bij Den Helder, waardoor de kans op samentreffen van maximaal stormeffect en een hoge astronomische stand groter is, en de geografische ligging van Den Helder, die maakt dat dit station zowel „meedoet” met stormen die specifiek in het Waddengebied waterstandsverhogingen veroorzaken, als met stormen die vooral langs de westelijke kust verhogingen veroorzaken.

7.2 Bespreking van de resultaten van de toegepaste statistische modellen

De verdelingvrije methoden

De methoden VVM-0 en VVM-c verschillen met name van elkaar in de eis voor de limietwaarde van pn voor grote n , te weten 0 resp. een positieve constante c .

In de praktijk vinden we bij eindige reeksen altijd een positieve waarde; de uitkomst van de grootte $pn \approx (p_0, q)/N = p_0 \cdot s$ bedraagt voor Vlissingen 0,01 voor $p_0 = 10^{-4}$. De vraag is dan welke benadering het beste bij deze situatie past. De grootte pn kan geïnterpreteerd worden als het verwachte aantal overschrijdingen van het p_0 -kwantiel (§ 5.2). Het is bekend dat het niveau dat correspondeert met bijv. het 10^{-4} -kwantiel in werkelijkheid niet is overschreden. Dit suggereert dat de benadering $c = 0$ hier het meest realistisch is.

Een bezwaar van de methode VVM-c is verder dat de variabiliteit van de γ -schatting niet in de betrouwbaarheidsintervallen van de

kwantielschattingen tot uitdrukking komt: er is immers een vaste waarde voor γ ingevoerd. De betrouwbaarheidsintervallen zijn daardoor relatief klein en niet erg realistisch.

Om deze twee redenen worden de uitkomsten van de methode VVM-c niet verder in beschouwing genomen.

Voor Vlissingen is een kwantielschatting waarbij $\gamma = 0$ een redelijke uitkomst gezien het verloop van de kwantielschattingen als functie van k (bijlage 20A).

Bij Hoek van Holland is dat minder duidelijk. Er is eigenlijk geen stabiel gebied aan te wijzen. In bijlage 17 vertoont de overschrijdingskromme voor Hoek van Holland, corresponderend met het verloop van de γ -schatting in bijlage 20B, boven u_{\min} een neerwaarts afbuigend verloop, gevolgd door een opwaartse afbuiging. In bijlage 20B is een „gemiddelde waarde” aangegeven over het traject tot waar de waarde voor γ significant gaat afwijken van 0. Deze gemiddelde uitkomst houdt dus rekening met beide tendenzen. Voor de overige stations zijn redelijk stabiele trajecten aanwezig.

Uit het bovenstaande blijkt dat de interpretatie van de uitkomsten van de methode VVM-0 een subjectief element bevat.

Het zaagtand-achtige karakter van de grafieken van de schattingen wordt veroorzaakt door groepen van gelijke waarnemingen.

GPV-methode

Bij deze methode wordt een GPV-verdeling opgelegd aan dat deel van de waarnemingen dat een bepaalde drempel overstijgt. De uitkomsten voor de 10^{-4} -kwantielen zijn daardoor gevoelig voor de mate waarin deze uitstekende delen van de waarnemingen (waarneming minus drempel) voldoen aan deze verdeling.

De berekeningen zijn uitgevoerd voor oplopende drempelwaarden met een interval van 5 cm, hetgeen veel minder gedetailleerd is dan voor alle oplopende waarden van k . Het zaagtand-achtige verloop bij de VVM-methoden is hiermee ook onderdrukt. De interpretatie van de bijlagen 21A t/m 21E tot een aan te bevelen kwantielschatting is bij deze methode voor de meeste stations behoorlijk subjectief. Alleen Harlingen en Delfzijl geven een redelijk ondubbelzinnige uitkomst.

CON-methode

De CON-methode heeft t.a.v het opleggen van een GPV-verdeling aan de boven een drempel uitstekende delen van de waargenomen HW-opzetten hetzelfde bezwaar als de GPV-methode. Voordelen zijn het grote aantal te gebruiken waarnemingen (HW-opzetten) en het gebruik van de bekende verdeling van de astronomische hoogwaterstanden. De methode weegt ook kansen mee van combinaties die nooit zijn opgetreden.

Opmerkelijk zijn de hoge pieken voor relatief kleine k bij Hoek van Holland en Den Helder (bijlagen 22B en 22C). De geschatte γ 's zijn hier duidelijk positief, hetgeen ook valt af te leiden uit de empirische frequentieverdelingen van de HW-opzetten van deze stations (bijlage 18). De gekozen kwantielschattingen zijn steeds die bij de hoogste k -waarde van een behoorlijk stabiel traject. De keuze ervan is minder subjectief dan bij de VVM- en GPV-methoden.

Bij de CON-methode zijn de betrouwbaarheidsbanden berekend volgens een parametrische bootstrap-methode (§ 6.6).

Uitgangspunt daarvoor was een GPV-verdeling voor de HW-opzetten waarvan de parameters geschat werden uit de waarnemingen behorende bij de gekozen k -waarde. T.o.v. de andere toegepaste methoden zijn deze k -waarden groot. Het aantal te gebruiken grootste waarnemingen wordt niet begrensd door een duidelijk afwijkend gedrag van de overschrijdingskromme beneden een bepaalde drempelwaarde (bijlage 18). Deze grote k -waarden leiden tot relatief kleine betrouwbaarheidsintervallen.

Methode GEV-s

Bij deze methode speelt k geen rol en is daardoor een objectieve methode. Het grootste bezwaar van deze methode is dat veel informatie ongebruikt blijft, omdat per stormseizoen alleen de hoogste waarde wordt beschouwd.

Conclusie:

In hoofdstuk 5 is op grond van theoretische overwegingen de voorkeur uitgesproken voor de methode VVM-0. Deze methode legt geen parametrische verdeling op aan de waarnemingen zelf. Gebleken is wel dat interpretatie van de uitkomsten van deze methode een subjectief element bevat. Gezien de bezwaren en/of subjectieve elementen die ook de andere methoden hebben blijft de voorkeur voor de methode VVM-0 overeind. De resultaten volgens deze methode geven de beste schattingen voor de kwantielen op basis van statistische analyse.

7.3 Betrouwbaarheidsintervallen.

Bij bestudering van de betrouwbaarheidsintervallen in tabel 6.6 en in de bijlagen 19 t/m 23 vallen enige zaken op:

1. de betrouwbaarheidsintervallen zijn groot;
2. gelet op de grootte van de betrouwbaarheidsintervallen liggen de kwantielschattingen per station volgens de verschillende methoden nog dicht bij elkaar;
3. de betrouwbaarheidsintervallen zijn symmetrisch, behalve bij de CON-methode;
4. de ondergrens vertoont een dalend verloop bij de methoden VVM-0, GPV en GEV-s;
5. de ondergrens voor het 10^{-4} -kwantiel ligt lager dan de hoogste bekende waarnemingen;

ad 1: Dat de betrouwbaarheidsintervallen van de 10^{-4} -kwantielen groot zijn heeft uiteraard te maken met de zeer grote mate van extrapolatie, maar komt vooral omdat de extreme-waardenindex γ niet op voorhand een bepaalde waarde is gegeven, dit in tegenstelling tot het „Deltarapport” waarin werd uitgegaan van $\gamma = 0$. Het meebeschouwen van de onzekerheid in de γ -schatting leidt tot veel grotere betrouwbaarheidsintervallen voor de kwantielschattingen (vergelijk ook de schattingen volgens de methode VVM-c). Het lijkt niet verstandig deze reële onzekerheid te verdoezelen door op voorhand een vaste waarde van γ te kiezen.

Ook hieruit blijkt dat een klein betrouwbaarheidsinterval niet garandeert dat het statistische model goed is.

ad 2: In beginsel geldt dat het betrouwbaarheidsinterval groter wordt naarmate minder veronderstellingen worden gedaan en dus meer parameters geschat moeten worden. Dat de betrouwbaarheidsintervallen volgens de GEV-s-methode voor de stations Vlissingen, Hoek van Holland en Delfzijl het grootst zijn heeft ook nog te maken met het beperkte aantal gebruikte waarnemingen (alleen stormseizoenmaxima). Bij de stations Den Helder en Harlingen zijn de betrouwbaarheidsintervallen relatief klein door het duidelijk negatief zijn van γ (de kansverdeling heeft in dat geval een bovengrens).

Dat de uitkomsten van de kwantielschattingen volgens de verschillende methoden in het licht van de grootte van de betrouwbaarheidsintervallen toch vrij dicht bij elkaar liggen komt omdat wordt uitgegaan van dezelfde basisgegevens en omdat de methoden een zekere verwantschap hebben door het gebruik van de extreme-waardenindex γ .

ad 3: De symmetrie van de betrouwbaarheidsintervallen volgt uit de manier waarop ze zijn berekend. Verwacht mag worden dat ze dat in werkelijkheid niet zijn.

De betrouwbaarheidsintervallen berekend volgens het VVM-0-model, het GPV-model en het GEV-s-model zijn symmetrisch omdat de kwantielschattingen asymptotisch normaal zijn, d.w.z. de ze hebben voor grote n bij benadering een normale verdeling. Voor het GPV-model en het GEV-s-model volgt dit mede uit het schatten van de parameters volgens de methode van de grootste aannemelijkheid. Dit laat zich het gemakkelijkst aantonen voor de GPV-methode omdat hierbij twee parameters geschat moeten worden (σ_v en γ , zie formule (6.15)); bij de GEV-s-methode zijn dat er drie (μ , σ en γ , zie formule (6.19)). De gebruikelijke procedure voor het schatten van de betrouwbaarheidsintervallen voor het GPV-model wordt beschreven in Aanhangsel 2. Voor het GEV-model geldt een soortgelijke beschouwing.

ad 4 en 5: De betrouwbaarheidsintervallen rond de kwantielschattingen van de overschrijdingslijnen van hoofdstuk 6 zijn telkens apart berekend voor een reeks waarden van p_0 . Voor zeer kleine p_0 leidt deze (asymptotisch normale) benadering in relatie tot de beschikbare lengte van de waarnemingsreeksen tot een absurditeit m.b.t. de ondergrenzen van de betrouwbaarheidsintervallen. Bij de VVM-0 methode treedt deze absurditeit ook op, zij het dat de afleiding van de betrouwbaarheidsintervallen heel anders is. Bij de convolutiemethode zijn de betrouwbaarheidsintervallen kleiner omdat in beginsel meer waarnemingen ter beschikking stonden. Bovendien zijn ze berekend volgens een parametrische bootstrap-methode. Omdat hierbij een zeer groot aantal overschrijdingslijnen wordt berekend, die elk uiteraard monotoon niet-dalend zijn, doet deze absurditeit zich bij deze aanpak niet voor.

De bootstrap-methode geeft in beginsel voor de onderhavige problematiek een beter beeld van de (scheve) vorm van de kansverdeling van de kwantielschattingen, en dus ook van de 95%-betrouwbaarheidsintervallen, dan de in hoofdstuk 5 besproken „normale” (symmetrische) benadering van deze kansverdeling. Om meer inzicht te krijgen in de vorm van de verdeling van de kwantielschattingen is de parametrische bootstrap-methode toegepast bij het (overigens niet gekozen) GPV-model. Voor de „verdelingsvrije”

methode VVM-0 is deze werkwijze in beginsel niet toepasbaar. Via een aangepaste procedure, waarvan de theoretische onderbouwing zwak is, zijn toch bootstrap-betrouwbaarheidsintervallen voor de VVM-0-kwantielschattingen bepaald.

De waarden van deze bootstrap-betrouwbaarheidsintervallen worden slechts ter kennisneming aangeboden. Zij kunnen nog niet worden gebruikt bij het vaststellen van de aan te bevelen betrouwbaarheidsintervallen. De toegepaste procedures, alsmede de resultaten ervan, worden besproken in Aanhangsel 3.

7.4 Bijzondere selecties

Jaarmaxima i.p.v. stormseizoenmaxima

De toepassing van de extreme-waardentheorie vereist gelijkverdeelde en onafhankelijke waarnemingen. In verband daarmee is het (lange) stormseizoen ingevoerd. Daarmee blijft een deel van de oorspronkelijk beschikbare reeks waarnemingen buiten beschouwing. Dat is verantwoord mits de kansbijdrage van dit deel aan de overschrijding van de extreem hoge stormvloedpeilen, waarover het in dit onderzoek gaat, verwaarloosbaar klein is. In dat geval kan voor dergelijke peilen de overschrijdingskans per stormseizoen gelijk worden gesteld aan de overschrijdingskans per jaar. Gezien de kansverdelingen en de aantallen van de geselecteerde HW-opzetten buiten het gekozen lange stormseizoen (§ 4.3) is dat in het onderhavige onderzoek het geval voor de methoden waarbij alle geselecteerde HW's dan wel HW-opzetten statistisch worden verwerkt. Bij de GEV-methode verdient dit aspect echter nadere beschouwing.

De GEV-methode wordt in de regel toegepast op jaarmaxima, waarbij in het onderhavige geval als jaar de periode van 1 juli t/m 31 juni („stormjaar”) wordt gekozen om onafhankelijkheid van de jaarmaxima te garanderen. Bij bestudering van de gegevensreeksen is gebleken dat niet alle jaarmaxima in het stormseizoen vallen. Bij Vlissingen is dat voor 17 seizoenen het geval, bij Hoek van Holland voor 10, bij Den Helder 3, bij Harlingen 4 en bij Delfzijl 13 (zie bijlage 34). Ter vergelijking is daarom de GEV-methode ook toegepast op de (storm)jaarmaxima (GEV-j), ofschoon daarmee uiteraard niet is voldaan aan de eis van gelijke verdeling van de waarnemingen. Door nu op voorhand $\gamma = 0$ te stellen wordt de uitkomst verkregen die is aangegeven bij methode GUMBEL-j; dit is dus een bijzonder geval van de GEV-j-methode. Ter vergelijking zijn ook de schattingen voor de stormseizoenmaxima, d.w.z. volgens de methode GEV-s (tabel 6.7), in tabel 7.1 opgenomen.

Voor de noordelijke stations liggen de schattingen volgens het GEV-j-model hoger dan voor het GEV-s-model en is de onzekerheid veel groter.

Tabel 7.1

Schattingen van het 10^{-4} -kwantiel [cm boven NAP] en het bijbehorende 95%-betrouwbaarheidsinterval [cm] m.b.v. het GEV- en GUMBEL-model op (storm)jaarmaxima (1 juli . . . 31 juni)

methode	Vlissingen	Hoek van H.	Den Helder	Harlingen	Delfzijl
GEV-j	565 ± 165	505 ± 170	420 ± 165	450 ± 125	750 ± 350
GUMBEL-j	540 ± 35	480 ± 40	465 ± 60	580 ± 70	680 ± 60
GEV-s	585 ± 190	510 ± 180	390 ± 110	420 ± 65	650 ± 200

Splitting van de reeksen in even en oneven stormseizoenen

Een manier om na te gaan hoe gevoelig de berekeningsresultaten van een bepaald station zijn voor de „toevallig“ voor dat station beschikbare gegevensreeks, die opgevat kan worden als een steekproef uit een populatie met onbekende kansverdeling, is het splitsen van de reeks van dat station in twee delen, om vervolgens de berekeningsresultaten voor de hele reeks en die voor de halve reeksen onderling te vergelijken. Een dergelijke berekening is uitgevoerd voor het station *Vlissingen*. Gekozen is voor een splitsing in even en oneven stormseizoenen; bij een even (oneven) stormseizoen valt de maand januari in een even (oneven) jaartal. Daarmee wordt de invloed van een eventuele (geringe) trend in de tijd op de resultaten geëlimineerd. Van een eventuele tweejaarlijkse periodiciteit in het stormvloedklimaat is niets bekend.

De verschillen tussen de resultaten van de gesplitste reeksen zijn opmerkelijk (zie tabel 7.2). De kwantielschattingen voor de even stormseizoenen liggen voor de verschillende methoden lager, die voor de oneven stormseizoenen in het algemeen hoger dan die voor de gehele reeks. De Gumbel-methode is uiteraard minder gevoelig omdat de γ hierbij op voorhand gelijk aan nul is gesteld.

Bij de convolutiemethode is het berekenen van betrouwbaarheidsintervallen achterwege gebleven. Voor de oneven stormseizoenen bleek het verloop van de kwantielschattingen als functie van de drempel dermate instabiel dat een waarde niet goed was vast te stellen.

Tabel 7.2

Schattingen van het 10^{-4} -kwantiel [cm boven NAP] en het bijbehorende 95%-betrouwbaarheidsinterval [cm] voor *Vlissingen* bij gesplitste reeksen

reeks	VVM-0	CON	GEV-j	GUMBEL-j
alle stormseizoenen	540 ± 140	540	565 ± 165	540 ± 35
even stormseizoenen	475 ± 130	480	500 ± 140	525 ± 45
oneven stormseizoenen	540 ± 200	-	665 ± 440	550 ± 50

Het verloop van de kwantielschattingen als functie van k volgens de methode VVM-0 zijn voor de even en oneven stormseizoenen gegeven in de bijlagen 24A en 24B. De reeks van de even stormseizoenen neigt duidelijk naar een negatieve γ -waarde; de reeks voor de oneven stormseizoenen geeft een redelijk stabiele schatting $\gamma = 0$.

Gegevens vanaf het stormseizoen 1932/1933

Van Den Helder en Harlingen zijn alleen gegevens vanaf het stormseizoen 1932/1933 bruikbaar voor statistische analyse vanwege het gereedkomen van de Afsluitdijk in 1932. Nagegaan is wat voor de andere drie stations het resultaat is volgens de VVM-0-methode bij het gebruik van de gegevens vanaf het stormseizoen 1932/1933. De bijlagen 25A, B en C geven de γ - en de 10^{-4} -kwantielschattingen als functie van k . De aangegeven k_{max} is ongeveer de helft van die van de gehele reeksen. De gekozen kwantielschattingen zijn duidelijk lager dan die van de gehele reeksen, hetgeen samenhangt met de gemiddeld wat lagere waarden van γ ; vergelijk ook de bijlagen 20A, B en E. De uitkomsten zijn verzameld in tabel 7.3.

Tabel 7.3

Schattingen van het 10^{-4} -kwantiel [cm boven NAP] en het bijbehorende 95%-betrouwbaarheidsinterval voor de gehele reeksen en de reeksen vanaf het stormseizoen 1932/33 (methode: VVM-0)

reeks	Vlissingen	Hoek van Holland	Delfzijl
geheel	540 ± 140	500 ± 190	600 ± 150
vanaf 1932/33	520 ± 170	460 ± 200	545 ± 165

Conclusie:

Uit de resultaten van de berekeningen met de verschillende deelreeksen kan geconcludeerd worden dat de invloed van de „toevallig“ beschikbare waarnemingsreeksen behoorlijk groot is. Dit is in overeenstemming met de brede betrouwbaarheidsintervallen van de kwantielschattingen.

De reeksen vanaf het seizoen 1932/33 geven alle voor de stations met langere reeksen lagere uitkomsten voor de geschatte 10^{-4} -kwantielen dan de volledige beschikbare reeksen. Gezien de toegepaste homogenisering kan dit verschil niet verklaard worden uit een systematische verandering van het HW-regime. De geconstateerde verschillen moeten dus als toevallig worden beoordeeld. Opgemerkt kan worden dat geringe, niet statistisch significant aantoonbare, verschillen in het HW- of HW-opzetklimaat door de grote mate van extrapolatie en de gevoeligheid voor de parameter γ toch dergelijke verschillen kunnen opleveren. Omdat het verschil zich bij alle drie stations met reeksen van ca. 100 jaar manifesteert, en er toch een bepaalde mate van samenhang tussen de stations aanwezig is, lijkt de conclusie gerechtvaardigd dat de uitkomsten voor de stations Den Helder en Harlingen enige decimeters lager zijn dan wanneer voor deze stations reeksen van vergelijkbare lengte met de andere stations beschikbaar zouden zijn geweest.

De betrouwbaarheidsintervallen zijn iets groter voor de reeksen vanaf het seizoen 1932/33. Enerzijds worden de betrouwbaarheidsintervallen kleiner naarmate γ kleiner is, anderzijds worden ze groter omdat minder waarnemingen beschikbaar zijn. Beide effecten compenseren elkaar hier min of meer.

7.5 Gevoeligheidsonderzoek m.b.t. de selectiemethode**De keuze van i in de D-S- i -selectie**

In § 4.3 is de D-S-4-selectie gekozen als de beste selectie voor de verwijdering van autocorrelatie uit de reeksen van HW-opzet voor alle stations. De keuze voor $i = 4$ is daarbij gebaseerd op de interpretatie van de grafieken van de bijlagen 9A t/m 9E. Het vaststellen voor welke waarde van i de krommen zich voldoende gestrekt hebben is echter niet geheel vrij van enige subjectiviteit. Voor sommige stations lijkt $i = 3$ ook een redelijke keuze. Daarom is nagegaan wat de uitkomst voor het 10^{-4} -kwantiel is volgens de methode VVM-0 voor de stations Hoek van Holland, Den Helder en Harlingen bij toepassing van een D-S-3-selectie.

De schattingen voor γ en het kwantiel als functie van k zijn gegeven in de bijlagen 26A, B en C. Vergelijking met de bijlagen 20B, C en D laat zien dat het verloop voor die k -intervallen, die bepalend zijn voor de kwantielschatting, vrijwel gelijk is (voor grotere k -waarden treden wel verschillen op). Een D-S-3-selectie levert vooral meer lage HW's en beïnvloedt daardoor het verloop van de overschrijdingskrommen bij de hoogste frequenties.

Voor Hoek van Holland, Den Helder en Harlingen werden respectievelijk 1830, 1031 en 1124 HW-opzetten en dus ook HW's geselecteerd.

Voor de D-S-4-selectie zijn dat er achtereenvolgens 1577, 866 en 936 (tabel 6.1).

Slotsom: Een keuze voor $i = 3$ i.p.v. $i = 4$ had niet tot andere uitkomsten geleid.

De keuze van de onderdrempel van HW-opzet

Om de gegevensreeksen te beperken tot relevante gevallen van waterstandsverhogingen t.o.v. het astronomisch getij werd voor HW-opzet een drempelwaarde ingevoerd van 30 cm. De Deltacommissie gebruikte als ondergrens voor de selectie van HW-opzetten 50 cm. Om na te gaan of de schattingen van de kwantielen gevoelig zijn voor de keuze van 30 cm of 50 cm zijn voor Hoek van Holland en Harlingen γ - en 10^{-4} -kwantielschattingen berekend met de methode VVM-0, waarbij geselecteerd werd met een drempelwaarde voor HW-opzet van 50 cm. De resultaten zijn gegeven in de bijlagen 27A en B. Vergelijking met de bijlagen 20B en D laat weer zien dat de eindresultaten voor de kwantielschattingen niet anders zouden zijn geweest.

7.6 Tijdens het onderzoek afgevalen methoden

De Pickands-schatter voor γ

Voor het schatten van de vormparameter γ bij de verdelingsvrije methoden is gebruik gemaakt van de zgn. momentenschatter, welke gegeven is in hoofdstuk 5, formule (5.15). Naast deze in de loop van het onderzoek ontwikkelde schatter is in het begin ook gewerkt met de „Pickands“-schatter voor γ , die gegeven is door formule (5.28). Het verloop van de kwantielschattingen als functie van k bij toepassing van de Pickands-schatter bleek echter dermate instabiel vergeleken bij dat voor de momentenschatter, dat besloten is de Pickands-schatter niet meer bij het onderhavige onderzoek te beschouwen. Bijlage 28 geeft als voorbeeld het verloop van beide schatters als functie van k voor de stations Hoek van Holland en Delfzijl. De grotere stabiliteit van de momentenschatter is gelegen in het grotere aantal waarnemingen dat voor de schatting gebruikt wordt. De momentenschatter gebruikt alle k grootste waarden, de Pickandsschatter gebruikt telkens slechts drie waarden - zie formule (5.28).

De koppelingsformule

Tijdens het onderzoek is getracht naast de CON-methode ook de verdelingsvrije methode VVM-c toe te passen op de HW-opzetten en deze te „koppelen“ aan de kansverdeling van de astronomische hoogwaterstanden m.b.v. de zgn. koppelingsformule, die gegeven is aan het slot van § 5.2. Deze aanpak bleek dermate afwijkende (soms lage) schattingen van het 10^{-4} -kwantiel op te leveren, dat aangenomen moet worden dat aan de voorwaarden voor het toepassen van de koppelingsformule niet is voldaan.

De aanpak veronderstelt in feite dat 10^{-4} -kwantielen alleen zullen voorkomen bij extreem hoge astronomische hoogwaterstanden, hetgeen strijdig is met de inzichten daaromtrent. Bovendien wordt niet altijd aan de voorwaarde voldaan dat de γ -schattingen voor HW-opzet kleiner dan nul moeten zijn. Deze weg is derhalve voor dit onderzoek niet begaanbaar gebleken.

7.7 Vergelijking met het statistische model van de Deltacommissie

In het kader dit onderzoek past een vergelijking met resultaten bij toepassing van het statistische model van de Deltacommissie op de huidige waarnemingsreeksen.

De werkwijze van de Deltacommissie is in het kort beschreven in § 3.1.1. Het mag hieruit duidelijk zijn dat het onverkort toepassen van deze werkwijze op de huidige reeksen behalve niet erg zinvol, zelfs niet mogelijk is. Het station Hellevoetsluis is niet meer bruikbaar voor het selecteren van depressies voor Hoek van Holland. Voor de andere stations bestaat geen „poort” voor depressiebanen. Bovendien moet bedacht worden dat nu voor de stations in de Waddenzee een reeks beschikbaar is met een lengte die vergelijkbaar is met die van Hoek van Holland ten tijde van de studie van de Deltacommissie. De langere reeksen maken het ook gewenst de gegevens te corrigeren voor relatieve zeespiegelrijzing.

De volgende werkwijze wordt daarom hier beschouwd als

overeenkomend met het statistische model van de Deltacommissie:

- per station wordt uitgegaan van de voor dat station geselecteerde en gecorrigeerde HW-standen;
- als drempel wordt de laagst aanvaardbare waarde aangehouden: u_{\min} (§ 6.2);
- de extreme-waardenindex γ wordt op nul gesteld. De GPV-verdeling wordt daarmee een exponentiële verdeling, de GEV-verdeling een Gumbel-verdeling;
- alleen HW-standen worden statistisch verwerkt.

De resultaten staan vermeld in tabel 7.4. Ter vergelijking is ook de methode VVM-0 toegepast op deze reeksen van gegevens, waarbij op voorhand $\gamma = 0$ is gesteld.

Tabel 7.4:

Schattingen van het 10^{-4} -kwantiel [cm boven NAP] en het bijbehorende 95%-betrouwbaarheidsinterval [cm] volgens het statistische model van de Deltacommissie

methode	Vlissingen	Hoek v.H.	Den Helder	Harlingen	Delfzijl
exponentieel	545 ± 30	520 ± 30	545 ± 45	720 ± 50	765 ± 40
Gumbel	545 ± 35	490 ± 40	485 ± 65	640 ± 85	710 ± 65
VVM-0 met $\gamma = 0$	550 ± 30	520 ± 30	520 ± 50	680 ± 55	740 ± 40

Zoals verwacht mocht worden zijn de 95%-betrouwbaarheidsintervallen klein door de veronderstelling $\gamma = 0$.

Met name voor de noordelijke stations levert deze aanpak conservatieve schattingen t.o.v. die gepresenteerd in hoofdstuk 6, geheel in overeenstemming met de gevonden waarden voor γ .

7.8 Resulterende overschrijdingslijnen en betrouwbaarheidsintervallen

In § 3.1.2 werd reeds opgemerkt dat de overschrijdingslijnen in het „Deltarapport” werden samengesteld uit deellijnen van verschillend karakter. Enerzijds zijn dat de overschrijdingslijnen berekend uit alle opgetreden hoogwaterstanden, anderzijds zijn dat de geëxtrapolerde overschrijdingslijnen, welke hun helling ontleen aan de overschrijdingslijn behorende bij de geselecteerde HW's voor Hoek van Holland. In dit rapport wordt afgezien van een dergelijke presentatie.

Bijlage 29 geeft de overschrijdingslijnen van de vijf onderzochte stations voor alle HW-standen, herleid naar de toestand van 1985, uit de volle beschikbare perioden (dus niet alleen de stormseizoenen). Daarnaast bevatten de bijlagen 30A t/m 30E de „empirische” overschrijdingslijnen van de wèl geselecteerde en naar 1985 herleide HW's uit (de lange stormseizoenen van) de betrokken waarnemingsperioden, met daarbij de geëxtrapoleerde overschrijdingslijnen met hun betrouwbaarheidsbanden (volgens de schattingen met de methode VVM-0). De betrouwbaarheidsbanden volgens de bootstrap-methode (zie aanhangsel 3) staan op bijlage 35. Schattingen zijn gemaakt voor frequenties van $5 \cdot 10^{-3}$ t/m 10^{-5} ; voor hogere frequenties is de toegepaste theorie niet meer geldig; daarvoor zijn andere methoden geëigend, die hier verder buiten beschouwing zijn gebleven.

In beginsel gelden de geëxtrapoleerde overschrijdingslijnen voor het lange stormseizoen; ze zijn daarvoor afgeleid. Omdat de kansbijdrage aan het overschrijden van dergelijke extreme stormvloedpeilen buiten het lange stormseizoen mag worden verwaarloosd (zie ook het begin van § 7.4) zijn de overschrijdingsfrequenties per stormseizoen hier gelijk aan die per jaar. Voor de hogere frequenties geldt dit niet.

In de praktijk blijkt er nog wel eens misverstand te bestaan over de interpretatie van de overschrijdingslijnen. De begrippen „overschrijdingsfrequentie” en „overschrijdingskans” worden namelijk nog al eens door elkaar gehaald. Voor de lage frequenties is dat geen probleem; ook in dit rapport zijn de berekende kwantiëlschattingen (stormvloedpeilen bij extreem lage overschrijdingskansen per stormseizoen/jaar) uitgezet als het gemiddelde aantal overschrijdingen per stormseizoen/jaar. Voor frequenties groter dan één keer per stormseizoen is al zonder meer duidelijk dat dit geen kans per stormseizoen kan zijn. Voor alle duidelijkheid: een frequentie is een aantal per tijdseenheid en een kans is een getal ≥ 0 en ≤ 1 ; een kans wordt bij onderhavige problematiek steeds betrokken op een willekeurig stormseizoen of jaar.

De overschrijdingsfrequenties van peilen berekend uit alle (dus niet geselecteerde) HW-standen (bijlage 29) mogen nooit als kans worden geïnterpreteerd omdat geen sprake is van onafhankelijke waarnemingen. Bovendien is dit waarnemingsmateriaal niet homogeen.

Voor de betrekking tussen de overschrijdingsfrequentie en de overschrijdingskans per stormseizoen voor onafhankelijke en gelijkverdeelde waarnemingen (HW's) het volgende:
Als m het gemiddelde aantal keren per stormseizoen is dat een bepaald niveau wordt overschreden, dan is de kans op r gebeurtenissen (gebeurtenis = overschrijden van het betreffende niveau) in een willekeurig stormseizoen volgens de wet van Poisson gelijk aan

$$P(X = r) = \frac{m^r e^{-m}}{r!}$$

De kans dat het betreffende niveau in een willekeurig stormseizoen niet wordt overschreden (dus $r = 0$) is dan

$$P(X = 0) = e^{-m}$$

De kans dat het betreffende peil in een willekeurig stormseizoen wel zal worden overschreden is dus

$$P(x \neq 0) = 1 - e^{-m}$$

Hiermee is dus de betrekking verkregen tussen m , het gemiddelde aantal keren per stormseizoen dat een bepaald niveau wordt overschreden en P , de kans per stormseizoen op overschrijden van dat niveau.

Ontwikkeld in een Taylor-reeks is deze uitdrukking te schrijven als

$$P(x \neq 0) = m - \frac{m^2}{2!} + \frac{m^3}{3!} - \dots$$

Hieruit valt eenvoudig te zien dat

$$P(x \neq 0) \approx m \quad \text{voor kleine } m$$

Voor waarden van m kleiner dan 0,1 is deze benadering in de praktijk reeds voldoende nauwkeurig.

7.9 Hydraulische aspecten en ruimtelijke samenhang

De uitkomsten van het statistische onderzoek, samengevat in tabel 6.7, geven het opmerkelijke beeld dat de geschatte 10^{-4} -kwantielen voor Den Helder en Harlingen aanmerkelijk lager uitvallen dan de bestaande basispeilen. Dat geldt nog sterker als wordt gekeken naar de 2,5%-bovengrenzen van de betrouwbaarheidsintervallen. Men kan zich afvragen of dergelijke verschillen vanuit hydraulisch oogpunt gezien wel reëel zijn. Ook als rekening wordt gehouden met het gestelde in § 7.4 t.a.v. de reeksen vanaf 1932/33 en in verband daarmee een correctie van enige decimeters wordt toegepast zijn de uitkomsten nog relatief laag.

In § 7.1 werd al opgemerkt dat de empirische overschrijdingslijnen voor Den Helder en Harlingen een duidelijk ander beeld vertonen dan die voor de andere stations. Omdat voor de statistische analyse geen andere mogelijkheid openstaat dan te veronderstellen dat het gedrag van deze lijnen in het te extrapoleren gebied niet anders is dan in het waarnemingsgebied, worden voor deze stations relatief lage kwantielschattingen gevonden. De vraag is dan of deze veronderstelling juist is. Als het „gedrag van het water“ (§ 2.1.2) bij buitengewoon hoge stormvloed duidelijk anders is dan bij opgetreden stormvloed, dan wordt daarmee de juistheid van de statistische extrapolatie ondergraven. Mogelijke effecten die in de ondiepe Waddenzee bij zeer hoge stormvloed een rol kunnen spelen zijn:

- een vertraging in het vollopen van de westelijke Waddenzee vanwege de beperkte doorstroomcapaciteit van de zeegaten (in dat geval speelt de stormduur een belangrijke rol);
- een verandering van de lokale opwaaiing in de Waddenzee, omdat bij hogere standen boven het relatief hoge voorland zich een krachtiger retourstroom zich kan ontwikkelen, hetgeen de opwaaiing vermindert;
- een verandering van de waterbeweging doordat de platengebieden meer zullen bijdragen aan het watertransport bij hogere standen;

Meer inzicht hieromtrent kan worden verkregen door middel van onderzoek met zgn. mathematisch-fysische rekenmodellen (het uitvoeren van een dergelijk onderzoek voor de westelijke Waddenzee lijkt dan ook gewenst). Daarbij dient gebruik te worden gemaakt van de samenhang tussen stations binnen en buiten de Waddenzee. Delfzijl is gelegen aan het Eems-Dollard-estuarium. Stormvloedstanden bij Delfzijl worden daarom sterk beïnvloed door de lokale windrichting: is deze juist het estuarium in gericht dan leidt dit tot extra hoge standen, zeker als dat tijdens het hoogtepunt van de storm het geval is. Vaak wijkt de lokale windrichting daar van af. Men zou dit kunnen uitdrukken door te zeggen dat Delfzijl een „smal doel" heeft t.o.v. bijvoorbeeld een station als Hoek van Holland. Of dit „doel" geraakt wordt is weer een kwestie van kans. In beginsel is dit effect in de waarnemingen aanwezig. Omdat de andere voor de westelijke Waddenzee genoemde aspecten voor Delfzijl niet zo'n grote rol spelen, lijkt er voor dit station geen aanleiding te twifelen aan de juistheid van de extrapolatie.

In dit rapport zijn de gegevens voor elk station afzonderlijk geanalyseerd. De verkregen resultaten gelden dan ook steeds alleen voor dat bepaalde station.

In werkelijkheid zal er tussen de stations een bepaalde mate van samenhang zijn en wel meer naarmate de stations dicht bij elkaar liggen. Zo is het uiterst onwaarschijnlijk dat de „werkelijke" overschrijdingslijn voor Vlissingen in de buurt van de bovengrens van de betrouwbaarheidsband zou liggen en die voor Hoek van Holland in de buurt van de ondergrens. Dergelijke beschouwingen zijn in dit onderzoek niet meegenomen.

Met behulp van recent ontwikkelde technieken is het mogelijk gebruik te maken van de samenhang die in de gegevens voor de verschillende stations zit. Daarmee kunnen vragen beantwoord worden als: wat is de kans dat in een bepaald jaar een bepaald extreem waterstandsniveau wordt overschreden in bijvoorbeeld Hoek van Holland en IJmuiden - of: wat is de kans dat òf in Hoek van Holland òf in IJmuiden een bepaald niveau wordt overschreden? Dergelijke vragen komen op als men wil weten wat de kans is dat ergens langs een dijkkring een bepaald niveau wordt overschreden.

De beantwoording van dit soort vragen, die thans binnen het bereik van de huidige stand van de theorie ligt [De Haan 1985; De Haan en Resnick 1991], maar waarvoor de toepassing veel extra werk vergt, is buiten het huidige onderzoek gebleven.

7.10 West-Terschelling en IJmuiden

Door de Hoofddirectie van de Rijkswaterstaat is in de laatste fase van het onderzoek gevraagd om ook voor de stations West-Terschelling en IJmuiden op grond van statistische analyse een schatting te geven van het 10^{-4} -kwantiel. Op de meetreeksen van de stations West-Terschelling en IJmuiden zijn daartoe dezelfde bewerkingen toegepast als op die van de stations waar het onderhavige onderzoek zich op gericht heeft. IJmuiden beschikt over een lange meetreeks.

Daarom is ook voor IJmuiden de reeks vanaf 1932/33 geanalyseerd. De resultaten van de analyses met de VVM-0-methode zijn gegeven in de bijlagen 31, 32 en 33.

De volgende schattingen van de 10^{-4} -kwantielen zijn daaruit bepaald:
Voor West-Terschelling(1932/33...1984/85): NAP + 350 (\pm 60) cm
en voor IJmuiden (1883/84...1984/85): NAP + 530 (\pm 210) cm
(1932/33...1984/85): NAP + 470 (\pm 240) cm.

De getallen tussen haken betreffen het 95%-betrouwbaarheidsinterval. Opmerkelijk is weer het verschil tussen de schattingen van IJmuiden voor de korte en de lange periode en de relatief lage uitkomst voor West-Terschelling. Deze uitkomsten ondersteunen de conclusie van § 7.4 m.b.t. de uitkomsten voor de stations aan de westelijke Waddenzee in verband met de beschikbare waarnemingsreeksen. Van alle in dit onderzoek beschouwde stations worden voor IJmuiden de grootste 95%-betrouwbaarheidsintervallen gevonden, zowel voor de volle periode als voor de periode na mei 1932. Dat is vooral het gevolg van de relatief hoge γ -waarden voor IJmuiden. In deze onzekerheid van de kwantielschattingen zit ook de relatieve onnauwkeurigheid van de HW's te IJmuiden besloten (zie § 4.1). In welke mate deze onnauwkeurigheid bijdraagt aan de breedte van deze betrouwbaarheidsintervallen is niet te zeggen.

8. Historische bevindingen, Zeeberen en Zeebevingen

8.1 Historische zware stormvloed

In de tweede helft der jaren tachtig werd bij de Dienst Getijdewateren van de Rijkswaterstaat besloten een inventarisering te maken van bewaard gebleven historische getijwaterstanden in Nederland en de onmiddellijk daarop aansluitende Belgische en Duitse gebieden. Daartoe werd niet alleen eigen materiaal nagelopen, veel literatuur geraadpleegd en gebruik gemaakt van welwillend ter beschikking gestelde aantekeningen van (nu wijlen) mevr. dr. M.K.E. Gottschalk en van dr. H. Schoorl, maar ook op ruime schaal archiefonderzoek verricht. In 1989 verscheen een overzicht van gevonden gegevens met een eerste analyse ervan [Jonkers 1989]. Een geheel herziene en sterk uitgebreide versie van het inventariseringsgedeelte van deze studie is in voorbereiding; het ligt verder in het voornemen een analyse van de gevonden stormvloedstanden en één van aangetroffen „reguliere” getijwaterstandswaarnemingen in afzonderlijke nota's te doen verschijnen. Vooruitlopend op een en ander worden in deze paragraaf enkele reeds thans duidelijke relevante bevindingen weergegeven. De stelselmatige inwinning van getijwaterstanden begon voor zover bekend in ons land en wel in Amsterdam op 1 januari 1700*. De aflezingen werden uurlijks en voor een deel halfuurlijks verricht en voortgezet tot in de tweede helft der 19^e eeuw een peilschrijver werd geplaatst. Uit deze lange reeks 1700-1932 (waarvan de waarden uit het tijdvak 1750-1766 echter verloren zijn gegaan) is een toonaangevende grafiek voor de rijzing van de gemiddelde zeespiegel gedurende die periode afgeleid [Van Veen 1945]. Ook op enkele andere plaatsen in ons land werden in de 18^e eeuw getijwaarnemingen verricht, maar de meeste reeksen zijn nogal kort en betreffen merendeels alleen dagwaarnemingen, soms alleen daghoogwaterstanden. In de 19^e eeuw neemt het aantal peilmeetstations in het tijgebied toe en zijn de bewaard gebleven reeksen vaak ook langer, maar het blijft in het algemeen bij dagwaarnemingen, die soms zijn aangevuld met een beperkt aantal nachtelijke stormstanden. Van enkele getijstations staan over een vrij lange tijd dergelijke reeksen ter beschikking (voor de basisstations is dit het geval met Delfzijl). Maar ook wanneer dergelijke reeksen zich over vele tientallen jaren uitstrekken kleeft er vanuit statistisch oogpunt een belangrijk nadeel aan, te weten de beperking tot vrijwel uitsluitend dagstanden. Voor bepaalde stations zijn er trouwens ook nog problemen met betrekking tot hun opstelling en/of hun referentiehoogte en soms zelfs het gebezigde maatstelsel. Bij de overgeleverde stormvloedhoogten past de kanttekening dat zij betrekking zullen hebben gehad op de hoogste bereikte standen van de betrokken stormvloeden; zo'n topstand kan uiteraard zeer wel de top van een bui-oscillatie zijn geweest.

* een reeks halfuurlijkse nachtelijke waarnemingen aldaar gedurende 10 maanden in 1683/4 was daaraan overigens voorafgegaan

Niet alleen hierom past enig voorbehoud bij het vergelijken van deze standen met die van veel latere stormvloed, ook de toenmalige hoogte van de gemiddelde zeespiegel vormt een onzekerheidsfactor terwijl op veel plaatsen de waterstaatkundige toestand sterk afweek van de huidige, wat evenzeer van invloed op de stormvloedstanden kan zijn geweest.

De oudste bewaard gebleven stormvloedstanden in ons land zijn voor zover dezerzijds bekend vergelijkende standen van 1421, 1446 en 1468 te Rotterdam, die echter moeilijk of niet zijn te herleiden. Beperkt men zich tot stormvloedstanden, die op aanvaardbare wijze zijn te relateren aan de huidige waterstaatkundige toestand dan betreffen de oudste de beruchte *Allerheiligenvloed van 1570* (die gerekend naar onze huidige kalender overigens niet op 1 maar op 11 november zou zijn opgetreden). Deze stormvloed is berucht vanwege de zeer omvangrijke overstromingen, de grote aantallen slachtoffers en de enorme schade, die hij veroorzaakte. Volgens een bron uit 1682 [Schotel 1922, bl. 99] zou te Middelburg de topstand van 1570 lager zijn geweest dan die van 26 januari 1682; zou dit inderdaad het geval zijn geweest (wat niet zonder meer vaststaat) dan zouden daar tot 1850 een keer of 10 hogere stormvloedstanden zijn opgetreden dan in 1570. In Zuid-Holland bleven de kort daarvoor (tot NAP + 3,60m ?) verhoogde *Maas- en IJsseldijken van Delfland en Schieland* waterkerend (de Zuidhollandse eilanden echter overstromden). In Scheveningen stroomde het water met veel geweld de kerk in en bereikte er een hoogte van 3 (of 3½) voet plus 2 duim oftewel van NAP + 400 cm c.q. 415 cm (in 1953: NAP + 397 cm) [Schotel 1922, bl. 69, 75, 76, 80; Gottschalk 1975, bl. 665, 666, 670, 705]. Het is intussen aannemelijk dat dit de topstand van een kortstondige secundaire golf is geweest en niet een stormvloedstand in de huidige betekenis (het onbeschadigd blijven staan van de kerk lijkt moeilijk te rijmen met een bij zo'n stormvloedstand behorend afslagprofiel van het duin). Nabij Petten bereikte het water in de Zijpepolder een hoogte van 13 voet, wat op NAP + 400 cm is geschat [Gottschalk 1975, bl. 677, 706]. Deze schatting heeft nog een nadere beschouwing* (in 1953 was het stormvloedpeil te Petten NAP + 350 cm [Rijkswaterstaat en KNMI 1961, bl. 67]). In het zeer zwaar getroffen Friesland stond het water in de (op een terp staande) kerk van Metslawier (6 km NW van Dokkum) 1 voet hoog, wat neerkomt op NAP + ong. 380 cm; dit overtreft de hoogte van de hoogste bekende standen van het dijkvak Nieuw Bildt - Holwerd (NAP + 370 cm) van 22 december 1954 (maar niet die aan de Friese stations aan de voormalige Lauwerszee, die 20 tot 50 cm hoger opliepen in 1906 en 1944). Voor *Emden* kan de bereikte stand worden geschat op NAP + 470 cm, d.i. ongeveer even hoog als in 1825 (de zwaarste stormvloed, waarvan voor Noord-Nederland de stormvloedstanden goed bekend zijn). Voor Amsterdam lijkt de meest aannemelijke waarde AP + 88 duim (= 227 cm); die stand kan beïnvloed zijn door vele dijkdoorbraken in de buurt. Tot 1826 is deze stand overigens 9 maal overschreden met als hoogste stand (N)AP + 250 à 252 cm op 5 februari 1825. Tussen 1570 en 1825 zijn een aantal zware stormvloed opgetreden, waarvan sommige forse overstromingen teweegbrachten.

* van de oorspronkelijke twee geulen tussen Petten en het noordelijk ervan gelegen Huisduinen was de Zijpe opgezand en in 1553 ingepolderd.

De belangrijkste van deze stormvloedendeden zich voor op

- 26 januari 1682 (met name in het Zuidwesten van het land);
- 25 december 1717 (de zgn. Kerstvloed);
- 14 en 15 november 1775;
- 20 en 21 november 1776 (vooral aan de oostelijke kust van de Zuiderzee).

De indruk, die deze stormvloedend maakten, blijkt o.a. hieruit dat latere stormvloedstanden veelvuldig met die van een der genoemde stormvloedend werden vergeleken. Deze standen werden echter weer overtroffen door de hoogste standen uit de stormvloedperiode 3 - 5 februari 1825.

In de stormvloedlijst van de Rijkswaterstaat figureren van een aantal stations de tot NAP herleide hoogste stormvloedstanden van februari 1825 [TO 1961-1970, bl. 414 e.v.]. Aan vrijwel al die stations werden bij de stormvloed van 1953 nog hogere standen bereikt. In Zeeland en West-Brabant beliep dit verschil 5 à 10 dm; in het tijdsverloop tussen beide stormvloedend werden daar de topstanden van 1825 overigens plaatselijk (in Bergen op Zoom en Zierikzee) enkele malen, elders in het geheel niet overtroffen. In Katwijk beliep het verschil tussen de topstand van 1953 en die van 1825 slechts 36 cm, terwijl in Delfzijl het peil van 1825 tot op heden niet is overtroffen. Langs de haven van die plaats deed zich toen een doorbraak voor, terwijl daar bovendien de zijlen (sluizen) naar het Damsterdiep geruime tijd overliepen. Aangezien de peilschaal zich zeer waarschijnlijk bij die sluizen bevond is het alleszins aannemelijk dat zich ter plaatse enige afzuiging heeft doen gevoelen, waardoor de hoogste buitenwaterstand een 5 à 10 cm hoger zou zijn geweest. Deze stand is verder herleidbaar naar de toestand 1985 omdat voor die plaats ook de daghoogwaterstanden over de periode 1827-1881 beschikbaar zijn. Aangezien het jaargemiddelde daarvan over die periode geen trend vertoont (bijlage 13) zou de opgetreden waterstand te Delfzijl op 4 februari 1825, herleid naar de toestand 1985, zijn geweest (zie tabel 4.3): $\text{NAP} + 460 + 5 \text{ à } 10 + 30 \text{ cm} = \text{NAP} + \sim 500 \text{ cm}$. Voor de andere basisstations is een dergelijke afleiding voor de stormvloedstand van 1825 niet mogelijk.

Opmerkelijk voor wat betreft Delfzijl zijn nog de topstanden van 1570, 1717 en 1776 en andere gegevens voor Emden [Rohde 1977, bl. 59-61]. Volgens een nadere analyse daarvan zouden de topstanden tijdens die drie stormvloedend in dit gebied de topstand, opgetreden in 1825, heel weinig hebben ontlopen. Dit zou inhouden dat in ruim 400 jaar tijds de, naar de toestand van 1825 herleide, stand $\text{NAP} + \sim 500 \text{ cm}$ vier maal ten naastebij zou zijn bereikt, wat voor dat peil een overschrijdingsfrequentie van ongeveer 10^{-2} p.j. zou medebrengen. Dit zou goed passen in de voor Delfzijl, toestand 1985, geldige overschrijdingslijnen (bijlagen 29 en 30). Bij deze bevinding past echter enig voorbehoud omdat zich in die 400 jaar een klimaatswijziging heeft voorgedaan, te weten het aflopen van de zogenaamde Kleine IJstijd, die ook het HW-opzet-klimaat in dit gebied beïnvloed zou kunnen hebben.

8.2 Bijzondere waterstandsverheffingen

Afwijkingen van enige betekenis ten opzichte van het astronomische getij zijn aan onze kusten in veruit de meeste gevallen gevolg van de

weersomstandigheden, die zich gelijktijdig of kort daarvoor hebben voorgedaan boven de Noordzee. Een nadere beschouwing over deze verschijnselen en omstandigheden is in algemene zin gegeven in § 2.1.2. Men kan over zo'n afwijking een verwachting uitspreken *tegelijk met de weersverwachting; door de gevoeligheid voor relatief kleine verschillen in bepaalde elementen der weersgesteldheid* (b.v. de windrichting) en door de wisselwerking tussen weer en astronomisch getij is de onnauwkeurigheid van zo'n verwachting voor zover zij grote verhogingen inhoudt echter in het algemeen pas aanvaardbaar voor uitspraken minder dan 8 uur vooruit. Forse waterstandsverhogingen tijdens stormen gaan verder vaak gepaard met bui-oscillaties en buistoten. In § 4.2.1, waar zij ter sprake kwamen, is uiteengezet dat de invloed van omstreeks HW optredende buistoten goeddeels in de waarden van HW en HW-opzet is begrepen en dus in de statistische analyses is „meegenomen”. Ook is daar vermeld dat bij het ontwerpen van zeekeringen met deze verschijnselen door middel van de „toeslagregel van Wemelsfelder” rekening moet worden gehouden. Deze secundaire golfverschijnselen zijn onvoorspelbaar; hooguit zou men uit de weersverwachting kunnen opmaken dat hun aanstaand optreden in meerdere of mindere mate waarschijnlijk is.

Volledigheidshalve:

zowel bui-oscillaties als buistoten worden verwekt door weerkundige verschijnselen boven zee (§ 4.2.1), niet ver van de kust.

Aan het slot van § 2.1.2 is gewag gemaakt van nog andere oorzaken, die waterstandsverhogingen van betekenis aan onze kusten kunnen teweegbrengen. Alvorens deze te bespreken verdient het volledigheidshalve de aandacht dat buistoten ook wel geheel *onderscheiden van het meest nabije HW optreden. Voor zover daarvan in de waarnemingsperioden sprake is geweest zijn zij niet in de bestanden van HW en HW-opzet (en dus ook niet in de geanalyseerde tijdreeksen) opgenomen, ook niet als hun topstand dat nabije HW overtrof.* De reden daarvan is dat zij losstaan van de algemene generering van het stormeffect (zie § 2.1.2); met de eerdergenoemde toeslagregel is hun mogelijke invloed bij extreme stormvloed en geacht te zijn verrekend.

Een ander punt is dat buistoten zich kunnen voordoen in situaties, die op zich geen belangrijke waterstandsverhogingen doen verwachten, b.v. bij forse zuidwestelijke winden en buig weer. Hun invloed kan overrompend zijn; zo te Petten op 26 april 1924, toen omstreeks laagwater geheel onverwachts het water in 15 minuten tijds eerst 130 à 160 cm wies en daarna omstreeks 110 cm daalde.

Een verklaring zou kunnen zijn dat een zich noordwaarts verplaatsende buistoot daar samentrof met een landwaarts gaande bui-oscillatie. Iets vergelijkbaars zou zich hebben voorgedaan bij Zandvoort op 21 mei 1792 [Van Malde 1990, par. 6.4.3 resp. 6.2.4 ad d en par. 7.2]. Voor dergelijke verschijnselen geldt hetzelfde als hiervoor over buistoten „vrij van het nabije HW” is opgemerkt.

Van de eerdergenoemde andere oorzaken van waterstandsverhogingen aan onze kusten heeft de belangrijkste als bron stormweer dat elders is opgetreden. Wanneer dat stormweer daar op volle zee een *omvangrijke (zeer platte) „waterheuvel” heeft gegenereerd kan deze „waterheuvel” zich onder daartoe gunstige omstandigheden verplaatsen naar een gebied waar het niet meer stormt of zelfs niet gestormd heeft.*

Nadert zo'n „waterheuvel" de kust dan zal de waterstand hoger worden, waarmee aan de kust een waterstandsverhoging van betekenis kan optreden.

Dit verschijnsel doet zich aan onze kust op gezette tijden voor. De oorspronkelijke „waterheuvel" kan in zo'n geval zijn ontstaan in het noordelijk deel van de Noordzee, op de Noorse Zee of op de Atlantische Oceaan. In het laatste geval benoemen de Britten, met hun zoveel langere kust, het verschijnsel; zij spreken dan van een „external wave". Het ligt in de rede voor dit verschijnsel aan onze kusten één benaming te bezigen, ongeacht welk van de genoemde zeegebieden ontstaangebied is. Een oude benaming is „moerzee", een nieuwere - in navolging van het Duits - is „zeebeer".

Een zeebeer kan zich manifesteren na en bij rustig weer op de zuidelijke Noordzee, maar evenzeer tijdens een stormvloed aan onze kust. Voor zover dit verschijnsel in een der beschouwde lange stormseizoenen invloed van betekenis heeft gehad op HW is het verwerkt in de bestanden van HW en HW-opzet en in voorkomende gevallen ook in de geanalyseerde tijdreeksen. Het behoeft in verband met het onderhavige onderzoek verder geen bijzondere aandacht.

Het is niet a priori uitgesloten dat ook *onderzeese aardbevingen* oorzaak van waterstandsverhogingen aan onze kusten kunnen zijn. Een dergelijke beving kan namelijk een seismische zee golf - een zogenaamde *tsunami* - teweegbrengen mits ze voldoet aan een aantal voorwaarden (wat bij veruit de meeste onderzeese bevingen niet het geval is)*. Het staat vast dat als gevolg van zware aardbevingen in de seismische gordel Gibraltar - Azoren tsunami's zijn opgetreden aan de zuidkusten van Engeland en Ierland op 1 november 1755 (aardbeving van Lissabon) en op 31 maart 1761; er zijn echter geen aanwijzingen dat deze golven de Noordzee hebben bereikt [o.c., par. 6.2.2 resp. 6.2.3]. Secundaire breuksystemen van de aardschollen bevinden zich voor een deel onder achtereenvolgens het Kanaal en de Noordzee. In het Kanaal deed zich op 6 april 1580 een grote waterbeweging voor - hoogstwaarschijnlijk een tsunami - als gevolg van een onderzeese aardbeving aldaar, die naar mag worden aangenomen ook op de Noordzee gevoeld werd. Verder zijn er aanwijzingen dat de haard der aardbeving van 18 september 1692 zich deels onder de Noordzee uitstreckte en misschien daar een kleine tsunami heeft gegenereerd [o.c., par. 6.2.1]. Ten slotte deed zich op 7 juni 1931 een vrij krachtige onderzeese aardbeving voor bij Doggersbank, die echter geen tsunami tot gevolg had [o.c., par. 6.4.1].

Op grond van theoretische beschouwingen, ontleend aan de literatuur, is een voorzichtige schatting gemaakt van de mogelijke tsunami's aan onze kusten. Volgens deze schatting zou een tsunami met aan de kust een golfhoogte van 30 à 50 cm gemiddeld eens per 500 jaar kunnen optreden en een met een golfhoogte van 50 à 75 cm gemiddeld eens per 5000 jaar [o.c., par. 5.1 en 5.2].

De hierboven vermelde empirische gegevens bevestigen noch weerspreken deze schatting.

* deze voorwaarden zijn: een geheel of deels onderzeese en niet erg diepe haard, voldoende water boven de haard, voldoende grote „magnitude" (= sterkte) en een belangrijke verticale component van de beweging van de aardkorst [o.c., hoofdstuk 3 en par. 4.1]

Ondanks de onzekerheidsmarge in deze schatting mag worden aangenomen dat er in het Noordzeegebied geen correlatie bestaat tussen aardbevingen en stormvloeden [o.c., par. 7.1.4 ad c]. In het kader van het onderhavige statistische onderzoek naar extreem hoge getijstanden kan de eventuele invloed van onderzeese aardbevingen op getijwaterstanden aan onze kusten dan ook verder buiten beschouwing blijven.

9. Samenvatting, conclusies en aanbevelingen

9.1 Samenvatting

9.1.1 Probleemstelling en voorberekingen van het waarnemingsmateriaal

In haar eindverslag, verschenen in 1960, heeft de Deltacommissie voor elk van een groot aantal peilmeetstations in het Nederlandse kustgebied en langs de Westerschelde twee bijzondere hoogwaterstanden vastgesteld, te weten het zgn. basispeil (overschrijdingsfrequentie 10^{-4} p.j.) en het zgn. ontwerppeil. De ontwerppeilen dienden en dienen als uitgangspunten voor de verzwareing der zeekeringen en van hoogwaterkeringen in het tiggebied. De verschillen tussen de plaatselijke basis- en ontwerppeilen vloeiden voort uit enerzijds de verhogende uitwerking van voorgenomen afdammingen en anderzijds verlagingen vanwege economische afwegingen. Grondslag voor de vaststelling der basispeilen vormden

- een studie van het KNMI naar de depressiebanen van stormvloeden, die zich te Hellevoetsluis als zwaar hadden gemanifesteerd en
- een wiskundig-statistische analyse van hoge hoogwaterstanden te Hoek van Holland uit de periode 1888-1957 door het Mathematisch Centrum te Amsterdam.

In verband met de onzekerheden in de basispeilen, in het bijzonder die voor de westelijke Waddenzee, beval de Deltacommissie nog aan de studies naar hun hoogten onder verwerking van beschikbaar komende nieuwe gegevens voort te zetten.

Mede naar aanleiding van deze aanbeveling verschenen er in de jaren zeventig en aan het begin der jaren tachtig enkele oriënterende studies ter zake, zowel van Rijkswaterstaatsdiensten als van het KNMI. Met een breed en diepgaand onderzoek werd halverwege 1984 begonnen door het „Samenwerkingsverband Frequentielijnen Kust“, bestaande uit vertegenwoordigers van het Centrum voor Wiskunde en Informatica (het vroegere Mathematisch Centrum), het KNMI en de Directie Waterhuishouding en Waterbeweging (vanaf medio 1985: Dienst Getijdewateren) van de Rijkswaterstaat en prof. dr. L. de Haan van de Erasmus Universiteit Rotterdam. Het voorliggende rapport brengt van deze studie en haar resultaten verslag uit.

De probleemstelling is gegeven in de eerste drie hoofdstukken. Op de oriënterende inleiding (hoofdstuk 1) volgt daarbij eerst een beschrijving der voor het onderzoek bepalende verschijnselen, te weten het astronomische getij, verwekt door de aantrekkingskrachten van maan en zon, en de meteorologische omstandigheden, die waterstandsverhogingen van belang teweegbrengen. Ook de opgetreden en optredende zeespiegelrijzingen en mogelijke toekomstige veranderingen op hydrologisch en meteorologisch gebied komen in dit hoofdstuk 2 ter sprake. De eerste twee paragrafen van het aansluitende hoofdstuk 3 handelen over verschillende gebezigde statistische benaderingswijzen om extreem hoge getijwaterstanden te schatten, te weten de methoden van de Deltacommissie (gevolgd door enkele kritische kanttekeningen) en methoden toegepast in het Verenigd Koninkrijk en de Duitse Bondrepubliek.

Het onderzoek van het „Samenwerkingsverband" betrof (zie § 4.1) de navolgende vijf getijstations, waarvan de hoogwaterstanden (HW's) uit zo lang mogelijke perioden - hieronder vermeld - zijn verwerkt:
 Vlissingen :1 juli 1881...1985; Den Helder: 1 juni 1932...1985
 Hoek van Holland :1 aug. 1887...1985; Harlingen: 1 juni 1932...1985
 Delfzijl :1 maart 1881...1985;

De genoemde perioden beginnen met de aanvang van de ononderbroken relevante reeksen betrouwbare hoog- en laagwaterstanden aan de betrokken stations. Bepalend voor de relevantie is de herleidbaarheid van de waarnemingen naar de waterstaatkundige toestand van 1985; om die reden valt het begintijdstip voor de beide Waddenzeestations even na de afsluiting der Zuiderzee aangezien die afsluiting in de westelijke Waddenzee niet alleen het getij, maar ook de waterstandsverhogingen („opzetten") tijdens stormen aanmerkelijk heeft beïnvloed.

Voor elk der 5 stations zijn HW's met zeer kleine overschrijdingskansen als zgn. kwantilen berekend met behulp van wiskundig-statistische modellen. Daartoe moest per station het beschikbare gegevensmateriaal eerst onderworpen worden aan selecties ten einde autocorrelatie uit te sluiten en homogeniteit te verzekeren.

In beginsel stonden voor de vaststelling der gezochte hoge HW's, naar in § 3.3 nader is uiteengezet, twee wegen open:

- a de statistische verwerking van HW's;
- b de statistische verwerking van relevante opzetten, waarvoor alleen de hoogwateropzetten in aanmerking bleken te komen. Deze grootheid HW-opzet is het verschil tussen het opgetreden hoogwater HW en het bijbehorende astronomische hoogwater ongeacht enig tijdsverschil tussen beide.

Elk der twee benaderingswijzen verdient enige toelichting:

- ad a: een HW is in beginsel de som van een deterministische component, het astronomische hoogwater, en een stochastische, de HW-opzet. De eerste component is louter gevolg van de aantrekkingskrachten van zon en maan, waarvan de uitwerking ter plaatse wisselt volgens een ingewikkeld samenstel van cycli. De sterker schommelende HW-opzet, teweeggebracht door meteorologische factoren (wind, luchtdruk), bezit - na te zijn onderworpen aan toereikende selecties - een toevalskarakter. De gebruikte statistische theorieën zijn ontwikkeld voor rijen stochastische grootheden die niet een cyclisch karakter hebben, zodat hun toepassing op HW niet ideaal lijkt. Daar echter ieder samentreffen van een bepaald astronomisch hoogwater en een op toereikende wijze geselecteerde HW-opzet een toevalskarakter heeft is dit geen bezwaar.
- ad b: Na de vaststelling per station van een verdelingsfunctie voor HW-opzet dient die gekoppeld te worden aan de verdelingsfunctie van het astronomische hoogwater voor dat station. Voor een dergelijke koppeling is bij de meeste toegepaste methoden (nog?) geen theoretische basis.

De voorbereidingen, verricht per station, omvatten:

- 1) Herberekening van de astronomische hoog- resp. laagwaterstanden voor de genoemde perioden volgens de tot 1985 gebruikelijke culminatiemethode (§ 4.2). Deze herberekening (met opeenvolgende analyseperioden van 9 jaar)

was nodig omdat de berekeningen uit het verleden in onvoldoende mate consistent waren. Voor elk der stations zijn verder voor een periode van 93 jaar de astronomische hoogwaterstanden berekend en in de vorm van kansdichtheidgrafieken gepresenteerd (§ 4.6).

- 2) Berekening van HW-opzet als het verschil tussen opgetreden HW en bijbehorend herberekend astronomisch hoogwater voor iedere HW (§ 4.2).
- 3) Herleiding van HW's naar de toestand 1985 i.v.m. de opgetreden veranderingen (vooral rijzing) van de jaargemiddelden van HW (§ 4.5).
- 4) Onderzoek naar de mogelijke geschiktheid voor statistische verwerking van enige andere opzet (b.v. laagwateropzet) naast HW-opzet t.b.v. het onderzoek (§ 3.3). Uitslag: negatief.
- 5) Ter homogenisering van het waarnemingsmateriaal zijn de volgende selecties toegepast:
 - a beperking tot gegevens uit het zogenaamde lange stormseizoen 1 oktober...15 maart (vastgesteld m.b.v. statistische procedures voor Hoek van Holland en Delfzijl - zie § 4.3.3);
 - b invoering van „drempels”, die inhouden:
 - 1e: HW-opzet ≥ 30 cm, dit op grond van meteorologische en hydrologische overwegingen (§ 4.3.2). Deze drempel geldt ongeacht de keuze van de statistisch veranderlijke (HW-opzet resp. HW);
 - 2e: HW \geq hoogste astronomische hoogwater, vervat in de betrokken reeks, + 30 cm. Tot deze drempel is besloten voornamelijk op grond van een statistische gedachtengang (§ 6.2).
- 6) Klimatologische, meteorologische en hydrologische overwegingen hebben geleid tot de zgn. D-S-i („dunnings”)selectie ter verwijdering van autocorrelatie. Deze luidt: „Selecteer alle HW-opzetten ≥ 30 cm uit het stormseizoen, waarvoor binnen een radius van i HW-opzetten geen hogere HW-opzetten voorkomen.” Een statistische benadering voerde tot de slotsom dat $i = 4$ de meest geëigende waarde is voor i (§ 4.3.4 en - voor de wiskundig-statistische behandeling - aanhangsel I). Uiteraard bleven ook na deze selectie alleen die HW's voor verwerking beschikbaar, waarvan de HW-opzetten tot de geselecteerde waarden behoren.

Naast deze voorbereidingen zijn nog twee vooronderzoekingen verricht met behulp van statistische toetsen en wel naar

- de homogeniteit van het klimaat van HW-opzet in de beschouwde perioden (§ 4.3.5). Voor alle stations bleek dit klimaat inderdaad als homogeen te mogen worden aangenomen;
- een mogelijke samenhang tussen het astronomische hoogwater en HW-opzet (§ 4.4.3). Voor alle stations bleek tot afwezigheid van een dergelijke samenhang besloten te mogen worden.

9.1.2 De toegepaste methoden en hun uitkomsten

De theorie met betrekking tot de toegepaste methoden - een onderdeel van de theorie der extreme waarden uit de wiskundige statistiek - treft men aan in hoofdstuk 5; in § 6.1 is deze stof in een meer beschrijvende zin behandeld.

In die toegepaste methoden kunnen twee typen worden onderscheiden t.w. de verdelingsvrije en de verdelingsgebonden methoden.

Er zijn twee verdelingsvrije methoden (§ 5.2) gebruikt. Bij beide berekent men een zekere „krommingsgrootheid” (de zgn. extreme-waardenindex γ) voor de k grootste van de beschikbare waarden der beschouwde tijdreeks van HW. Uit de (k, γ) -grafiek (specifiek voor elk station) leest men γ af voor het grootste gebied van k , waarvoor γ constant is. Bij de eerste methode VVM-c (zie ook § 6.3) kan men vervolgens slechts een kleiner deel van deze k grootste waarden gebruiken voor de berekening van het gezochte kwantiel x_0 ; in de (k, x_0) -grafiek leest men vervolgens x_0 af op soortgelijke wijze als eerder γ . Bij de tweede methode VVM-0 (zie ook § 6.4) berekent men x_0 gelijktijdig met γ voor iedere k ; verder leest men de aangehouden waarde van x_0 op vergelijkbare wijze af. Beide methoden, die in samenhang met het onderhavige onderzoek zijn ontwikkeld, verschillen verder nog voor wat betreft de berekening der betrouwbaarheidsintervallen.

De methoden van het tweede type (§ 5.3) berusten daarentegen op een parametrisch model (wat een minpunt is t.o.v. verdelingsvrije methoden). Bij een dezer methoden beschouwt men alleen die beschikbare waarden X_i , die uitsteken boven een goed gekozen drempel u . Is u voldoende hoog dan voldoen de uitstekende waarden $y_i = (X_i - u)$ met goede benadering aan de zgn. gegeneraliseerde Pareto-verdeling („GPV”), die gekenmerkt wordt door een vormparameter γ (= de extreme-waardenindex) en een schaalparameter $\sigma(u)$. In dit geval kunnen de parameters en y_0 via de methode van de grootste aannemelijkheid worden berekend voor verschillende u . Daar k wordt bepaald door de waarde van u kan y_0 weer in een grafiek tegen k worden uitgezet. De aangehouden waarde van de y_0 van het kwantiel is afgeleid uit een (k, y_0) -grafiek; de betrouwbaarheidsintervallen worden in deze procedure gelijk meeberekend en bepaald (zie verder § 6.5).

Bij de tweede parametrische methode worden alleen de maxima der opeenvolgende stormseizoenen gebruikt, waaraan een extreme-waardenverdeling („GEV”) met (onbekende) locatie μ en vorm- en schaalparameters γ en σ wordt aangepast (zie ook § 6.7). Voor de schattingen van μ , γ en σ bezigt men weer de methode der grootste aannemelijkheid. Op basis van een schatting voor de variantie van het berekende kwantiel kan vervolgens een benaderend betrouwbaarheidsinterval worden bepaald.

M.b.v. de geschetste methoden zijn nu voor ieder station vier berekeningen gemaakt met HW als statistisch veranderlijke. In het vijfde geval (de zgn. CON-methode, zie § 6.6) werd de GPV-methode toegepast op HW-opzet met als voordeel: meer beschikbare waarden (de beperking van een drempel voor HW vervalt immers). De gevonden verdeling moest vervolgens gekoppeld worden aan de verdeling van de astronomische hoogwaterstand gedurende het lange stormseizoen aan het betrokken station (convolutie); het betrouwbaarheidsinterval is bepaald door een simulatietechniek („Monte-Carlo-procedure”).

De uitkomsten volgens de 5 genoemde methoden, verzameld in tabel 9.1 (= tabel 6.7), zijn vervolgens beoordeeld aan de hand van enkele maatstaven (zie ook § 7.2): VVM-c is op theoretische gronden in het onderhavige geval minder bruikbaar; bovendien zijn de

verhoudingsgewijs kleine) betrouwbaarheidsintervallen weinig realistisch, waardoor deze methode relatief laag wordt gewaardeerd. VVM-0 is theoretisch het meest aantrekkelijk. Bij de GPV bleek de keuze van y_0 resp. x_0 relatief het meest subjectief. Aan de toepassing van GEV-s is het bezwaar verbonden dat een groot deel van het waarnemingsmateriaal buiten beschouwing blijft. Een zwak punt bij CON is het gebruik van GPV; de aantrekkelijk lijkende toepassing van VVM-0 i.p.v. GPV op HW-opzet stuit af op het feit dat koppeling met de verdeling van het astronomisch hoogwater praktisch niet mogelijk is. Voor de goede orde zij er verder op gewezen dat aan de breedte van het berekende betrouwbaarheidsinterval geen beslissende waarde mag worden toegekend voor de keuze van de kwantielen.

Tabel 9.1:

Schattingen van het 10^{-4} -kwantiel (cm t.o.v. NAP) en de bijbehorende 95%-betrouwbaarheidsintervallen (cm) voor de toestand 1985

	Vlissingen	Hoek van H.	Den Helder	Harlingen	Delfzijl
methode	1881/82... 1984/85	1887/88... 1984/85	1932/33... 1984/85	1932/33... 1984/85	1881/82... 1984/85
VVM-c	540 ± 50	460 ± 40	360 ± 30	430 ± 30	600 ± 40
VVM-0	540 ± 140	500 ± 190	370 ± 100	405 ± 80	600 ± 150
GPV	555 ± 140	475 ± 125	415 ± 155	440 ± 75	615 ± 150
CON	540 ± 65	460 ± 80	390 ± 70	445 ± 45	660 ± 120
GEV-s	585 ± 180	510 ± 180	390 ± 110	420 ± 65	650 ± 195
bp	565	500	505	580	640
hs (herleid)	466	402	331	379	480
hs ₀	455(1953)	385(1953)	325(1953)	369(1954, 1976)	453(1901)

- bp - huidige basispeil
 hs (herleid) - hoogste stand van het gebruikte, naar 1985 herleide HW-bestand
 hs₀ - hoogste waargenomen stand uit de beschouwde periode, met jaar (de hoogste bekende stand van Delfzijl is NAP + 460 cm uit 1825, te stellen op NAP + 500 cm voor de toestand 1985-zie § 8.1)

Op grond van bovenstaande overwegingen kiezen de betrokken onderzoekers de kwantielwaarden volgens de methode VVM-0 als beste schattingen.

N.B.: Voor een goede vergelijking van de huidige basispeilen (bp) met de geschatte waarden in tabel 9.1 dient men in aanmerking te nemen dat die basispeilen daartoe in feite eveneens naar de toestand 1985 zouden moeten worden herleid (lees: verhoogd!).

9.1.3 Evaluatie en afsluitende opmerkingen

Ten einde de uitkomsten der stations onderling te kunnen vergelijken is de methode VVM-0 ook nog toegepast op de stations

- Ijmuiden, waarnemingsperiode 1883-1985 (§7.10);
- Vlissingen, Hoek van Holland, Ijmuiden en Delfzijl, „korte” waarnemingsperiode mei 1932-1985 (§§ 7.9 en 7.10);
- West-Terschelling, waarnemingsperiode mei 1932-1985 (§ 7.10).

De uitkomsten voor het 10^{-4} -kwantiel treft men aan in tabel 9.2, waaruit duidelijk blijkt dat de resultaten voor de reeksen van 53 jaar

aanmerkelijk lager zijn dan voor de zoveel langere reeksen, beginnend in de 19e eeuw. Ervan uitgaande dat er sprake is van samenhang tussen zeer hoge hoogwaterstanden aan de stations (zie § 7.9) ligt de slotsom voor de hand dat de geschatte waarden voor de Waddenzee stations Den Helder, Harlingen en West-Terschelling waarschijnlijk een 50 à 60 cm lager zijn dan gevonden zou zijn indien ook voor deze stations overeenkomstige lange reeksen beschikbaar geweest zouden zijn.

Tabel 9.2:

Schattingen van het 10^{-4} -kwantiel (VVM-0-methode) met betrouwbaarheidsintervallen. Perioden: lang (~ 100 jr.) en kort (mei 1932-1985), stormseizoen: 1 okt. . . . 15 mrt.

periode	Vlissingen	Hoek van Holl.	IJmuiden	Den Helder	Harlingen	West-Tersch.	Delfzijl
lang	540 ± 140	500 ± 190	530 ± 210	*	*	*	600 ± 150
Kort	520 ± 170	460 ± 200	470 ± 240	370 ± 100	405 ± 80	350 ± 60	545 ± 165

*: aanbeveling hiervoor in § 9.2 (ad 8e)

Vergelijkbaar met de verschillen tussen de schattingen voor die lange en korte perioden zijn de verschillen, die zijn gevonden voor het station Vlissingen tussen enerzijds de HW's uit alleen de even jaren der lange periode en anderzijds die uit alleen de oneven jaren (§ 7.4).

Als afsluiting van de studie zijn nog een aantal aanvullende beschouwingen gegeven en onderzoeken verricht, waarvan hier vermelding verdienen:

- statistische technieken, die gebleken zijn ten behoeve van het onderhavige vraagstuk niet voor toepassing in aanmerking te komen (§ 7.6);
- schattingen m.b.v stormjaarmaxima (§ 7.4) en schattingen volgens het statistische model gebruikt voor Hoek van Holland in het Deltarapport (§ 7.7). Deze exercities gaven geen aanleiding tot enige relativering van de eerder genoemde schattingen;
- gevoeligheidsonderzoeken m.b.t.
 - i In de selectie D-S-i (§ 7.5) en
 - de drempel u 30 cm voor HW-opzet;
- beschouwingen over de betrouwbaarheidsintervallen (§ 7.8);
- bepaling der overschrijdingslijnen voor extreem hoge stormvloedstanden voor de verschillende stations en hun onderlinge vergelijking (§ 7.8);
- onderzoek naar de waterhoogten bereikt bij beruchte stormvloeden uit het verleden, met name die van 1 nov. 1570, 14 op 15 nov. 1775 en 4 op 5 februari 1825 (§ 8.1). Het is aannemelijk dat de eerste tot - waar schijnlijk zeer kortdurende - overschrijding van het peil van 1953 heeft geleid langs de gesloten Hollandse kust en dat de hoogste bekende standen in Noordoost Nederland (van 1825) in de daaraan voorafgaande 2¾ eeuw enkele malen ten naastebij zijn geëvenaard of enigszins overschreden;
- andere oorzaken voor hoge getijwaterstanden dan stormen aan onze kust, te weten zeeberen en tsunami's (§ 8.2). Er blijkt geen

aanleiding te zijn om aan deze verschijnselen afzonderlijke aandacht te wijden in het raam van het beheer van en het beleid betreffende zeeweringen en hoogwaterkeringen in het Nederlandse getijgebied.

9.2 Gevolgtrekkingen en aanbevelingen

Het onderzoek heeft geleid tot de volgende conclusies:

- 1e Voor de bepaling van getijhoogwaterstanden van zeer lage overschrijdingsfrequenties door middel van statistische analyses lenen zich zowel de hoogwaterstand (HW) als de door weersinvloeden opgetreden verhoging van de hoogwaterstand (HW-opzet). De bruikbaarheid van HW-opzet wordt echter sterk beperkt door moeilijkheden bij de dan noodzakelijke koppeling van zijn kansverdeling aan de verdeling van de astronomische hoogwaterstanden.
- 2e Om onafhankelijke en gelijkverdeelde waarden te verkrijgen, nodig voor de toegepaste statistische analyses dient het waarnemingsmateriaal opeenvolgend aan onderstaande drie selecties te worden onderworpen:
 - beperking tot het zgn. lange stormseizoen 1 oktober...15 maart;
 - beperking tot de gevallen met HW-opzet ≥ 30 cm;
 - beperking tot de gevallen met een HW-opzet, die groter is dan die van de 4 onmiddellijk voorafgaande en van de 4 onmiddellijk volgende HW opzetten (selectie D-S-4).Bij verwerking van HW-opzet zijn deze bewerkingen voldoende. Voor de waarden van HW gelden echter bovendien de eisen dat zij hoger moeten zijn dan de hoogste astronomische hoogwaterstand uit alle beschouwde stormseizoenen plus 30 cm en dat zij herleid moeten worden naar het laatste lange stormseizoen in verband met de rijzing van de jaargemiddelden van HW.
- 3e Er is noch statistisch noch anderszins ook maar enige aanwijzing dat aan onze kusten het klimaat van HW-opzet in het lange stormseizoen tijdens de periode 1880-1985 op enigerlei wijze aan een trend onderhevig is ge weest. Wel is dit „HW-opzet-klimaat“ voor de stations aan de westelijke Waddenzee in de loop van enige jaren gewijzigd als gevolg van de bouw van de Afsluitdijk (dichting laatste sluitgat: mei 1932). Dit kunstwerk veranderde ook en in korter tijd de hoogten van de astronomische hoog en laagwaterstanden in dit gebied.
- 4e Van alle beschouwde methoden levert de verdelingsvrije methode VVM-0 , toegepast op HW, thans de beste schattingen voor de gezochte kwantielen.
- 5e Bij alle analyses is voor de selectie D-S-i toegepast de meest geëigende waarde $i = 4$, maar gebleken is dat de bedoelde schattingen voor $i = 3$ (dus met meer waarnemingen) weinig afwijken van die voor $i = 4$. Evenmin leidt een (drempel)keuze beperking tot HW-opzet ≥ 50 cm - dus met minder verwerkbare waarnemingen! - in plaats van HW-opzet ≥ 30 cm tot duidelijk afwijkende schattingen voor de gezochte kwantielen.
- 6e Gebruikt men bij toepassing der methode VVM-0 van de beschikbare tijdreeksen van HW, die 100 jaar of meer

omspannen, ongeveer de helft van het beschikbare materiaal dan blijken de uitkomsten aanmerkelijk te verschillen van de schattingen uit de tijdreeksen over de volle periode van omstreeks 100 jaar (de gehalveerde reeksen zijn daarbij bepaald door alleen de even resp. alleen de oneven jaren en door de (gehele) periode mei 1932...1985 te beschouwen - zie tabel 9.2).

- 7e Voor de stations met omstreeks 100 jaar lange reeksen verwerkbare waarnemingen zijn de beste schattingen voor de 10^{-4} -kwantielen, toestand 1985 - met tussen haakjes hun 95%-betrouwbaarheidsintervallen:

Vlissingen	NAP + 540 (± 140) cm;
Hoek van Holland	NAP + 500 (± 190) cm;
Ijmuiden	NAP + 530 (± 210) cm;
Delfzijl	NAP + 600 (± 150) cm.

Deze waarden zijn bepaald met de methode VVM-0 onder toepassing van de „normale” benadering voor de betrouwbaarheidsintervallen.

- 8e Voor de stations aan de westelijke Waddenzee beslaan de voorhanden bruikbare reeksen ruim 50 jaar. Om de hiermee mogelijke schattingen voor de 10^{-4} -kwantielen vergelijkbaar te maken met die voor de andere stations zijn zij „aangepast”, d.w.z. verhoogd met 55 cm - dit naar analogie van de bevindingen ter zake (zie ad 6e) voor die andere stations. *Bij deze „aangepaste schattingen” kunnen geen betrouwbaarheidsintervallen worden berekend.* Nu zijn voor de andere 4 stations de betrouwbaarheidsintervallen voor de reeksen vanaf mei 1932 niet zeer veel groter dan voor hun volledige reeksen (voor Vlissingen, Hoek van Holland, Ijmuiden en Delfzijl achtereenvolgens 20%, 5%, 14% en 10%). Gelet hierop zou men als werkhypothese voor de „aangepaste schattingen” de betrouwbaarheidsintervallen voor de korte periode kunnen aanhouden; daarmee worden de beste schattingen voor de bedoelde stations:

Den Helder	NAP + 425 (± 100) cm;
Harlingen	NAP + 460 (± 80) cm;
West-Terschelling	NAP + 405 (± 60) cm.

De bevindingen naar aanleiding van het onderzoek geven aanleiding tot het doen van onderstaande *aanbevelingen*:

- A** De resultaten van het ingestelde statistische onderzoek dienen geïnterpreteerd te worden tegen de fysische achtergrond van het verschijnsel van buitengewoon hoge waterstanden aan onze kusten. Daarbij kunnen uitkomsten van fysische rekenmodellen van de Noordzee en van de Waddenzee van groot nut zijn.
- B** Het is gewenst dat de onderhavige studie wordt voortgezet met de verwerking van de samenhang, die bestaat tussen de gegevens van verschillende stations, een en ander volgens kortelings hiervoor ontwikkelde theorieën. De grote hoeveelheid werk, die hiermee gemoeid is, zou men voor lief dienen te nemen.
- C** Zodra een theorie ter bepaling van de betrouwbaarheidsintervallen volgens de „bootstrap”-methode (§ 7.3) zal zijn ontwikkeld voor de VVM-0 methode dient zij te worden toegepast op de gebruikte reeksen, dit ter verfijning van de gegeven betrouwbaarheidsintervallen.

- D Gezien de gevoeligheid van de kwantielschattingen voor de halvering van het waarnemingsmateriaal (zie de 6e conclusie) is herhaling van het onderzoek nodig zodra de reeks waarnemingen in voldoende mate (met zeg 15 of 20 jaar) is verlengd, Van dit onderzoek dienen in elk geval ook de dan beschikbare statistische technieken voor ruimtelijke samenhang (zie onder B) deel uit te maken.
- E Het is - o.a. vanuit een oogpunt van gemeenschappelijke belangen - gewenst over de ontwikkelingen met betrekking tot de wijzen van bepaling van uiterst hoge getijstanden nauw contact te onderhouden met de hierbij betrokken diensten in de andere oeverstaten van de Noordzee.

10. Literatuurlijst

- P. Ackers en T.D. Ruxton (1975):*
„Extreme levels arising from meteorological surges". Proceedings 14th Coastal Engineering Conference Copenhagen 1974, Vol.1, bl. 69...86.
- A.C. Bijlsma (1989):*
„Investigation of surge-tide interaction in the storm surge model CSM-16". Delft Hydraulics/Waterloopkundig Laboratorium Delft, rapport Z311.
- J.M. Chambers, W.S. Cleveland, B. Kleiner en P.A. Tukey (1983):*
„Graphical methods for Data Analysis". Wadsworth.
- D. van Dantzig, J. Hemelrijk e.a. (1960):*
„Extrapolatie van de overschrijdingslijn van de hoogwaterstanden te Hoek van Holland met behulp van geselecteerde stormen". Rapport Deltacommissie, deel 3, bijdrage 11.1.
- A.L.M. Dekkers en L. de Haan (1989):*
„On the estimation of the extreme-value index and large quantile estimation". The Annals of Statistics, 1989, 17, 4, bl. 1795-1832.
- A.L.M. Dekkers, H.J.J. Einmahl en L. de Haan (1989):*
„A moment estimator for the index of an extreme-value distribution". The Annals of Statistics, 1989, 17, 4, bl. 1833-1855.
- Deltacommissie (1960):*
„Eindverslag en interimadviezen". Rapport Deltacommissie, deel 1 (= RDC 1)
„Deltarapport":
zie noot na "Gebruikte afkortingen".
- P.J. Diggle (1983):*
„Statistical Analysis of spatial point patterns". Academic Press.
- J. Doekes (1985):*
„Invloed van de afsluiting van de Zuiderzee op het getij in de Waddenzee". Rijkswaterstaat, Dienst Getijdewateren, nota GWIO-85.001.
- J. Doekes (1986):*
„Zijn HW-opzet en astronomisch hoogwater onafhankelijk?"
Rijkswaterstaat, Dienst Getijdewateren, nota GWIO-86.013.
- J.J. Dronkers (1955):*
„De invloed van de afsluiting van de Zuiderzee op de verhoging van de stormvloedstanden te Den Helder, Terschelling en Harlingen". Rijkswaterstaat, Centrale Studiedienst, nota CSD 55-20.
- A. Führböter (1978):*
„Frequencies and probabilities of extreme storm surges". Proceedings 16th Coastal Engineering Conference, Hamburg 1978, Vol.1, bl. 949...964.
- Gezondheidsraad (1983):*
„Deeladvies inzake CO₂-problematiek". Den Haag.
- M.K.E. Gottschalk (1975 resp. 1977):*
„Stormvloed en rivieroverstromingen in Nederland". Deel 2 resp. 3. Van Gorcum, Assen.

- J. Graff (1981):*
„An investigation of the frequency distributions of annual sea level maxima at ports around Great Britain". *Estuarine, Coastal and Shelf Science* 1981, nr. 12, bl. 389-449.
- P.M.E.M. van der Grinten en J.M.H. de Lenoir (1973):*
„Statistische procesbeheersing". Uitg. Het Spectrum.
- P. Groen (1961):*
„De wateren van de wereldzee". Tweede druk, C. de Boer jr., Hilversum.
- L. de Haan (1985):*
„Extremes in higher dimensions: the model and some statistics". *Proceedings ISI-meeting Amsterdam*.
- L. de Haan (1986):*
„The Arch-enemy Attacked Mathematically". *CWI Monograph 4, Mathematics and Computer Science II* (eds. M. Hazewinkel, J.K. Lenstra & L.G.L.T. Meertens), bl. 51-60.
- L. de Haan (1990):*
„Fighting the arch-enemy with mathematics". *Statistica Neerlandica*, 1990, 44, 2, bl. 45-68.
- L. de Haan en H. Rootzén (1991):*
„On the estimation of high quantiles". *Verschijnt in: Journal of Statistical Planning and Inference*.
- L. de Haan en S. Resnick (1991):*
„Consistent empirical estimators of multivariate extreme value distributions". *Verschijnt in: Statistical Models*.
- J. Hájek (1969):*
„A Course in Nonparametric Statistics". Holden Day, San Francisco.
- C.J. van der Ham (1960):*
„Onderzoek naar de bij stormvloed behorende depressiebanen en naar de extreme mogelijkheden van noordwesterstormen". *RDC 2*, bijdrage 1-2.4.
- R. Helmers (1984):*
„Jackknife en bootstrap methoden. Vacantiecursus 1984 Hewet - plus wiskunde". *CWI syllabus 1*, Amsterdam.
- IPCC - ed.: J.T. Houghton, G.J. Jenkins, J.J. Ephraem (1990):*
„Climatic Change". *The IPCC Scientific Assesment*. Cambridge. IPCC = Intergovernmental Panel on Climatic Change.
- J. Jensen (1985):*
„Über instationäre Entwicklungen der Wasserstände an der deutschen Nordseeküste". *Mitteilungen Leichtweiss-Institut für Wasserbau, Braunschweig*, Heft 88/1985, bl.141...319 (Proefschrift TU Braunschweig).
- A.R.T. Jonkers (1989):*
„Over den schrikkelijken watervloed". Een onderzoek naar hoge waterstanden vóór 1825. Stageverslag DGW.
- S. Kruizinga (1978):*
„Objectieve classificatie van dagelijkse 500 mbar-patronen". *KNMI, De Bilt, Wetenschappelijk Rapport WR-78.8*.
- M.R. Leadbetter (1983):*
„Extremes and Local Dependence in Stationary Sequences". *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 65, bl. 291-306.
- M.R. Leadbetter, G. Lindgren, H. Rootzen (1983):*
„Extremes and Related Properties of Random Sequences and Processes". *Springer Series in Statistics*.

- G.K. Lorenz, W. Groenewoud, F. Schokking, M.W. v.d. Berg, J. Wiersma, F.J.J. Brouwer, S. Jelgersma (1991):*
„Heden en Verleden. Nederland naar beneden??? Interim-rapport Onderzoek bodembeweging in Nederland". Rijkswaterstaat, Rijks Geologische Dienst; Delft/Haarlem/Rijswijk.
- J.W. van der Made (1987):*
„Analysis of some criteria for design and operation of surface water gauging networks". Academisch Proefschrift Technische Universiteit Delft.
- J.W. van der Made (red.) (1989):*
„Frequentielijnen Kust. Tussentijds verslag 2". RWS, DGW, nota GWIO-89.001.
- J. van Malde (1984):*
„Voorlopige uitkomsten van voortgezet onderzoek naar de gemiddelde zeeniveaus in de Nederlandse kustwateren". RWS, dir. W&W, nota WW-WH 84.08.
- J. van Malde (1990):*
„Onderzoek naar de mogelijke beïnvloeding van Nederlandse getijwaterstanden door aardbevingen en naar historische "moerzeeën". RWS, DGW, nota GWAO 90.009.
- J. van Malde (1992):*
„Relative rise of Mean Sea-Levels in the Netherlands in recent times". In: M.J. Tooley en S. Jelgersma: Impact of Sea-Level Rise on European Coastal Lowlands. Blackwell, Oxford (V.K.) en Cambridge (V.St.).
- McGill, Tukey en Larsen (1978):*
„Variations of Box Plots". The American Statistician, Vol. 32, no.1, bl. 12-16.
- J. Pickands (1975):*
„Statistical inference using extreme order statistics". Annals of Statistics, Vol. 3, no.1, bl. 119-131.
- D.T. Pugh en J.M. Vassie (1979):*
„Extreme sea levels from tide and surge probability". Proceedings 16th Coastal Engineering Conference, Hamburg 1978, Vol.1, bl. 911-930.
- D.T. Pugh (1987):*
„Tides, surges and mean sea-level". John Wiley & Sons, Chichester etc.
- P.J. Rijkooort (1960):*
„De frequentieverdeling van de tijdsintervallen tussen opeenvolgende noordwesterstormen en de onderlinge afhankelijkheid van deze stormen". RDC 2, bijdrage 1-2.7.
- Rijkswaterstaat en KNMI (1961):*
„Verslag over de stormvloed van 1953". Staatsdrukkerij- en uitgeverijbedrijf.
- H. Rohde (1977):*
Sturmfluthöhen und säkularer Wasserstandsanstieg an der deutschen Nordseeküste. Die Küste, Heft 30.
- J.G. de Ronde (1984^a):*
„Grenspeilen 1981.0". RWS, dir. W&W, nota WW-WH-82.12.
- J.G. de Ronde (1984^b):*
„Correcties t.b.v. de „Getijtafels voor Nederland 1985". RWS, dir. W&W, nota WW-WH 84.05.
- J.G. de Ronde (1985):*
„Wisselwerking tussen opzet en verticaal getij". RWS, DGW, nota GWIO 85.003.

Samenwerkingsverband Basispeilen (1986):

„Onderzoek Basispeilen. Tussentijds verslag juni 1986". RWS, DGW, nota GWIO-86.008.

P. Schotel (1922):

„Historisch overzicht van de hooge vloed en overstromingen tot het jaar 1868". Gebr. J. & H. van Langenhuysen, 's-Gravenhage.

R.L. Smith (1984):

„Threshold Methods for Sample extremes" in: Statistical Extremes and Applications (editor J. Tiago de Oliveira), bl. 621-638. Reidel, Dordrecht.

R.L. Smith (1987):

„Estimating tails of probability distribution". The Annals of Statistics 15, bl. 1174-1207.

J.A. Tawn (1988):

„An extreme-value theory model for dependent observations". Journal of Hydrology 1988, bl. 227 ... 250.

H. Tennekes en G.P. Können (red.) (1990):

„Aanhoudend warm. Klimaatvoorspellingen vanuit De Bilt". KNMI, Thieme Baarn.

H. Timmerman (1977):

„Meteorological Effects on Tidal Heights in the North Sea". Mededelingen en Verhandelingen no. 99, hoofdstuk 10. KNMI. TO 1951-1960. TO 1961-1970. TO 1971-1980.

(Zie onder "Gebruikte afkortingen").

A. van Urk en J. van Malde (1977):

„Onderzoek basispeilen. Gevolgtrekkingen en aanbevelingen". RWS, dir. W&W, nota WH-77.21.

A. van Urk en J.G. de Ronde (1982):

„Getijtafels voor Nederland vanaf 1980". Rijkswaterstaat-serie nr. 37.

Joh. van Veen (1945):

„Bestaat er een geologische bodemdaling te Amsterdam sedert 1700?" Tijdschrift Koninklijk Nederlandsch Aardrijkskundig Genootschap, 2e Reeks, deel LXII.

J.A. Vogel (1992):

„Statistiek van seiches". RWS, DGW, nota GWAO-91.012.

M.P.H. Weenink (1960^a):

„De stormvloed van 21 t/m 24 december 1954. Een geval van resonantie". RDC 2, bijdrage I-3.2.

M.P.H. Weenink (1960^b):

„Berekening van de windeffecten veroorzaakt door de verplaatst gedachte storm van 31 januari-1 februari 1953". RDC 2, bijdrage I-3.4.

P.J. Wemelsfelder (1960^a):

„De overschrijdingslijnen van de hoogwaterstanden in het Nederlandse getijgebied". RDC 4, bijdrage III.2.

P.J. Wemelsfelder (1960^b):

„Bui-oscillaties en buistoten tijdens stormvloed". RDC 4, bijdrage III-5.

Gebruikte afkortingen:

- CWI = Centrum voor Wiskunde en Informatica.
DGW = Dienst Getijdewateren van de Rijkswaterstaat.
ISI = International Statistical Institute.
RDC = Rapport Deltacommissie*, 5 delen. Staatsdrukkerij- en Uitgeverijbedrijf, 's-Gravenhage. RDC 1 = deel 1 enz..
RWS = Rijkswaterstaat.
TO = Tienjarig Overzicht der Waterhoogten, Afvoeren en Watertemperaturen.
TO 1961-1970: uitgave RWS, dir. W&W (1985).
TO 1971-1980: uitgave RWS, DGW (1989).
dir. W&W = voormalige Directie Waterhuishouding en Waterbeweging RWS.

* In de tekst van het voorliggende rapport wordt het Rapport Deltacommissie aangeduid als „Deltarapport“.

N.B.: De in de tekst gebezigde afkortingen a.w. (= aangehaald werk) en o.c. (= opere citato) doelen op de laatste eraan voorafgaande literatuurverwijzing.

Aanhangsel 1:

Clustering van hoge HW-opzetten in het korte stormseizoen

Afhankelijkheid van HW-opzetten (en dus ook van de bijbehorende HW-standen) voor een gegeven station leidt in de regel tot *clustering* van HW-opzetten, welke in de tijd dicht bij elkaar liggen en een voldoende hoge drempelwaarde overschrijden. Bekend is dat in het geval van onafhankelijkheid het puntproces van de stochastische tijdstippen waarop HW-opzetten - welke een voldoende hoge drempelwaarde overschrijden - optreden, benaderd kan worden door een *Poisson-proces*. In het bijzonder betekent dit dat géén clustering zal optreden van 'hoge' HW-opzetten. Ook de kansverdeling van het genormeerde maximum (de normeringsconstanten a_n en b_n zijn dezelfde als in het geval van onafhankelijkheid - zie (§ 5.2) ziet er voor afhankelijke waarnemingen anders uit: voor grote waarden van n geldt (5.7) met G_γ , de extreme-waardenverdeling (5.8), nu vervangen door G_γ^θ , waarbij θ Leadbetter's *extremaal-index* ($0 \leq \theta \leq 1$) voorstelt.

Leadbetter (1983) heeft laten zien dat voor *afhankelijke* waarnemingen inderdaad (5.7) geldt met G_γ vervangen door G_γ^θ , mits X_1, X_2, \dots (in ons onderzoek: de HW-opzetten aan een gegeven station) een stationaire rij vormen, met de eigenschap dat X_i 's die in de tijd ver uit elkaar liggen, vrijwel onafhankelijk zijn (preciezer gezegd: voorwaarde (D) van Leadbetter moet gelden; deze voorwaarde is zeer zwak en lijkt in ons onderzoek géén probleem) en bovendien aan een kleine extra aanname is voldaan, namelijk dat de stationaire rij X_1, X_2, \dots een *extremaal-index* θ heeft (preciezer gezegd: indien $\{u_n\}$ een rij constanten is met de eigenschap dat $nP(X_1 > u_n) \approx \tau$, voor een constante $0 < \tau < \infty$ en voldoende grote waarden van n , dan geldt $P(\max_{1 \leq i \leq n} X_i \leq u_n) \approx e^{-\theta\tau}$, voor voldoende grote waarden van n).

Anders gezegd: Indien $\tilde{X}_1, \tilde{X}_2, \dots$ fictieve, onderling onafhankelijke stochastische grootheden voorstellen, met als gemeenschappelijke verdelingsfunctie, de (marginale) verdelingsfunctie van de afzonderlijke X_i 's in de stationaire rij X_1, X_2, \dots , dan geldt, voor n voldoende groot,

$$P\left(\frac{\max(X_1, \dots, X_n) - b_n}{a_n} \leq x\right) \approx \left(P\left(\frac{\max(\tilde{X}_1, \dots, \tilde{X}_n) - b_n}{a_n} \leq x\right)\right)^\theta$$

De parameter θ - de extremaal-index - kan geïnterpreteerd worden als de reciproke van de *gemiddelde clusteromvang* van het puntproces van de 'hoge' HW-opzetten. In het klassieke geval van onafhankelijkheid is $\theta = 1$. Een intuïtief plausibele maat voor het optreden van clustering van hoge HW-opzetten wordt gegeven door (zie Diggle(1983), par 2.2, voor een soortgelijke benadering),

$$V_r(u) = \frac{\#\{(i,j): i \neq j, |i-j| \leq r, Y_i = Y_j = 1\}}{\#\{i: Y_i = 1\}} \quad (\text{A1.1})$$

waarbij $Y_i = 1$ dan en slechts dan als $X_i > u$ (en $Y_i = 0$ anders). M.a.w.: $V_r(u)$ is het quotiënt van het aantal paren hoge HW-opzetten, welke hoogstens op een afstand van r tijdseenheden van elkaar optreden en het totaal aantal hoge HW-opzetten.
Men gaat eenvoudig na dat

$$E[\#\{(i,j): i \neq j, |i-j| \leq r, Y_i = Y_j = 1\}] = 2n \sum_{j=2}^{r+1} P(Y_1 = Y_j = 1) \quad (A1.2)$$

en

$$E[\#\{i: Y_i = 1\}] = nP(Y_1 = 1) \quad (A1.3)$$

zodat we mogen verwachten dat

$$EV_r(u) \approx \frac{2 \sum_{j=2}^{r+1} P(Y_1 = Y_j = 1)}{P(Y_1 = 1)} \quad (A1.4)$$

In het klassieke geval van *onafhankelijkheid* reduceert (A1.4) tot:

$$EV_r(u) = \frac{2rP(Y_1 = 1)P(Y_2 = 1)}{P(Y_1 = 1)} = 2rP(Y_1 = 1). \quad (A1.5)$$

Indien we definiëren

$$V_r^*(u_n) = nV_r(u_n) \quad (A1.6)$$

en de drempel u_n zo gekozen wordt dat $P(X_1 > u_n) = \frac{\tau}{n}$, voor zekere constante $\tau > 0$, dan vinden we in het geval van onafhankelijkheid voor de genormeerde grootheid $V_r^*(u_n)$ de benadering

$$EV_r^*(u_n) \approx 2r\tau \quad (A1.7)$$

In het algemene geval dat het optreden van hoge HW-opzetten positief gecorreleerd is, vinden we, voor grote waarden van n , de ongelijkheid

$$EV_r^*(u_n) \approx \frac{2n^2}{\tau} \sum_{j=2}^{r+1} P(Y_1 = Y_j = 1) \geq \frac{2n^2}{\tau} \sum_{j=2}^{r+1} P(Y_1 = 1)P(Y_j = 1) = 2r\tau \quad (A1.8)$$

d.w.z. $EV_r^*(u_n)$ zal, in het geval van *positieve afhankelijkheid*, groter zijn dan $2r\tau$.

Een eenvoudige berekening leert verder dat in het geval van 'afhankelijkheid' - d.w.z. (hoge) HW-opzetten welke meer dan m tijdseenheden van elkaar optreden zijn onafhankelijk - geldt

$$EV_r^*(u_n) = \begin{cases} 2r\tau + 2n \sum_{i=2}^{r+1} \rho_i & \text{voor } r \leq m \\ 2r\tau + 2n \sum_{i=2}^{m+1} \rho_i & \text{voor } r > m \end{cases} \quad (A1.9)$$

waarbij $\rho_i = \rho_i(n)$ de correlatiecoëfficiënt tussen Y_1 en Y_i voorstelt. Voor grote waarden van n vinden we de benadering:

$$\rho_i \approx \frac{n}{\tau} P(Y_1 = Y_i = 1) - \frac{\tau}{n} \quad (\text{A1.10})$$

Ook leidt men eenvoudig de volgende ongelijkheid voor de extremaalindex θ van een m -afhankelijke stationaire rij af:

$$1 - \tau^{-1} \lim_{n \rightarrow \infty} \sum_{i=2}^{m+1} P(Y_1 = Y_i = 1) = 1 - \lim_{n \rightarrow \infty} \sum_{i=2}^{m+1} \rho_i(n) \leq \theta \leq 1 \quad (\text{A1.11})$$

Merk op dat $P(Y_1 = Y_i = 1) = P(X_1 > u_n, X_i > u_n)$ waarbij $u_n = u_n(\tau)$, zodanig dat $P(X_1 > u_n) = \frac{\tau}{n}$. In het geval van onafhankelijkheid reduceert (A1.11) tot het gewenste resultaat: $\theta = 1$. Bij m -afhankelijkheid kan de benedengrens in (A1.11) echter in principe iedere waarde van het interval $[0, 1]$ aannemen.

In dit onderzoek hebben we te maken met een stochastisch aantal N_i , het aantal geselecteerde HW-opzetten in het i^{de} ($i = 1, \dots, s$) korte stormseizoen, die boven een geschikt gekozen drempel u uitsteken. We vervangen n door N_i en berekenen de bijbehorende functie $V_r^*(u)$ voor het i^{de} korte stormseizoen. Het gemiddelde van de $V_r^*(u)$'s over de s beschikbare stormseizoenen is nu onze schatting voor $V_r^*(u)$ op basis van het beschikbare waarnemingsmateriaal voor een gegeven station.

Aanhangsel 2:

De gebruikelijke wijze van schatten van betrouwbaarheidsintervallen bij het GPV-model

Voor het schatten van de parameters van de GPV-verdeling is het gebruikelijk de methode van de grootste aannemelijkheid toe te passen.

De methode van de grootste aannemelijkheid voor het schatten van parameters uit waarnemingen beschouwt de simultane kansverdeling van de waarnemingen X_1, \dots, X_N , gedefinieerd door:

$$F(x_1, x_2, \dots, x_N) = Pr\{X_1 \leq x_1 \text{ en } X_2 \leq x_2 \text{ en } \dots \text{ en } X_N \leq x_N\} \quad (\text{A2.1})$$

De bijbehorende simultane kansdichtheidsfunctie is

$$f(x_1, x_2, \dots, x_N) = \frac{\delta^N F(x_1, x_2, \dots, x_N)}{\delta x_1 \delta x_2 \dots \delta x_N} \quad (\text{A2.2})$$

Opgevat als een functie van de onbekende parameters noemt men deze simultane kansdichtheidsfunctie de *aannemelijkheidsfunctie*.

Omdat de waarnemingen als onafhankelijk en gelijkverdeeld mogen worden beschouwd (door de bewerkingen van hoofdstuk 4) kan, in overeenstemming met de notatie van § 6.1, deze aannemelijkheidsfunctie voor het GPV-model geschreven worden als

$$L(\sigma_u, \gamma) = f(y_1, y_2, \dots, y_m | \sigma_u, \gamma) = \prod_{i=1}^m f(y_i | \sigma_u, \gamma) \quad (\text{A2.3})$$

met $f(y | \sigma_u, \gamma) = \frac{\delta G(y; \sigma(u), \gamma)}{\delta y}$ (zie ook formule (6.13)).

Hierin is y dus het boven drempel u uitstekende deel van een waarneming; m is het aantal waarnemingen dat boven u uitsteekt. De meest aannemelijke schatters van σ_u en γ zijn nu die waarden van σ_u en γ waarvoor de aannemelijkheidsfunctie maximaal is. Ze worden berekend door de partiële afgeleiden van de aannemelijkheidsfunctie naar σ_u en γ gelijk aan te stellen aan 0. Meestal wordt niet de afgeleide van de aannemelijkheidsfunctie zelf bepaald, maar die van de natuurlijke logaritme ervan. Er ontstaan in het geval van twee te schatten parameters twee vergelijkingen waaruit de parameters langs iteratieve weg moeten worden opgelost. Een gebruikelijke iteratieprocedure is de methode van Newton-Raphson. In het algemeen geldt nu dat parameters die geschat zijn volgens de methode van de grootste aannemelijkheid voor grote N normaal verdeeld zijn.

Het gaat echter niet om de verdeling van de parameters maar om die van de kwantiel-schattingen. Voor de berekening daarvan wordt gebruik gemaakt van de volgende algemene regel uit de waarschijnlijkheidsleer: Als z een willekeurige lineaire functie is van de variabelen x en y , dan geldt voor het gemiddelde van z :

$$\mu_z = a + b\mu_x + c\mu_y \quad (\text{A2.4})$$

en voor haar variantie:

$$\sigma_z^2 = b^2 \sigma_x^2 + 2bc \operatorname{cov}(x, y) + c^2 \sigma_y^2 \quad (\text{A2.5})$$

Bovendien geldt dat z normaal verdeeld is als x en y normaal verdeeld zijn. Nu is een kwantielschatting x_p een niet-lineaire functie van σ en γ - zie formule (6.15) - en wordt een benadering gebruikt op basis van een linearisering van deze functie ontwikkeld vanuit het punt $(\hat{\sigma}, \hat{\gamma})$. Daartoe dienen de partiële afgeleiden van deze functie te worden bepaald naar de beide parameters. De gelineariseerde functie heeft de gedaante

$$h(\sigma, \gamma) \approx h(\hat{\sigma}, \hat{\gamma}) + (\sigma - \hat{\sigma}) \left\langle \frac{\delta h}{\delta \sigma} \right\rangle_{(\hat{\sigma}, \hat{\gamma})} + (\gamma - \hat{\gamma}) \left\langle \frac{\delta h}{\delta \gamma} \right\rangle_{(\hat{\sigma}, \hat{\gamma})} \quad (\text{A2.6})$$

waarbij de functie $h(\sigma, \gamma)$ gegeven wordt door

$$h(\sigma, \gamma) = -\frac{\sigma}{\gamma} \left(1 - \left(\frac{P_0}{Q_u} \right)^{-\gamma} \right) \quad (\text{A2.7})$$

Combinatie van de formules (A2.5) en (A2.6) leidt tot de uitdrukking voor de variantie van de kwantielschatting volgens formule (5.45). Omdat de schattingen voor σ en γ bij benadering normaal verdeeld zijn, heeft de kwantielschatting volgens formule (5.45) ook bij benadering een normale verdeling.

Aanhangsel 3:

Bootstrap-betrouwbaarheidsintervallen voor het GPV-model en het VVM-0-model

Het schatten van betrouwbaarheidsintervallen van een kwantielschatting d.m.v. een zgn. bootstrap-procedure wordt vaak toegepast wanneer:

- geen theoretisch model voorhanden is voor het schatten van de betrouwbaarheidsintervallen;
- een theoretisch model een symmetrische benadering geeft (bijv. asymptotisch normaal), terwijl men toch de scheefheid (asymmetrie) van de verdeling van de kwantielschatting tot uiting wil laten komen.

Bij de CON-methode zijn om de eerstgenoemde reden bootstrap-betrouwbaarheidsintervallen berekend; de berekeningswijze is beschreven in § 6.6. Men spreekt wel van een parametrische bootstrap-methode, omdat de waarnemingen worden beschreven d.m.v. een (geschatte) parametrische kansverdelingsfunctie, waarvan vervolgens wordt aangenomen dat die exact geldig is, en waaruit telkens steekproeven van k onafhankelijke trekkingen worden genomen.

In verband met de ongerijmdheden die optreden m.b.t. de ondergrenzen van de methoden GPV, VVM-0 en GEV (zie § 7.3) is, om meer inzicht te krijgen in de werkelijke vorm van de verdeling van de kwantielschattingen, ook voor het GPV-model de parametrische bootstrap-methode toegepast voor het schatten van de betrouwbaarheidsintervallen. De gevolgde procedure bestaat per station uit de volgende stappen:

- uitgangspunt is de GPV-verdeling met de parameters zoals vastgesteld in § 6.5 (tabel 6.4);
- uit deze „moeder“-verdeling wordt een steekproef van k (uit tabel 6.4) onafhankelijke trekkingen genomen volgens een Monte-Carlo-procedure, worden uit deze steekproef de parameters γ en σ geschat d.m.v. een grootste aannemelijkheidsprocedure en worden de verschillende kwantielen geschat;
- deze simulatie wordt 10.000 keer uitgevoerd. Uit de 10.000 schattingen per kwantiel kunnen de gewenste karakteristieken van de kansverdeling van de kwantielschatting worden bepaald.

De resultaten voor de 10^{-4} -kwantielen staan in tabel A3.1.

Tabel A3.1:

Karakteristieken van de kansverdeling van het 10^{-4} -kwantiel volgens het GPV-model, bepaald met de bootstrap-methode ($N = 10.000$)

station	min.	ond.	med.	gem.	bov.	max.	st.a.	sch.c.
Vlissingen	406	451	546	558	734	1219	74,1	1,215
Hoek van Holland	316	369	467	479	651	1135	74,4	1,128
Den Helder	249	299	399	417	644	1312	91,4	1,822
Harlingen	341	373	432	437	522	649	37,5	0,696
Delfzijl	438	502	607	616	785	1044	72,3	0,854

Hierin is:

- min. = laagste geschatte waarde (NAP + cm)
- ond. = 2,5%-ondergrens (NAP + cm)
- med. = mediaan (NAP + cm), d.i. het 50%-kwantiel
- gem. = gemiddelde waarde (NAP + cm)
- bov. = 2,5%-bovengrens (NAP + cm)
- max. = hoogste geschatte waarde (NAP + cm)
- st.a. = standaardafwijking (cm)
- sch.c. = scheefheidscoëfficiënt

Opmerking: de scheefheidscoëfficiënt van een verdeling van een stochastische variabele x met gemiddelde μ en standaardafwijking σ is per definitie:

$$\text{sch.c.} = E[(x - \mu)/\sigma]^3 \quad (E = \text{verwachtingswaarde van}).$$

Hij is een maat voor de asymmetrie van een kansverdeling.

De veranderingen t.o.v. tabel 6.7 zijn verzameld in tabel A3.2.

Tabel A3.2:

Verschillen uitkomsten bootstrap-methode GPV-model t.o.v. tabel 6.7

station	gemiddelde	95%-betrouwbaarheidsinterval	2,5%-ondergrens	2,5%-bovengrens
Vlissingen	+ 3 cm	+ 3 cm	+ 36 cm	+ 39 cm
Hoek van Holland	+ 4 cm	+ 32 cm	+ 19 cm	+ 51 cm
Den Helder	+ 2 cm	+ 35 cm	+ 39 cm	+ 74 cm
Harlingen	- 3 cm	- 1 cm	+ 8 cm	+ 7 cm
Delfzijl	+ 1 cm	- 17 cm	+ 37 cm	+ 20 cm

Omdat de VVM-0-methode de beste kwantielschattingen geeft (§ 7.2) is vervolgens getracht ook voor deze methode bootstrap-betrouwbaarheidsintervallen te schatten. De hierbij gevolgde procedure is theoretisch niet geheel onderbouwd. Het is nu nog niet bekend hoe een bootstrap-methode precies moet worden toegepast bij een verdelingsvrije methode als de methode VVM-0, waarbij geen parametrische kansverdeling wordt aangepast aan de waarnemingen zelf. De gevolgde procedure geeft naar de huidige inzichten echter de beste schattingen.

Deze gevolgde procedure bestaat per station uit de volgende stappen:

- uitgangspunten vormen de (verdelingvrije) schattingen van de parameters γ , $a_{n/k}$ en $b_{n/k}$ van tabel 6.3 (§ 6.4);
- deze schattingen worden opgevat als de parameters van de GPV-verdeling: γ wordt overgenomen, terwijl $a_{n/k}$ wordt genomen voor de schaalparameter σ en $b_{n/k}$ voor de drempel u . De aldus verkregen GPV-verdeling vormt de „moeder” verdeling van de simulaties. De rechtvaardiging hiervan volgt uit de vergelijking van de formules (6.12) en (6.15). Formule (6.12) is namelijk gelijk aan formule (6.15) als de parameters op de zojuist aangegeven wijze worden vervangen;
- Volgens een Monte-Carlo-procedure wordt uit deze moederverdeling een steekproef van k onafhankelijke trekkingen genomen (k uit tabel 6.3) en worden hieruit (na correctie voor drempel u) de parameters γ , $a_{n/k}$ en $b_{n/k}$ (momentenschatters volgens de formules 5.15 en 5.18) en daarmee de verschillende kwantielen „verdelingsvrij” geschat;
- deze simulatie wordt 10.000 keer uitgevoerd. Uit de 10.000 schattingen kunnen per kwantiel de gewenste karakteristieken van de kansverdeling van de kwantielschatting worden bepaald.

De uitkomsten voor de 10^{-4} -kwantielen staan in tabel A3.3.

Tabel A3.3

Karakteristieken van de kansverdeling van het 10^{-4} -kwantiel volgens het VVM-0-model, bepaald met een bootstrap-procedure ($N = 10.000$)

station	min.	ond.	med.	gem.	bov.	max.	st.a.	sch.c.
Vlissingen	399	449	524	539	680	1086	73,2	1,250
Hoek van Holland	308	376	484	505	703	1329	105,2	1,267
Den Helder	229	284	348	357	464	791	56,1	1,102
Harlingen	284	334	383	387	452	577	36,4	0,623
Delfzijl	417	484	572	583	717	1026	71,0	0,905

De betekenis van de kolomhoofden en de gebruikte eenheden zijn gelijk aan die van tabel A3.1.

De bijlagen 35A t/m 35E geven de overschrijdingslijnen volgens de VVM-0 methode met de asymptotisch normaal benaderde betrouwbaarheidsintervallen volgens § 6.4, samen met die volgens de bovenbeschreven bootstrap-benadering.

De ondergrenzen van die „bootstrap-betrouwbaarheidsintervallen” vertonen een monotoon stijgend verloop. In de laatstgenoemde bijlagen zijn tevens de histogrammen gegeven van de 10.000 bootstrap-schattingen van de 10^{-4} -kwantielen en worden deze vergeleken met die der normale benadering volgens § 6.4.

Deze resultaten worden met name gegeven om het inzicht in de vorm van de betrouwbaarheidsintervallen te vergroten. De theoretische basis is echter te zwak om de bootstrap-betrouwbaarheidsintervallen te koppelen aan de kwantielschattingen van het VVM-0-model. De zwakte is gelegen in het hybride karakter van deze bootstrap-benadering: er worden steekproeven getrokken uit een parametrische verdeling (de GPV-„moeder”-verdeling), terwijl de VVM-0-methode juist niet uitgaat van een bepaalde parametrische kansverdeling van de waarnemingen. De gemiddelde waarden van de bootstrap-kwantielschattingen blijken dan ook enigszins af te wijken van de schattingen volgens het VVM-0-model (voor Vlissingen en Hoek van Holland slechts -1 resp. +5 cm, voor Den Helder, Harlingen en Delfzijl evenwel achtereenvolgens -13, -18 en -17 cm).

Hier spelen overigens twee effecten een rol, te weten de modelfout van de gevolgde bootstrap-procedure en de (onbekende) onzuiverheid van de VVM-0-kwantielschatting.

Opmerkelijk is verder dat de 95%-betrouwbaarheidsintervallen volgens de bootstrap-benadering kleiner zijn dan die volgens de „normale” benadering. Men zou eerder verwachten dat met de ondergrens ook de bovengrens hoger komt te liggen. Vermoedelijk speelt ook hierbij de modelfout een rol.

N.B.: Bij de hierboven besproken bewerkingen zijn voor Den Helder en Harlingen uiteraard de lange stormseizoenen vanaf mei 1932 beschouwd en voor elk van de drie andere stations die uit de volle beschikbare periode.

Verwijzingslijst van belangrijke begrippen

De onderstaande verwijzingen betreffen de plaats waar het desbetreffende begrip is toegelicht.

	§, c.q. hfdst.		§, c.q. hfdst.
apogeum	2.1.1	HW-opzetbestand	4.2.2
astronomisch getij	2.1.1	knopenlijn	2.1.1
attractiedomein	6.1	„koppelingsformule”	5.2
autocorrelatie	3.3.2	kwantiel (b.v. 10^{-4}	
basispeil	1	-kwantiel)	4.3.3
basisstation	2.1.2		en 5.1
bestand astronomische		LW-opzet	2.1.2
hoogwaterstanden	4.2.2	maansdeclinatie	2.1.1
bootstrap-methode	6.6	momentenschatter	5.2
boxplot	4.3.3	nodale tij	2.1.1
bui-oscillatie	4.2.1	ontwerppeil	1
bui-stoot	4.2.1	opwaaiing	2.1.2
circulatie(patroon)	2.2.1	(rechte) opzet	2.1.2
circulatietype	2.2.1	perigeum	2.1.1
clustergrootheid V_r^* (u)	4.3.4	piekwaarde	4.3.4
CON(volutiemodel)	5.3	Poisson-proces	4.3.4
culminatiemethode	2.1.1	poolgetij	2.1.1
cyclonaliteit	2.2.1	„poort”	3.1.1
dagelijkse ongelijkheid	2.1.1	selectie Van der Ham	3.1.1
depressie	2.2	slotgemiddelde	2.1.1
depressiebaan	2.2	springtij	2.1.1
doodtij	2.1.1	stationariteit	4.3.3
drempel	3.1.1	stormeffect	2.1.2
D-S-i-selectie	4.3.4	stormjaar	7.4
extreme-waardenindex γ	5.2	stormklimaat	4.3.5
extreme-waardenverdeling	5.2	stormseizoen	4.3.3
gegeneraliseerde Pareto-		lange stormseizoen	4.3.3
verdeling (GPV)	5.3	korte stormseizoen	4.3.3
gelijkverdeelde		stormvloed	2.1.2
waarnemingen	4.4.1	tijverschil	2.1.1
getij	2.1.1	verdelingsvrije methode	5.2
getijkromme	2.1.1	verticaal getij	2.1.1
getijverwekkende kracht	2.1.1	vervroeging van het	
grenspeil	1	HW e.d.	2.1.2
harmonische analyse	2.1.1	voorwaardelijke	
herhalingstijd	3.2	verdelingsfunctie	5.3
homogeniteit	3.3.2	VVM-c-methode	5.2
HW = hoogwaterstand	2.1.2	VVM-o-methode	5.2
hoogwater-		zeebeer	8.2
overschrijdingslijn	1		
horizontaal getij	2.1.1		
HW-bestand	4.2.1		
HW-opzet	2.1.2		

Grafische verzorging:
Meetkundige Dienst
Afdeling Grafische Technieken