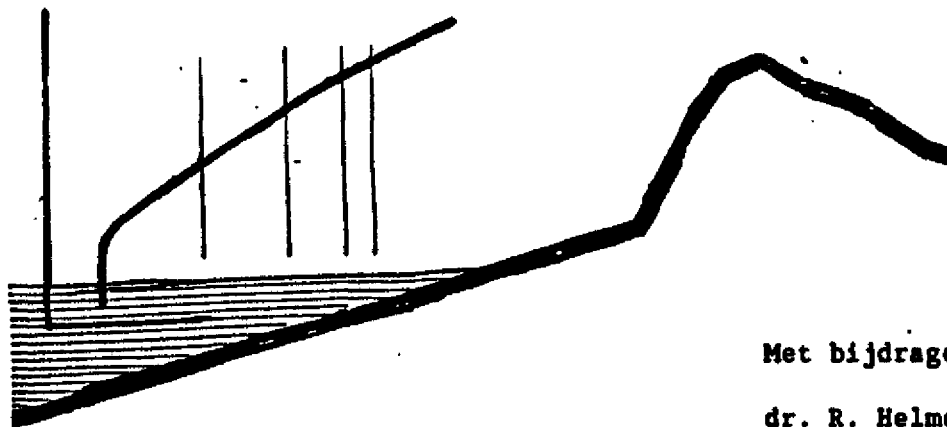


nota GWIO 89.001

FREQUENTIELIJNEN KUST

Tussentijds verslag 2



Met bijdragen van:

dr. R. Helmers, CWI
drs. A.L.M. Dekkers, CWI
prof.dr. L. de Haan, EUR
hr. H. Daan, KNMI
ir. J. van Malde, RW-DGW
ir. J.G. de Ronde, RW-DGW
ir. D. Dillingh, RW-DGW

Redactie:
dr.ir. J.W. van der Made, RW-DGW

Februari 1989

Deze nota vormt een tussentijdse vastlegging van het onderzoek "Frequentielijnen Kust" voor medio 1988. Uit de opgenomen cijfers, grafieken en tabellen kan evenwel geen enkele conclusie worden getrokken over de uiteindelijk vast te stellen frequentielijnen voor de kuststations.

INHOUD

blz

| | |
|---|----|
| 1. Inleiding | 4 |
| 2. Meteorologische gezichtspunten (KNMI) | |
| 2.1 Het verband tussen circulatiepatronen en hoge opzetten | 7 |
| 2.2 De persistentie van circulatietypen | 8 |
| 2.3 Circulatietypen en het selectieprobleem | 10 |
| 2.4 Selectie zonder meteorologische criteria | 10 |
| 3. HW-opzet als variabele | |
| 3.1 Inleiding | 13 |
| 3.2 Hoogwaterstanden of hoogwateropzetten (RW-DGW) | |
| 3.2.1 Probleemstelling | 13 |
| 3.2.2 De hoogwaterstand (HW) als statistisch veranderlijke | |
| 3.2.2.1 Algemene beschouwing | 13 |
| 3.2.2.2 De draagwijdte van de deterministische aard der astronomische hoogwaterstanden | 14 |
| 3.2.3 HW-opzet als veranderlijke | |
| 3.2.3.1 Algemene overwegingen | 15 |
| 3.2.3.2 Te stellen eisen aan de HW-opzet | 17 |
| 3.2.4 HW-opzet als enige opwaaiingsparameter | 18 |
| 3.3 Koppeling HW-opzetten en astronomische standen (EUR) | 19 |
| 3.4 Beschikbaarheid en verdeling van de astronomische standen (RW-DGW) | |
| 3.4.1 Inleiding | 20 |
| 3.4.2 De waarnemingsperiode | 21 |
| 3.4.3 De berekening van de astronomische hoogwaterstanden | 21 |
| 3.4.4 Frequentieverdeling van de astronomische hoogwaters in de huidige toestand | 22 |
| 3.5 Toetsing van onafhankelijkheid van astronomisch HW en HW-opzet (RW-DGW) | |
| 3.5.1 Inleiding | 27 |
| 3.5.2 De Spearman toets voor de hypothese van onafhankelijkheid | 27 |
| 3.5.3 Groepen van gelijke waarnemingen (ties) | 28 |
| 3.5.4 Toepassing op de astronomische HW's en HW-opzetten | 28 |
| 4. Waarnemingsperiode en stormseizoen | |
| 4.1 Waarnemingsperiode (RW-DGW) | 31 |
| 4.2 Het stormseizoen (RW-DGW;CWI) | |
| 4.2.1 Inleiding | 31 |
| 4.2.2 Criteria voor de keuze van een stormseizoen | 31 |
| 4.2.3 Boxplotten | 32 |
| 4.2.4 Het beschouwde waarnemingsmateriaal van de HW-opzetten | 32 |
| 4.2.5 De bepaling van de lengte van het stormseizoen | 36 |
| 5. Selectiemethode | |
| 5.1 Inleiding | 44 |
| 5.2 Enkele opmerkingen over autocorrelatie bij HW-opzetten (RW-DGW) | |
| 5.2.1 Probleemstelling | 44 |
| 5.2.2 Afhankelijkheden binnen een tijdreeks | 44 |
| 5.2.3 Beperking der afhankelijkheid | 45 |
| 5.2.4 Invloed selectiemethode | 46 |
| 5.3 Selectie van HW-opzetten (CWI) | |
| 5.3.1 Afhankelijkheid en clustering van hoge HW-opzetten | 49 |
| 5.3.2 Resultaten | 51 |
| 5.3.3 Conclusie | 51 |
| 5.4 Extreme-waarden-theorie voor afhankelijke waarnemingen en de selectiemethode Daan (CWI) | 58 |
| 6. Invloed van de drempelwaarde | |
| 6.1 Inleiding | 62 |
| 6.2 Het drempelniveau voor de HW-opzet (RW-DGW) | 62 |

| | | |
|-------|---|----|
| 6.3 | Onderzoek naar het verloop van het 10^{-4} -kwantiel bij oplopende drempel (CWI) | |
| 6.3.1 | Vroegere resultaten | 65 |
| 6.3.2 | Het geschatte 10^{-4} -kwantiel als functie van de drempelwaarde in het Gumbelmodel | 65 |
| 6.3.3 | Het geschatte 10^{-4} -kwantiel als functie van de drempelwaarde in het EV-model | 67 |
| 6.3.4 | De simulaties | 68 |
| 6.3.5 | Conclusie | 69 |
| 7. | Conclusies | 77 |
| 8. | Literatuur | 78 |
| | Bijlage | 79 |

Afkortingen:

KNMI Koninklijk Nederlands Meteorologisch Instituut, De Bilt
RW-DGW Rijkswaterstaat Dienst Getijdewateren, 's-Gravenhage
EUR Erasmus Universiteit, Rotterdam
CWI Centrum voor Wiskunde en Informatica, Amsterdam.

1. INLEIDING

Na de verschijning van het eerste "Tussentijdse verslag" in het kader van het onderzoek "frequentielijnen kust" (Rijkswaterstaat, 1986) is er inmiddels aanleiding een tweede verslag te doen verschijnen. Dit betreft:

- de keuze van de HW-opzetten als de voornaamste voor dit onderzoek te gebruiken variabele;
- de vaststelling van de bruikbare waarnemingsperiode per station;
- de vaststelling van het lange stormseizoen (1 oktober ... 15 maart) en van het korte stormseizoen (1 november ... 31 januari);
- de te gebruiken selectiemethode voor de hoogwaters (HW'S) en de HW-opzetten ;
- de verdeling van het astronomisch getij *).

De argumenten voor deze beslissingen vindt men in de volgende hoofdstukken.

Ten tijde van de samenstelling van dit tussentijds verslag zijn nog geen beslissingen genomen over de volgende punten:

- de voor de HW-opzetten toe te passen drempelwaarde.
- de keuze van de frequentieverdelingen.
- de samenvoeging van beide frequentieverdelingen tot één enkele voor HW-standen.
- de vaststelling van de betrouwbaarheidsgrenzen.

Er wordt verder op gewezen dat gestreefd wordt naar een resultaat, geldend voor de huidige toestand. Eventuele wijzigingen ten gevolge van toekomstige zeespiegelrijzing vallen buiten het kader van het lopende onderzoek; deze zullen afzonderlijk in rekening moeten worden gebracht.

Dit 2e tussentijdse verslag begint met een verhandeling over meteorologische gezichtspunten (hoofdstuk 2). Een verstoring van de waterbeweging ten gevolge van een storm blijkt nog 3 & 4 getijden door te werken. Van de HW's die binnen zo'n periode vallen zou men er slechts één (de hoogste) mogen verwerken om vrij te zijn van autocorrelatie. Gezien de persistentie van de atmosferische omstandigheden zou echter een termijn van tenminste twee weken meer in de rede liggen. In dat geval zou men selecteren op circulatiepatronen en niet op afzonderlijke stormen.

In hoofdstuk 3 vindt men de motivering voor de keuze van de HW-opzet als de te onderzoeken variabele in plaats van de HW-stand zelf. Aan beide benaderingen zijn voor- en nadelen verbonden. Deze zijn hieronder samengevat.

HW-standen

Voordelen:

- geven direct de gezochte waarden.
- de enige foutenbron komt voort uit de metingen.

*) De getijbeweging zoals die op zou treden bij "gemiddelde" weersomstandigheden.

Nadelen:

- de invloeden van het weer en van het getij zijn niet duidelijk te onderscheiden;
- er is de invloed van zeespiegelrijzing en van trendbreuken, bijv. als gevolg van havenaanleg;
- er zijn in beginsel twee drempelwaarden nodig, nl. voor de opzet én voor het HW zelf.

HW-opzetten**Voordelen:**

- er spelen alleen weersinvloeden, waardoor er beter inzicht ontstaat in de autocorrelatie.
- er hoeft slechts één drempelwaarde gekozen te worden, zodat:
- er meer waarnemingen beschikbaar zijn.

Nadelen:

- kennis is nodig van de astronomische getijstanden; de berekening ervan brengt extra onnauwkeurigheden met zich mee.
- uiteindelijk dienen de conclusies toch overgebracht te worden op de HW-standen.

Opzetten bij LW of op andere tijdstippen in de getijcyclus buiten HW zijn niet gebruikt. Dit vanwege optredende wisselwerking tussen getij en opzet. De veranderde hydraulische omstandigheden, die een gevolg zijn van een verhoging van de waterstand leiden nl. in het algemeen tot een vormverandering van de ongestoorde getijkromme, zowel naar tijd (vervroeging) als naar amplitude. Verwezen wordt ook naar par. 2.4.1 van het eerste Tussentijdse verslag (juni 1986).

Hoofdstuk 3 geeft verder een aantal beschouwingen over de samenvoeging van beide frequentieverdelingen. De definitieve behandeling hiervan zal in een latere publicatie volgen.

De perioden die het waarnemingsmateriaal leveren vormen het onderwerp van hoofdstuk 4. Het betreft zowel de totale periode in jaren als de periode binnen het jaar, het stormseizoen. Het stormseizoen binnen een winter werd vastgesteld op de periode 1 oktober ... 15 maart. Dit volgde uit een onderzoek naar de grootte van de opzetten, en naar het gedrag van de overschrijdingsfrequenties in de loop van het jaar, zoals gedemonstreerd door de zg. boxplotten, opgenomen in hoofdstuk 4. Een korter stormseizoen, nl. 1 november ... 31 januari, waarin de stormen min of meer stationair over de tijd zijn verspreid, werd gebruikt om de mate van clustering te bepalen, nodig in verband met de toe te passen selectieprocedure.

Hoofdstuk 5 is gewijd aan de wijze van selecteren. De destijds door de Deltacommissie toegepaste methode van selectie op depressiebanen is hierbij verlaten vanwege de subjectieve aanpak ervan en omdat sommige gevaarlijke depressies wellicht ten onrechte buiten beschouwing blijven. Bovendien blijft er enige autocorrelatie en clustering aanwezig. Verwezen wordt hierbij naar par. 3.4.1 van het eerste Tussentijdse verslag. Er is nu een meer statistisch gefundeerde methode toegepast.

Het optreden van autocorrelatie en clustering heeft invloed op de vorm van de frequentiekromme. Autocorrelatie blijkt nagenoeg te verdwijnen bij toepassing

van de selectiemethode, aangeduid als "Daan-S-4" (zie ook par. 2.4). De resterende reeks omvat de elementen voor het verdere frequentieonderzoek.

Tenslotte komt in hoofdstuk 6 de invloed van de toegepaste drempelwaarde aan de orde. Deze blijkt vooral voor de uitkomsten van de naar een frequentie van 10^{-4} /jaar geëxtrapoleerde waarde zeer belangrijk te zijn. Zoals eerderopgemerkt is over de drempelwaarde nog geen besluit genomen.

Het onderzoek vindt momenteel nog voortgang ten aanzien van de nog niet opgeloste vragen, aangegeven in het begin van de Inleiding. Een verslag hierover zal binnenkort verschijnen.

De werkgroep is sedert het verschijnen van het 1^e Tussentijds Verslag uitgebreid met ir. D. Dillingh (RW-DGW).

2. METEOROLOGISCHE GEZICHTSPUNTEN (KNMI) =).

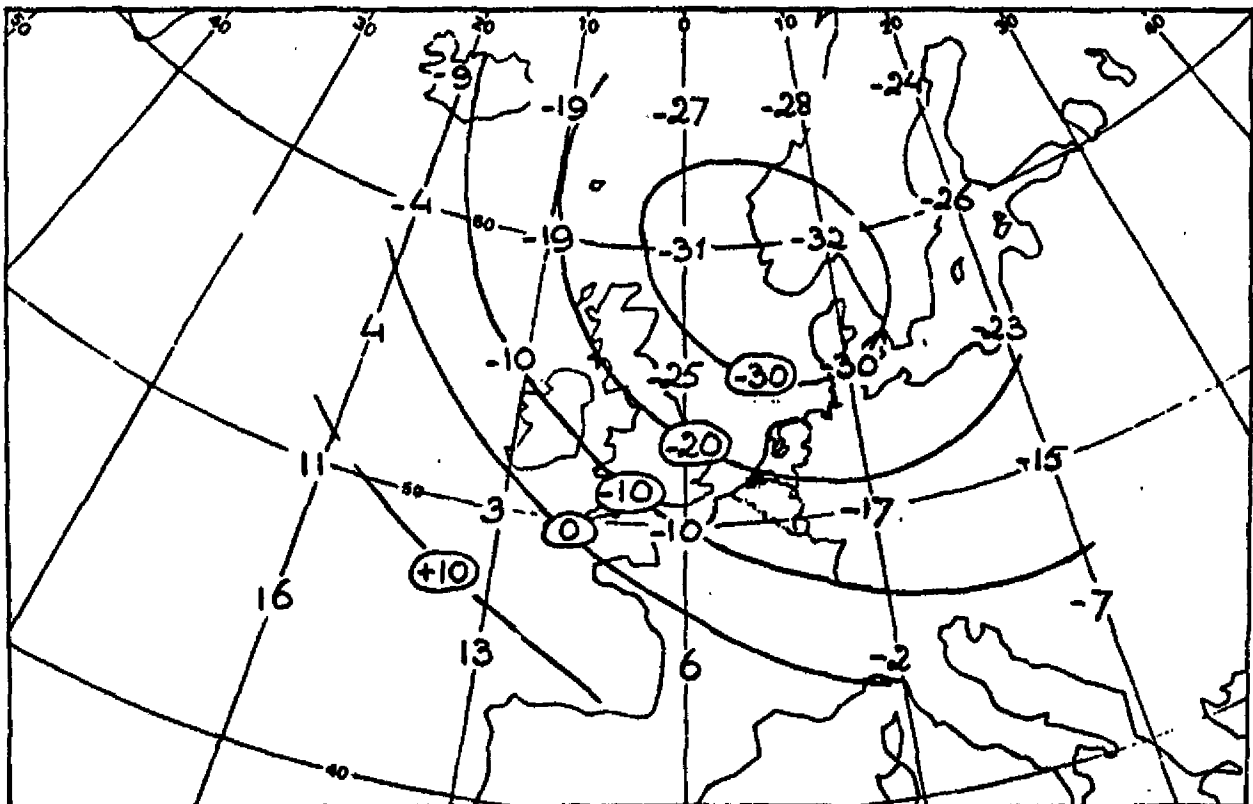
2.1 Het verband tussen circulatiepatronen en hoge opzetten

De weersontwikkeling aan het aardoppervlak wordt min of meer "gestuurd" door de luchtstromingen op het 500-mbar vlak. 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 2.1. Hierin is een schets gegeven van de geografische verdeling van de correlatiecoëfficiënt van de 500-mbarhoogte ter plekke met de opzet in Hoek van Holland.

N.B. De correlaties zijn berekend over de jaren 1949 ... 1980. Voor alle HW's met een opzet < 50 cm is een opzet 0 verondersteld.

Uit de figuur blijkt duidelijk een negatief verband tussen de 500-mbarhoogte boven zuid-Noorwegen en de HW-opzet in Hoek van Holland. De afbeelding kan zelfs gezien worden als een illustratie van het voorkeurs-circulatie-patroon voor hoge opzetten.



figuur 2.1 Correlatie coëfficiënten van de 500-mbarhoogte boven West Europa met de HW-opzet te Hoek van Holland (in %).

x) In dit hoofdstuk is steeds gewerkt met het korte stormseizoen 1 november ... 31 januari.

Het verband kan ook verduidelijkt worden door gebruik te maken van de circulatietypen volgens KRUIZINGA (1978). Deze classificatie is gebaseerd op een hoofdcomponenten-analyse van het 500-mbar-patroon, waarbij de eerste 3 hoofdcomponenten in hoge mate bepalend bleken voor de circulatie. Deze drie zijn als volgt te karakteriseren:

- 1e 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 combinatie 27 circulatietypen, die tezamen de zgn. P27-classificatie vormen. De typen zijn geïdentificeerd met een nummer uit de reeks 1 ... 27.

N.B. De classificatie heeft de eigenschap dat alle klassen in principe even frequent voorkomen. Dat geldt echter niet wanneer alleen naar bepaalde seizoenen gekeken wordt, zoals wanneer we vergelijken met HW-opzetten in de maanden november t/m januari.

Een samenvatting van de frequenties van opzetten > 50 cm en > 100 cm per circulatietype over de periode 1949 ... 1979 is opgenomen in Tabel 2.1.

Type 3 (dat weliswaar in de wintermaanden relatief veel voorkomt: ruim 7% van alle dagen tegen 3,7% over het hele jaar) blijkt samen te vallen met meer dan de helft van alle opzetten > 100 cm. Dit type is gekarakteriseerd door:

- sterke west-stroming
- noordelijke stroming
- cyclonaliteit (stroming rond lagedrukgebied)

Verder blijkt een zeer beperkt aantal types samen te gaan met het overgrote deel van de hoge opzetten. Maar naast type 3 leveren slechts de types 2 en 6 een "bovennormale" kans op een hoge opzet.

Een verdeling van de frequenties over de tercielen van de hoofdcomponenten is gegeven in Tabel 2.2. Daarbij blijkt de sterkte van de west-circulatie de belangrijkste factor te zijn voor hoge opzetten.

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 bij een willekeurige toedeling van typen over de dagen zou moeten worden verondersteld. De mate van persistentie kan worden uitgedrukt in de gemiddelde duur van een periode met een bepaald weertype. Voor de types met een verhoogde kans op een hoge 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 types 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.

* Tabel 2.1 : frequentie van hoge opzetten bij verschillende circulatietypen.

| C | Nc | Fc% | G = 50 cm | | | G = 100 cm | | |
|----------|-----|-----|-----------|-----|----|------------|-----|----|
| | | | Tg | Fg% | p% | Tg | Fg% | 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 |
| N = 2582 | | | T = 460 | | | T = 76 | | |
| | | | 100 | 16 | | 100 | 3 | |

C = circulatietype

Nc = aantal dagen met type C in nov ... jan

N = totaal aantal dagen

Fc% = relatieve frequentie van type C: $(Nc/N) \cdot 100\%$

G = drempelwaarde opzet

Tg = aantal opzetten per circulatie-type

T = totaal aantal opzetten

Fg% = $(Tg/T) \cdot 100\%$; verdeling van de opzetten over de circulatietypenp% = $(1/2) \cdot (Tg/Nc) \cdot 100\%$, frequentie van de opzetten per circulatietype; waarbij in de factor 1/2 het optreden van twee HW's per dag is verdisconteerd.

Tabel 2.2: Frequenties van hoge opzetten bij verschillende tercieilen van de hoofdcomponenten.

| H | omschrijving | Nc | FcZ | G = 50 cm | | | G = 100 cm | | |
|-----|---------------|------|-----|-----------|-----|----|------------|-----|----|
| | | | | Tg | FgZ | PZ | Tg | FgZ | PZ |
| I | sterk west | 1938 | 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 |

H = hoofdcomponent
verder als bij tabel 2.1

2.3 Circulatietypen en het selectieprobleem

Het doel van selectie van gegevens is onderlinge onafhankelijkheid te verkrijgen. Wanneer we gegevens, die optreden in een periode met een persistent circulatietype, als afhankelijk beschouwen, dan moet uit zo'n periode in principe slechts één gegeven worden geselecteerd (en wel de hoogste opzet in die periode).

Circulatietypen zijn slechts over een relatief klein deel van de totale reeks opzetten bekend. Een selectie-methode volgens bovengenoemd uitgangspunt zou dus slechts gebaseerd kunnen worden op een gemiddelde duur van circulatietypen. Voor het "gevaarlijke" type ligt deze op ongeveer 3 HW's, voor de drie typen met verhoogde kans op een grote opzet op 4 à 5 HW's. Een selectie criterium zou dus kunnen zijn dat van 3 tot 5 opeenvolgende HW's slechts één opzet geselecteerd zou mogen worden.

2.4. Selectie zonder meteorologische criteria x)

Een andere vraag is of verschijnselen die veroorzaakt worden door toestanden die een zekere persistentie hebben en daardoor geclusterd optreden, om die reden als onderling afhankelijk moeten worden beschouwd. Men zou pas van werkelijke afhankelijkheid kunnen spreken als de verschijnselen zelf elkaar beïnvloeden.

Indien dit laatste als enig uitgangspunt zou moeten gelden, dan heeft het slechts zin om naar de onderlinge beïnvloeding van HW-standen en opzetten te kijken, en kan snel, geconcludeerd worden dat het meest langdurige effect (de terugslingering) de periode van onderlinge afhankelijkheid tot 3 HW's beperkt.

*) Zie lit. 13, par. 3.4.

Het selectie criterium kan dan uiterst simpel als volgt geformuleerd worden:
 Geselecteerd worden alle opzetten die hoger dan de i voorafgaande en de i opvolgende opzetten, en bovendien hoger zijn dan een zekere drempelwaarde.

Deze methode wordt aangeduid met "Daan- i selectie". Een bezwaar van deze methode is dat in verscheidene gevallen een HW-opzet wordt geselecteerd en een aangrenzende hogere HW-opzet niet. Dit is het gevolg van het feit dat deze laatste binnen de radius i ligt van een nog hogere HW-opzet. In de figuur 2.2 hieronder is een voorbeeld gegeven.

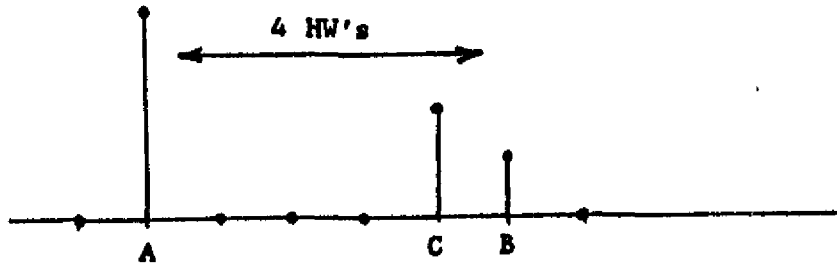


Fig. 2.2 Ten onrechte uitgevoerde selectie van HW-opzet B.

Bij een Daan-4 selectie wordt A in eerste instantie geselecteerd. Op grond daarvan wordt C geëlimineerd, terwijl B in een later stadium weer wél voor selectie in aanmerking komt. Dit is erg onlogisch wanneer we er van uitgaan dat tussen twee opeenvolgende HW-opzetten (i.c. tussen B en C) een sterke correlatie aanwezig is. Indien een interval van 4 HW's wordt aangehouden als de duur waarover nog correlatie optreedt ($i=4$), dan dienen B en C beide geëlimineerd te worden.

Een ander probleem doet zich voor als er na een eerste piek A een opzet B volgt, zelf geen piek zijnde, maar toch hoger dan een 2e piek C, die ten opzichte van A juist buiten het selectie criterium valt, maar ten opzichte van B niet. Zie fig. 2.3.

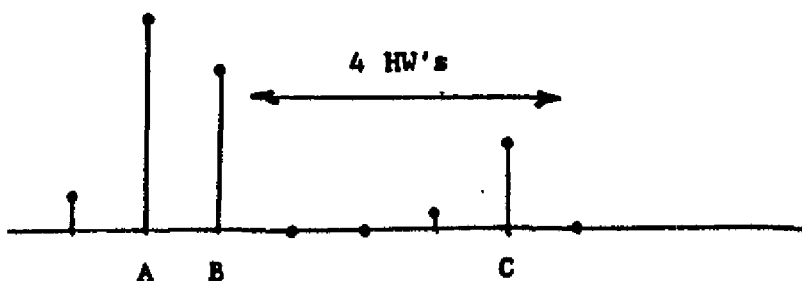


Fig. 2.3 Eliminatie van een piek-HW, C, door een hoger HW, B.

x) Het woord "eliminatie" moet niet opgevat worden als "weglaten" of "buiten beschouwing laten". Een geëlimineerde piek blijft wel meetellen als HW in de tijdreeks, echter zijn waarde telt niet meer mee, maar wordt als het ware op nul gesteld.

Indien als selectie criterium weer een duur van 4 HW's wordt genomen dan zal in dit geval piek C moeten vervallen.

Aldus komt men tot de volgende regel, aangeduid als DAAN-S-i selectie:

- Selecteer alle HW-opzetten boven de gekozen drempel waarvoor binnen een radius van i HW-opzetten geen hogere HW-opzetten voorkomen.

In hoofdstuk 5 wordt de bepaling van het getal i verder behandeld.

3. HW-OPZET ALS VARIABELE

3.1. Inleiding.

Bij het frequentieonderzoek komt de vraag naar voren van welke grootheid de frequenties in eerste aanleg moeten worden onderzocht: van de hoogwaterstanden zelf (HW's) of van de hoogwateropzetten. In par. 3.2 zal blijken dat het gebruik van de hoogwateropzetten de voorkeur verdient.

Par. 3.3 behandelt de samenvoeging van de frequentie verdelingen van de hoogwateropzetten met die van de astronomische hoogwaterstanden, leidende tot de frequentieverdeling van de te verwachten hoogwaterstanden. Deze werkwijze moet nog nader worden uitgewerkt.

De berekening van de astronomische hoogwaterstanden en de frequentieverdeling ervan vormt het onderwerp van par. 3.4. Par. 3.5 tenslotte is gewijd aan de vraag of het astronomisch hoogwater en de hoogwateropzet onderling al dan niet onafhankelijk zijn.

3.2. Hoogwaterstanden of hoogwateropzetten.

3.2.1. Probleemstelling

Bij het "Onderzoek Frequentielijnen Kust" heeft men voor de statistisch te verwerken grootheid de keuze tussen de hoogwaterstand ("HW") en de hoogwateropzet ("HW-opzet"); de laatste is daarbij gedefinieerd als het verschil tussen de opgetreden hoogwaterstand en de voor het betrokken hoogwater voorspelde astronomische hoogwaterstand, dit ongeacht een eventuele vervroeging of verlating van het werkelijke t.o.v. het astronomische getij (lit. 13 par. 2.1.3, blz. 7/8). De voor de tweede fase van het onderzoek gemaakte keuze voor HW-opzet als de voornaamste te verwerken veranderlijke wordt in het hiernavolgende nader gemotiveerd. Aan het slot komen voorts nog ter sprake de overwegingen om er van af te zien de LW-opzetten mede in de beschouwingen te betrekken.

3.2.2. De hoogwaterstand (HW) als statistisch veranderlijke

3.2.2.1. Algemene beschouwing

In de vijftiger jaren heeft het toenmalige Mathematisch Centrum in zijn desbetreffende statistische bijdrage aan het rapport der Deltacommissie HW als te verwerken grootheid gekozen en zich m.b.t. HW-opzet beperkt tot een oriënterend onderzoek. Dit laatste voerde tot de slotsom dat verwerking van HW-opzet "tot (nagenoeg) dezelfde resultaten zou leiden" als verwerking van HW en dit vormde een der hoofdargumenten om dit statistisch onderzoek (voor Hoek van Holland) te verrichten met als variabele HW en niet HW-opzet (lit. 2, par. 5.0, bl. 20). Die handelwijze biedt de volgende voordelen:

- a de benadering is rechtstreeks (lit. 2, par. 5.0); ieder gezocht peil is immers een (bijzonder) HW en in de voorgekomen HW's zijn naast alle periodieke verschijnselen ook de geregeld opgetreden en moeilijk ontwarbare complicaties (zoals vervroeging of verlating van het getij) geïntegreerd;
- b HW-waarnemingen zijn verhoudingsgewijze nauwkeurig (foutenbronnen: bedieningsfouten en de onnauwkeurigheden van de waarneming, in de nulpunts-

handhaving en bij de bewerking); als standaardmeetafwijking van de vastgestelde meetwaarde wordt 2,5 cm aangehouden (zie lit. 13, par. 2.4.2).

Bij deze benaderingswijze dient men nog een selectie toe te passen om te lage HW-waarden uit te sluiten aangezien hun verdeling afwijkt van die der hoge waarden. Een dergelijke homogenisering is b.v. mogelijk m.b.v. onderwaarden (zgn. drempels), waarvan er dan per reeks niet alleen één voor HW zelf wordt vastgesteld maar vanwege de vrij sterk wisselende waarden der astronomische HW-standen ook één voor HW-opzet. Bij eerdere verwerkingen (voor Vlissingen en Hoek van Holland) in het kader van het Onderzoek Frequentielijnen Kust is een zodanige selectie ook toegepast met als drempel voor HW achtereenvolgens 265 en 165 cm (boven NAP) en voor HW-opzet: 50 cm - zie lit. 13, par. 4.3. Intussen kleven aan de keuze van HW als variabele ook enkele belangrijke nadelen:

- 1e Een HW op enige plaats is per definitie de som van het bijbehorende astronomische hoogwater en de HW-opzet. Zoals ook in de reeds genoemde literatuur vermeld is het eerste bestanddeel deterministisch bepaald en bovendien gekenmerkt door een ingewikkeld samenstel van cyclische schommelingen (waaronder die van het nodale getij met een periode van 18,6 jaar).
- 2e HW, het jaargemiddelde van HW, vertoont in samenhang met de zeespiegelrijzing een rijzende trend (bijlage, staat II), welke voor vele Nederlandse (en Duitse!) getijstations sinds ten minste ettelijke decennia die van de gemiddelde zeestand ter plaatse overtreft; daarnaast komen in de grafieken van HW en van andere zeeniveaus als gevolg van uitgevoerde werken ook discontinuïteiten voor. Beide verschijnselen worden teweeggebracht door veranderingen in het verticale getij (lit. 9). Hangende de uitkomsten van lopende onderzoekingen hiernaar zijn correcties hierop, die per station moeten worden vastgesteld, onzeker.

Opmerking: Het statistisch onderzoek van het Deltarapport betreft alleen Hoek van Holland. Afgezien van enkele Waddenzeestations is men voor de andere stations (lit. 13, par. 2.2) uitgegaan van HW-overschrijdingslijnen 1901...-1950, waarbij voor de stations tussen Haringvliet en Belgische grens - dus ook voor Vlissingen - de waarnemingen zijn herleid tot gemiddeld tij. Dit laatste betekent dat deze gevonden overschrijdingslijnen evenwijdig lopen aan die voor de HW-opzet ter plaatse; de procedure is theoretisch niet correct omdat de verdeling van HW-opzet gekoppeld moet worden aan die van de astronomische hoogwaterstand (lit. 13, par. 4.6; zie verder hierna par. 3.2.3.2).

3.2.2.2. De draagwijdte van de deterministische aard der astronomische hoogwaterstanden

De praktische betekenis van de deterministische aard der astronomische hoogwaterstanden is in algemene zin niet eenvoudig aan te geven. Uiteraard wordt die betekenis groter naarmate

- het astronomisch hoogwater hoger is;
- de windinvloed (i.c. HW-opzet) kleiner is.

M.a.w. voor ieder station zal de betreffende betekenis van de deterministische component in het algemeen reeds van hoogwaterstand tot hoogwaterstand verschillen (dagelijkse ongelijkheid! - zie bijlage, staat III) en evenzeer kan ze voor twee gelijke HW's aanmerkelijk uiteenlopen. Verder bestaan er tussen de stations onderling ook nog grote verschillen: aan de stations in het Noorden des lands met deels een betrekkelijk zwakke astronomische getijbeweging is de opwaaiing niet alleen relatief maar ook in volstrekte zin al gauw groter dan in het Zuidwesten, waar de astronomische getijbeweging matig

tot krachtig is. Dit laatste blijkt (bijlage 2, staat I) uit de vergelijking van de waarden der zgn. grenspeilen G^2) met de overeenkomende astronomische hoogwaterstanden voor gemiddeld springtij (HWS), gemiddeld tij (HWM) en gemiddeld doodtij (HWD). Louter ter duiding van de betekenis van dit deterministische aspect is in de tabel nog per station aangegeven welk percentage van G overeenstemt met HWM; dit percentage schommelt van ong. 60% voor Vlissingen tot zo'n 25% voor Den Helder. Vanzelfsprekend liggen deze percentages hoger voor HWS en lager voor HWD; overigens liggen de grenspeilen van Vlissingen en Hoek van Holland nog achtereenvolgens 75 en 90 cm boven de aangehouden drempels voor HW, vermeld onder par. 3.2.2.1.

Het belang van het deterministische aandeel van de HW's wordt overigens wel wat ingeperkt door de reeds beschreven selectieprocedure d.m.v. drempels, waardoor het waarnemingsmateriaal wordt teruggebracht tot een aantal losse HW's en "plukjes" HW's, van elkaar in de tijd gescheiden door de weggeselecteerde HW's. Toch blijft ook na de hierop volgende tweede selectie, nodig om de autocorrelatie binnen de in eerste aanleg gevormde tijdreeks zo goed mogelijk weg te nemen (methode Daan-i of Daan-S-i), het deterministische aspect der astronomische bestanddelen van HW een niet-verwaarloosbare factor, die uiteraard voor Vlissingen het sterkst doorwerkt en voor Den Helder het minst.

3.2.3. HW-opzet als veranderlijke

3.2.3.1. Algemene overwegingen

In tegenstelling tot HW is HW-opzet, zij het onder zekere beperkingen (par. 3.2.3.2), in belangrijke mate stochastisch van aard. Hoezeer nu HW de invloed van HW-opzet kan verhullen blijkt niet zozeer uit de gegevens van gemiddeld spring- en doodtij (bijlage, staat I) als wel uit die voor de hoogste en laagste astronomische hoogwaterstanden. Volgens par. 3.4.4 zou, in een periode van 19 jaar, van het astronomisch hoogwater de hoogte t.o.v. NAP (berekend volgens de culminatiemethode met 1978...1986 als analyseperiode) voor Vlissingen schommelen tussen +113 en +279 cm en voor Hoek van Holland tussen +57 en +152 cm.

Op grond van deze omstandigheid lijkt het zonder meer voor de hand te liggen de statistische analyses per station te verrichten voor de verzameling van HW-opzetten, welke opzet trouwens de overheersende bijdrage levert aan de zeer hoge standen. Zo'n verzameling moet dan wel homogeen zijn, wat neerkomt op de eis dat de astronomische hoogwaterstanden over de gehele te beschouwen periode op een consistente wijze zijn berekend. Voor de tot voor enige jaren beschikbare astronomische hoog- (en laagwater)-standen

1) dit zijn de HW's, die gemiddeld eens per 2 jaar worden bereikt of overschreden.

gold dit niet of niet voldoende, terwijl er bovendien (vrijwel) geen getijtafels uit de vorige eeuw bestonden. Ten behoeve van het onderhavige onderzoek zijn de 5 betrokken stations de astronomische hoog- en laagwaterstanden voor de periode 1885-1985 dan ook opnieuw berekend en wel d.m.v. de culminatiemethode onder toepassing van opeenvolgende negenjarige analyseperiode (lit. 13, par. 2.4.1 en bijl. 3). Voordelen van de zo verkregen reeksen zijn:

- a door de toepassing van de onderscheidene analyseperioden is de verstorende werking der zeespiegelrijzing geneutraliseerd;
- b de negenjarige analyseperioden zijn zo gekozen dat zij gelet op het nodale tij steeds representatief zijn;
- c een eventuele eerste selectie (weglaten niet-passende lage waarden) kan nu worden verricht met slechts één drempel, waardoor er meer statistisch materiaal beschikbaar is dan bij de HW-reeksen (zie overigens opmerking aan het slot van deze paragraaf);
- d m.b.t. de sterk verstorende weerkundige omstandigheden is het gegevensmateriaal veel representatiever dan dat van HW;
- e HW-opzet is weliswaar niet autocorrelatievrij (zie onder), maar de selectie om de autocorrelatie te verwijderen (Daan-I of "Daan S") zal toch beter zijn dan voor HW omdat nu de deterministische bepaaldheid der astronomische hoogwaterstanden (vrijwel) niet verstorend werkt (zie ook het slot van par. 3.2.2.2 en hierna onder 2e).

Aan het gebruik van HW-opzet kleven echter ook enkele bezwaren, terwijl bij verwerking van deze grootheid op een aantal punten (par. 3.2.3.2) klaarheid is vereist. De bezwaren zijn de volgende:

- 1e De astronomische hoog- en laagwaterstanden, berekend m.b.v. de culminatiemethode, zijn geen "zuivere astronomische getijgegevens", maar zijn t.o.v. die -theoretische- waarden behept met onbekende fouten (lit. 13, par. 2.1.2). Nu bepaalt de onbekende onnauwkeurigheid van de astronomische hoogwaterstanden tezamen met de onnauwkeurigheid van HW de onnauwkeurigheid van HW-opzet; bij gevolg zal de onnauwkeurigheid van HW-opzet aanmerkelijk groter zijn dan die van HW.
- 2e De grootheid HW-opzet voor een willekeurig station bezit binnen de deelreeksen der stormseizoenen ongetwijfeld autocorrelatie, die gevolg is van enkele, in belang en in onderlinge verhouding, wisselende factoren (lit. 10).

Ter toelichting het volgende:

ad 1e: bij de culminatiemethode sorteert men voor elk der 4 parameters (tijdstippen en standen van hoog- en laagwater) de waarden, opgetreden in de analyseperiode, per kalendermaand afzonderlijk voor onder- en bovenmaansdoorgang. Vervolgens worden van elk der zo ontstane (4 x 12 x 2) deelverzamelingen de waarden ingedeeld in klassen ter lengte van één (culminatie)uur; die hoog- resp. laagwaterstanden, die per klasse meer dan 50 cm van het desbetreffende klassegemiddelde afwijken, worden tenslotte uit de klasse verwijderd (lit. 15, par. 5.3.2). Op deze wijze elimineert men de door weersinvloeden sterk afwijkende gevallen; afziende van verdere bijzonderheden der methode kan men stellen dat de berekende astronomische HW-standen klaarblijkelijk betrekking hebben op een onbekende "toestand van middelbaar weer", die voor de onderscheidene maanden nog kan verschillen. Dit bezwaar laat zich niet ondervangen door de harmonische analyse, de andere beschikbare getijtafelberekeningsmethode, waarbij alle uurstanden uit de analyseperiode worden

gebruikt bij vereffeningsprocedures ²⁾. Deze methode ligt ten grondslag aan de uitgaven vanaf 1986 der "Getijtafels voor Nederland", maar is voor het onderhavige onderzoek niet gebruikt (hiermee zou namelijk vanwege de opgetreden veranderingen in het verticale getij - par. 3.2.2.1, ad 2^e - een uitermate grote arbeidsinspanning zijn gemoeid). Voor de goede orde: de verschillen tussen de uitkomsten van beide methoden wijzen eens te meer op het onbepaalde in het begrip "toestand van middelbaar weer" en lijken vergelijking van uitkomsten van het onderhavige onderzoek met waarden in de "Getijtafels voor Nederland 1986" en latere uitgaven voorshands uit te sluiten.

ad 2e: het verband tussen opeenvolgende HW-opzetten is op zich al veranderlijk en mede daardoor ingewikkeld; de relatief grote onnauwkeurigheid van HW-opzet, samenhangend met de onbekende en verschillende "toestanden van middelbaar weer", vormt daarbij nog een verdere complicatie. Dit versterkt de slotsom van lit. 10 dat "er heel weinig reden is aan te nemen dat een "waterdichte" (Daan-) selectie (ter onderdrukking van de autocorrelatie) tot de mogelijkheden behoort".

3.2.3.2. Te stellen eisen aan HW-opzet

Bij de keuze van HW-opzet als statistisch veranderlijke leiden de uit te voeren verwerkingen tot de vaststelling van de kansdichtheids- en verdelingsfuncties van deze grootheid per station. Deze functies dienen per station gekoppeld te worden aan de overeenkomstige (deterministische) functies der astronomische hoogwaterstanden. Een voorwaarde daarbij is de onafhankelijkheid van HW-opzet van het astronomisch getij. Op grond van een eerste onderzoek ter zake (alleen voor Vlissingen) werd eerder voorlopig aangenomen dat aan deze eis wordt voldaan (lit. 13, par. 2.4.1, blz. 12), waarbij overigens uitbreiding van dit deelonderzoek voorzien werd (lit. 13, par. 5, bl. 30, ad 4^e).

Deze voorgenomen werkzaamheden, betrekking hebbende op de 5 te onderzoeken stations plus IJmuiden, vonden inmiddels een afronding ³⁾; het verslag hiervan (lit. 3) voert tot de volgende conclusies:

- a de gemiddelde HW-opzet per culminatieuur is het kleinst voor de beide culminatieuren van spring- en doottij en het grootst omstreeks de kwartieren; de verschillen wisselen daarbij van enkele cm voor de Waddenzeestations via 5 cm bij IJmuiden en Hoek van Holland tot 12 cm bij Vlissingen (lit. 3, bijl. 5);
- b men kan de getijden naar grootte rangschikken volgens de maansdoorgangsuren in drie groepen, die hier achtereenvolgens springtij-, resttij- en doottijgroep worden genoemd en die elk één derde der 2 x 12 (klok)uren betreffen (factor 2 vanwege de onder- en bovensculminatie). De aantallen verschillen per groep (volgens het onderzoek voor de periode 1901...1980)

2) Lit. 3, par. 7, bl. 6/7 maakt aannemelijk dat de harmonische analyse in dit opzicht een betere benadering biedt; echter de daar (bijlagen 6, 7, 8) gemaakte vergelijking is niet helemaal zuiver.

3) perioden: 1933...1980 voor Den Helder en Harlingen;
1901...1950 en 1960...1980 voor Delfzijl;
1901...1980 voor de overige.

enigszins; voor ieder station blijken nu de overschrijdingslijnen van HW-opzet per groep onderling niet verwaarloosbare verschillen te vertonen voor $P < 2,5 \text{ à } 5 \cdot 10^{-1}$ p.j. (voor Hoek van Holland voor $P < 7 \cdot 10^{-2}$ p.j.). In de verschillen zit geen vast stramien (lit. 3, bijlagen 10 a...f): grosso modo overheersend kunnen de HW-opzetten van de springtijgroep zijn (IJmuiden, Den Helder, Harlingen), maar ook die van de doottijgroep (Vlissingen) en evenzeer kunnen zowel de HW-opzetten van de resttijgroep (IJmuiden, Den Helder) als die van de doottijgroep (Harlingen) in het algemeen achterblijven bij de andere. Ook een abrupte overgang (Delfzijl bij $P = 10^{-1}$ p.j.) komt trouwens voor.

Opmerking: de verschillen in gemiddelde HW-opzet tussen de onderscheiden tijgroepen (ad b) zijn ongeveer de helft van die tussen de naar culminatieuur gesorteerde getijden (zie onder a).

De bij deze verwerkingen gebezigde reeksen HW-opzetten zijn behoudens een drempel van 50 cm ongeselecteerd en hebben betrekking op het gehele kalenderjaar; de verschillen tussen de overschrijdingslijnen zijn verder misschien niet verontrustend groot gelet op de onstelselmatigheid in de lijnenpatronen. Toch lijkt het geboden het onderzoek te herhalen voor het stormseizoen en met de volgens de te bezigen Daan-methode geselecteerde waarden (lit. 3, par. 10). Daarbij verdient overigens, gelet op de zojuist gemaakte opmerking, een verdergaande onderverdeling der getijden dan die der drie getijgroepen ad b overweging. Het gevaar op deze wijze per geval te weinig elementen over te houden moet dan in eerste instantie maar voor lief worden genomen.

3.2.4. HW-opzet als enige opwaaiingsparameter

De praktische onmogelijkheid de harmonische analyse toe te passen op de volledige periode 1885...1985 (par. 3.2.3.1. onder ad 1^e) beperkt de gegevens over de opwaaiing tot HW-opzet en LW-opzet. Zouden beide voor een willekeurig station tot dezelfde verzameling behoren dan zou - gezien de duur van het opwaaiingsverschijnsel bij flinke opzetten en de eis van minimalisering der autocorrelatie - medegebruik van LW-opzet weliswaar de omvang van het waarnemingsverschijnsel niet belangrijk vergroten, maar wel zou men zo de representativiteit der verzamelingen van extreme waarden kunnen verbeteren.

Onderzoek naar de wisselwerking tussen opzet en getij over de periode 1933...1980 (voor Delfzijl: 1971...1978) leidde in 1985 tot de slotsom (lit. 14) dat bij Vlissingen gemiddeld genomen HW-opzet $\sim 0,75$ LW-opzet, maar dat vrijwel overal elders, althans bij benadering, gemiddeld zou gelden HW-opzet = LW-opzet. Op grond van de uitkomst voor Vlissingen is daarop besloten alleen HW-opzet in beschouwing te nemen en LW-opzet verder niet te gebruiken (lit. 13, par. 2.4.1. blz. 11 en par. 5, gevolgtrekking b, blz. 29). Nadien is het onderzoek van lit. 8 uitgebreid tot IJmuiden (1901...1980) en de eerdergenoemde 71 jaar voor Delfzijl (lit. 3, par. 3, blz. 1/2; bijlagen 1, 2). De bevindingen van beide studies gezamenlijk luiden nu:

gemiddeld geldt te

Vlissingen en Delfzijl : LW-opzet $\sim 1,33$ resp. $1,2$ HW-opzet;
Hoek van Holland en Den Helder: LW-opzet \sim HW-opzet;
IJmuiden en Harlingen : LW-opzet $\sim 0,93$ resp. $0,95$ HW-opzet.

Dat de verhoudingen zo uiteenlopen is opvallend; klaarblijkelijk zijn de waterdiepten niet alles-bepalend (anders zou immers overal LW-opzet $>$ HW-opzet). Die verhoudingen hangen dan ongetwijfeld mede samen met de mate van

yervroeging van het getij en kombergingsverhoudingen. Volledigheidshalve wordt trouwens nog vermeld dat voor $P < 2 \cdot 10^{-1}$ p.j. ook voor Hoek van Holland geldt LW-opzet $>$ HW-opzet (lit. 14, bijlage 17).

De nieuwe uitkomsten bevestigen nogmaals de juistheid der eerder genomen beslissing alleen HW-opzet als opwaaiingsparameter te gebruiken. Het onderzoek kan op dit punt als afgesloten worden beschouwd.

3.3 Koppeling HW-opzetten en astronomische standen

De procedure zoals die nu gevolgd is leidt tot een schatting van de kansverdeling van het jaarmaximum van de HW-opzetten, i.h.b. van de vormparameter van de benaderende EV-kansverdeling. We willen echter een schatting hebben van de kansverdeling van het jaarmaximum van de HW's zèlf. Daarvoor hebben we behalve de genoemde schatting voor de opzetten ook de getijtafel tot onze beschikking.

Om tot een koppeling van beide te komen dienen we het volgende te overwegen. Aan te nemen valt dat de kansverdeling van de hoogte van een HW-opzet niet afhangt van het astronomisch tij op dat tijdstip. We veronderstellen verder dat na de selectieprocedure de HW-opzetten te beschouwen zijn als onafhankelijke trekkingen uit één kansverdeling. Verder worden dankzij de selectieprocedure maar weinig opzetten gebruikt voor de schattingsprocedure. We kunnen ons nu voorstellen dat deze (geselekteerde) opzetten op willekeurige manier toegewezen worden aan de dagen van de beschouwde winterperiode. D.w.z. uit de verzameling dagen van een winterperiode nemen we een trekking zonder teruglegging en beschouwen dat als de dagen waarop de geselekteerde HW-opzetten plaatsvinden. Dit is een gedachte-rekonstruktie van het kansmodel waaruit de waarnemingen komen. Aangezien er maar weinig dagen per winter mee doen, is trekking zonder teruglegging nauwelijks te onderscheiden van trekking met teruglegging d.w.z. aan alle geselekteerde HW-opzetten wijzen we willekeurig (overeenkomstig frequentie van voorkomen) een astronomische stand toe. Dit geeft het volgende model. Laten X_1, X_2, \dots de geselekteerde HW-opzetten zijn en Y_1, Y_2, \dots de bijbehorende astronomische standen. Dan zijn zowel de X-en als de Y's stochastische grootheden, de X-en hebben alle dezelfde kansverdeling en de Y's ook; verder zijn $X_1, Y_1, X_2, Y_2, X_3, Y_3, \dots$ onafhankelijk. We zijn geïnteresseerd in een schatting van de kansverdeling van de maxima van de rij $X_1 + Y_1, X_2 + Y_2, \dots$ n.l. de HW's zèlf. We hebben een dergelijke schatting voor de rij X_1, X_2, \dots en ook de rij Y_1, Y_2, \dots (de laatste via de getijtafel).

Laten we nu veronderstellen (wat waarschijnlijk lijkt) dat zowel k_1 (de vormparameter geschat via de X-en) als k_2 (de vormparameter geschat via de Y's) positief zijn. Laat F_1 de kansverdeling zijn van X_1 en F_2 die van Y_1 . Als $k_1, k_2 > 0$, dan bestaan er absolute bovengrenzen voor X_1 en Y_1 . Laten we die bovengrenzen beide nul nemen (dit is alleen voor 't gemak van schrijven). De kansverdeling van $X_1 + Y_1$ is de convolutie van F_1 en F_2 die aangegeven wordt als $F_1 * F_2$. Men kan nu aantonen dat als $k_1, k_2 > 0$,

$$\lim_{t \rightarrow 0} \frac{1 - F_1 * F_2(t)}{(1 - F_1(t))(1 - F_2(t))} = \frac{\Gamma(1 + k_1^{-1})\Gamma(1 + k_2^{-1})}{\Gamma(1 + k_1^{-1} + k_2^{-1})}$$

waarbij het voornamelijk van belang is dat het rechterlid positief en eindig is (de Haan 1986). Dat wil zeggen dat maxima uit de kansverdeling $F_1 * F_2$ waarin we geïnteresseerd zijn, zich net zo gedragen als maxima uit de kansverdeling

$$1 - (1 - F_1(t))(1 - F_2(t))$$

d.w.z. de kansverdeling van $Z_1 = \min(X_1, Y_1)$. Dit is een belangrijk gegeven. Het zou te ver voeren om de technische details te geven maar dit betekent i.h.b. dat $(k_1^{-1} + k_2^{-1})^{-1}$ de vormparameter is van de EV-verdeling behorende bij de waarnemingen $X_1 + Y_1, X_2 + Y_2, \dots$ d.w.z. de astronomische standen. Het koppelingsprobleem is daarmee opgelost.

Het is tussen haakjes niet zo moeilijk de bovenstaande redenering te verifiëren voor 't geval $F_1(x) = (-x)^{-k}$ voor $x < 0$ ($i = 1, 2$).

3.4. Beschikbaarheid en verdeling van de astronomische HW-standen.

3.4.1. Inleiding.

De aanpak waarbij de HW-opzetten gebruikt worden als basiswaarnemingen voor de statistische analyse vereist dat de verzamelingen waargenomen hoogwaterstanden voor elk station worden omgerekend tot verzamelingen HW-opzetten. Daartoe wordt van elke opgetreden hoogwaterstand de bijbehorende berekende astronomische hoogwaterstand afgetrokken, ongeacht een vervroeging of verlatting van het tijdstip van hoogwater. In het hierna volgende wordt aangegeven hoe de astronomische hoogwaterstanden zijn berekend.

Bij de opsomming van de analyse- en bijbehorende berekeningsperioden worden begin- en einddatum van de waarnemingsreeksen genoemd. Deze worden daarom eerst besproken.

Voor de koppeling van de verdeling van extreme HW-opzetten en astronomische hoogwaters, immers het gaat uiteindelijk om uitspraken over de overschrijdingskansen van extreme hoogwaterstanden, dienen de verdelingen van de astronomische hoogwaterstanden der verschillende stations gegeven te zijn. Daarbij mag er van uit worden gegaan dat de HW-opzetten en de astronomische HW's onafhankelijk zijn (par. 3.5). Hoewel het astronomisch getij deterministisch is bepaald, kan zij in het gedachtenmodel als stochastische grootheid worden opgevat, omdat elke combinatie van HW-opzet en HW-astronomisch even waarschijnlijk is.

3.4.2. De waarnemingsperiode.

De uitkomsten van de statistische analyse zullen betrouwbaarder zijn naarmate het aantal waarnemingen waarop zij gebaseerd is groter is. Daarom wordt voor elk station een zolang mogelijke aaneengesloten waarnemingsreeks gebruikt. De HW-standen zijn voor de verschillende stations in volledige en bruikbare vorm thans beschikbaar vanaf:

| | | |
|------------------|---------|-------|
| Vlissingen | 1 juli | 1881 |
| Hoek van Holland | 1 aug. | 1887 |
| Den Helder | 1 juni | 1932 |
| Harlingen | 1 juni | 1932 |
| Delfzijl | 1 maart | 1881. |

De HW-opzetten van voor de afsluiting der Zuiderzee (28 mei 1932) te Den Helder en Harlingen hadden een aanzienlijk lagere overschrijdingsfrequentie en zijn derhalve niet bruikbaar.

De andere data geven niet het begin van de waarnemingsreeksen, maar het tijdstip van waaraf bijgissing van hiaten m.b.v. waarnemingen aan omliggende stations met voldoende betrouwbaarheid mogelijk was.

Het einde van de waarnemingsperiode betreft in principe het laatste jaar waarover definitief vastgestelde gegevens beschikbaar zijn. Voor dit onderzoek is hiervoor het jaar 1985 aangehouden.

3.4.3. De berekening van de astronomische hoogwaterstanden.

De astronomische hoogwaterstanden van de vijf stations zijn in principe beschikbaar in de vorm van getijtafels, waarin de voorspelde hoogten en tijdstippen van hoog- en laagwater zijn opgenomen. De getijtafels voor Nederland t/m 1985 werden berekend volgens de methode van de culminatie-analyse. Deze methode levert alleen de hoog- en laagwaterstanden en niet de gehele getijkromme, dit in tegenstelling tot de voor de getijtafels vanaf 1986 gebruikte methode der harmonische analyse.

Omdat in het verleden de toegepaste culminatiemethode enige malen is herzien, zijn, teneinde homogeen materiaal te verkrijgen, de astronomische standen opnieuw berekend volgens de laatst toegepaste culminatiemethode (lit. 15). Toepassing van de harmonische analyse hiervoor zou buitensporig veel werk met zich meebrengen, omdat dan de oude peilbladen gedigitaliseerd zouden moeten worden op uurstanden. Voor de getijstations zijn de uurstanden in het elektronisch databestand opgeslagen vanaf 1971; van de periode daarvoor zijn alleen hoog- en laagwaterstanden in het bestand opgenomen.

De voor de culminatiemethode gebezigde negenjarige (ongeveer de halve knopen-cyclus) analyseperioden zijn wanneer mogelijk zo gekozen dat halverwege deze perioden het maansbaanvlak de middenstand innam (schommeling maansdeclinatie). Discontinuïteiten t.g.v. uitgevoerde werken doorkruisen in een enkel geval deze gedragslijn.

Bij de ontkoppeling van astronomisch hoogwater en HW-opzet is gebruik gemaakt van bestaande programmatuur voor analyse en voorspelling m.b.v. de culminatiemethode, welke niet voorzag in toepassing op andere eeuwen dan de 20^e. Deze is met enige aanpassingen hiertoe bruikbaar gemaakt. Het was echter niet zonder onevenredig veel inspanning mogelijk

- december 1899 en het jaar 1900 met culminatie-analyse te voorspellen;
- een periode die over een eeuwgrens valt te analyseren.

Tabel 3.1 geeft een overzicht van de gehanteerde analyseperioden en -methoden.

3.4.4. Frequentieverdeling van de astronomische hoogwaters in de huidige toestand

De bij de koppeling van het astronomisch hoogwater en de HW-opzet te gebruiken kansverdeling van de astronomische hoogwaterstanden dient karakteristiek te zijn voor hetzelfde stormseizoen als bij de analyse van HW-opzetten wordt gebruikt (par. 4.2). Op basis van de culminatie-analyses over 1978...1986 (voor Delfzijl 1979...1986) zijn (pseudo-) astronomische hoogwaters over de periode 1901...1994 voorspeld (93 lange stormseizoenen). De analyseperiode kan geacht worden de toestand van voor 1985 voldoende te representeren. Er is weer gebruik gemaakt van de culminatiemethode, waardoor kleine systematische afwijkingen inherent aan deze methode (lit. 3) gecompenseerd worden. Het aantal 93 is ingegeven door het feit dat dit een veelvoud is van 18,6 jaar (invloed helling maansbaan), en van dezelfde grootte-orde is als het aantal jaren van de langere waarnemingsreeksen. Een kortere reeks van astronomische hoogwaterstanden (19 stormseizoenen) zal overigens een verdeling opleveren die niet veel afwijkt van die op basis van de 93 seizoenen.

De kansdichtheden van de astronomische hoogwaterstanden in het lange stormseizoen zijn weergegeven in de figuren 3.1...3.5. Ze zijn bepaald door de frequenties van voorkomen in de 93 seizoenen per klasse met een breedte van 1 cm zodanig te normeren dat de gesommeerde kansdichtheid juist 1 is.

De hoogste en laagste in de 93 seizoenen voorkomende astronomische hoogwaterstanden zijn voor de verschillende stations in onderstaande tabel weergegeven.

Hoogste en laagste astr. HW's in de 93 seizoenen in cm's:

| station | hoogste HW-astr. | laagste HW-astr. |
|------------------|------------------|------------------|
| Vlissingen | 279 | 113 |
| Hoek van Holland | 152 | 57 |
| Den Helder | 97 | 22 |
| Harlingen | 139 | 40 |
| Delfzijl | 183 | 83 |

De hoogste HW's van de stormseizoenen zijn tevens de hoogste van het hele jaar; de laagste HW's zijn daarentegen beduidend hoger dan de minima voor het hele jaar, die buiten het stormseizoen (n.l. in april-mei) plegen te vallen. De verdelingen over het stormseizoen verschillen ook belangrijk van die van de astronomische HW's over het hele jaar.

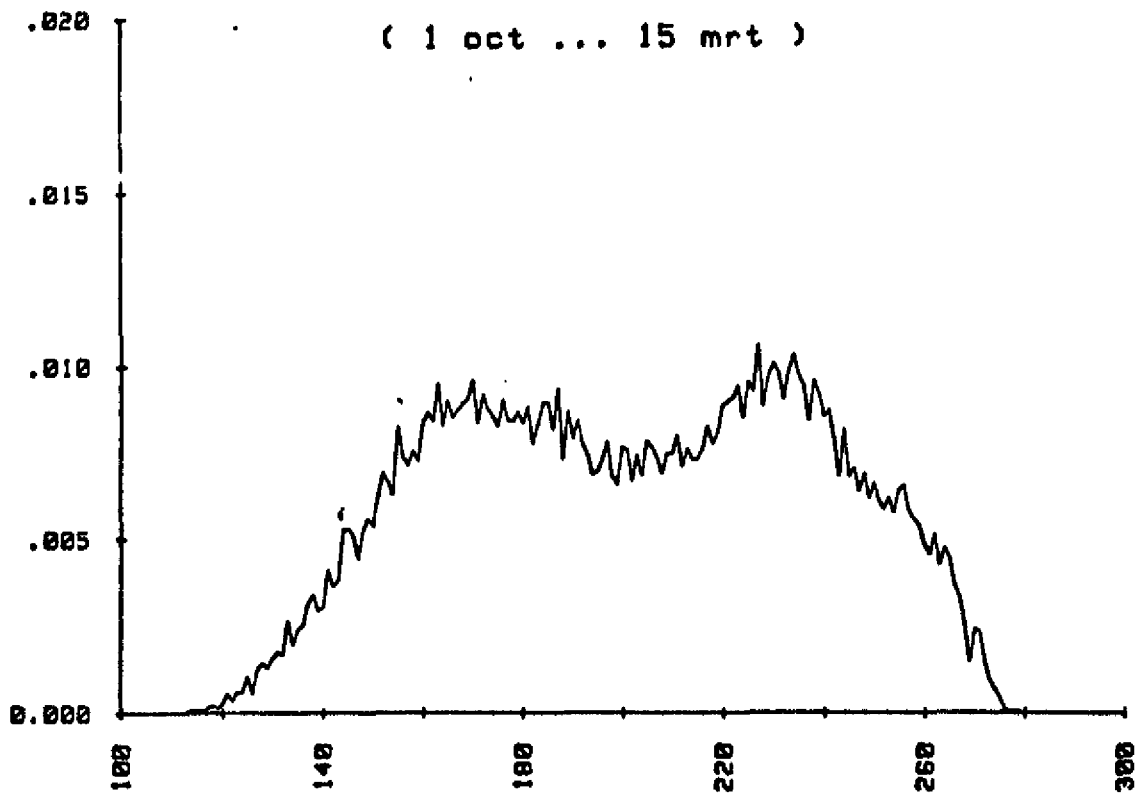
De springtij/doodtij-cyclus (en voor Den Helder en Harlingen de dagelijkse ongelijkheid) veroorzaakt een duidelijke tweetoppigheid.

Tabel 3.1

Gehanteerde methoden, analyse- en berekeningsperioden voor het astron. HW

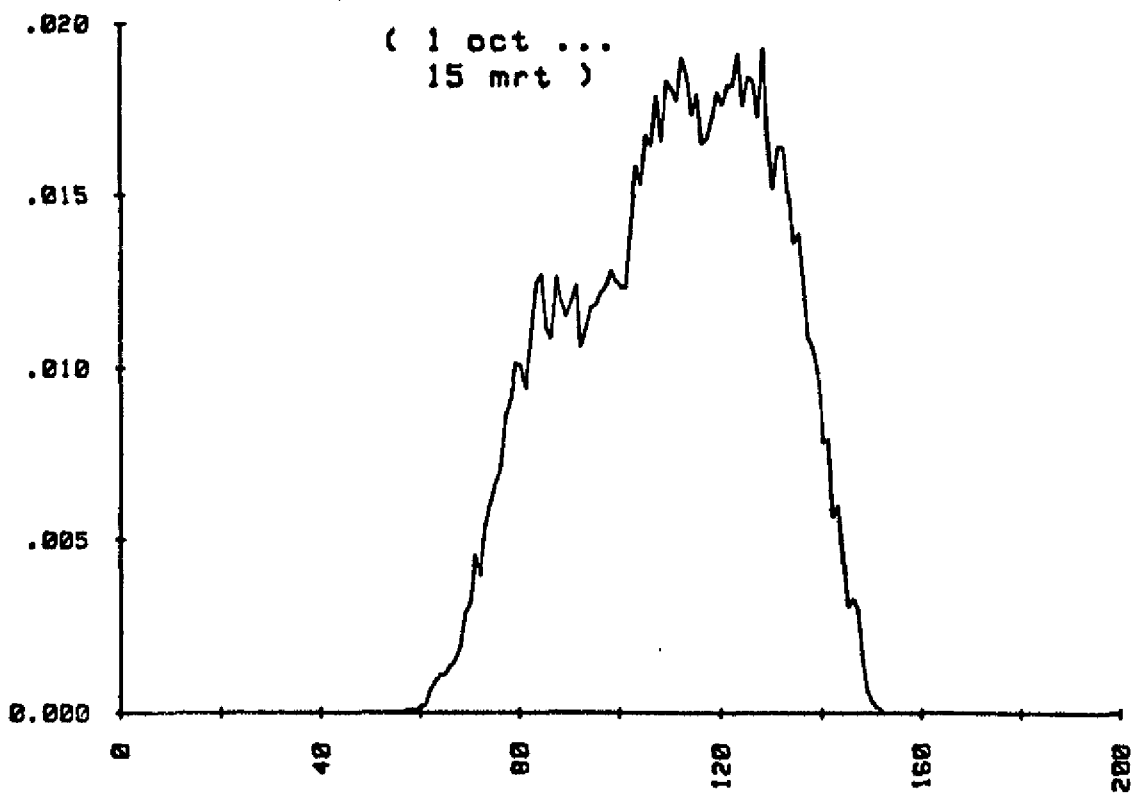
| station | berekeningsperiode | analyseperiode | berekeningsmethode | |
|------------------|---------------------|------------------|--|----|
| Vlissingen | 1-07-1881... | 1885...1893 | culminatie-analyse | |
| | 30-11-1899 | | | |
| | 1-12-1899...1900 | 1900 | harmonische analyse op uurwaarden | |
| | 1901...1913 | 1905...1913 | culminatie-analyse | |
| | 1914...1922 | 1914...1922 | .. | |
| | 1923...1931 | 1923...1931 | .. | |
| | 1932...1941 | 1932...1941 | .. | |
| | 1942...1950 | 1942...1950 | .. | |
| | 1951...1959 | 1951...1959 | .. | |
| | 1960...1968 | 1960...1968 | .. | |
| | 1969...1977 | 1969...1977 | .. | |
| | 1978...1985 | 1978...1986 | .. | |
| | Hoek van Holland | 1-08-1887... | 1888...1896 | .. |
| 30-11-1899 | | | | |
| 1-12-1899...1900 | | 1900 | harmonische analyse op halfuurwaarden | |
| 1901...1913 | | 1905...1913 | culminatie-analyse | |
| 1914...1922 | | 1914...1922 | .. | |
| 1923...1932 | | 1923...1931 | .. | |
| 1933...1941 | | 1933...1941 | .. | |
| 1942...1950 | | 1942...1950 | .. | |
| 1951...1964 | | 1951...1959 | .. | |
| 1965...1977 | | 1969...1977 | .. | |
| 1978...1985 | | 1978...1986 | .. | |
| Den Helder | | 1-06-1932...1941 | 1933...1941 | .. |
| | | 1942...1950 | 1942...1950 | .. |
| | 1951...1959 | 1951...1959 | .. | |
| | 1960...1968 | 1960...1968 | .. | |
| | 1969...1977 | 1969...1977 | .. | |
| | 1978...1985 | 1978...1986 | .. | |
| Harlingen | 1-06-1932...1941 | 1933...1941 | .. | |
| | 1942...1950 | 1942...1950 | .. | |
| | 1951...1959 | 1951...1959 | .. | |
| | 1960...1968 | 1960...1968 | .. | |
| | 1969...1977 | 1969...1970 | .. | |
| | 1978...1985 | 1978...1986 | .. | |
| Delfzijl | 1-03-1881... | 1885...1893 | .. | |
| | 30-11-1899 | | | |
| | 1-12-1899...1900 | - | oude getijtafels | |
| | 1901...1913 | 1905...1913 | culminatie-analyse | |
| | 1914...1922 | 1914...1922 | .. | |
| | 1923...1932 | 1923...1932 | .. | |
| | 1933...1941 | 1933...1941 | .. | |
| | 1942...1950 | 1942...1950 | .. | |
| | 1951...1959 | 1951...1959 | .. | |
| | 1960...1968 | 1960...1968 | .. | |
| | 1969...1978 | 1969...1977 | .. | |
| | 1979...1985 | 1978...1986 | .. | |

Astronomisch HW in stormseizoen
(1 oct ... 15 mrt)



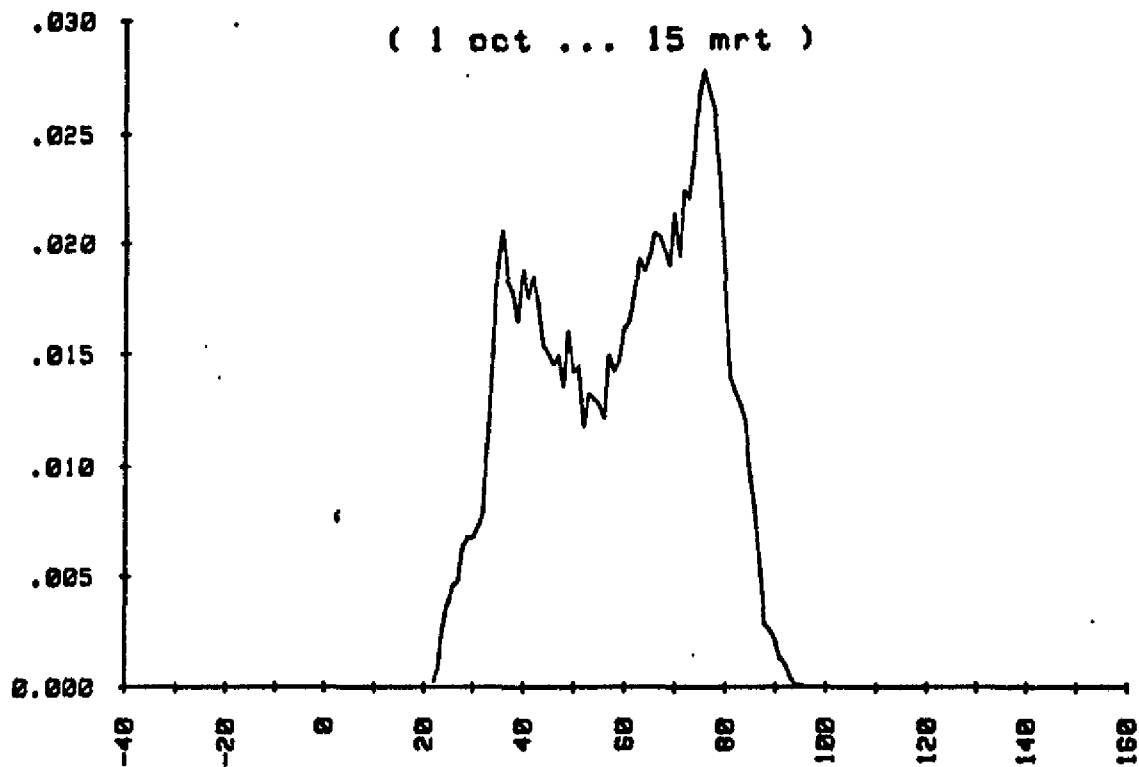
Figuur 3.1 VLISSINGEN : HW in cm + NAP

Astronomisch HW in stormseizoen
(1 oct ...
15 mrt)



Figuur 3.2 HOEK VAN HOLLAND : HW in cm + NAP

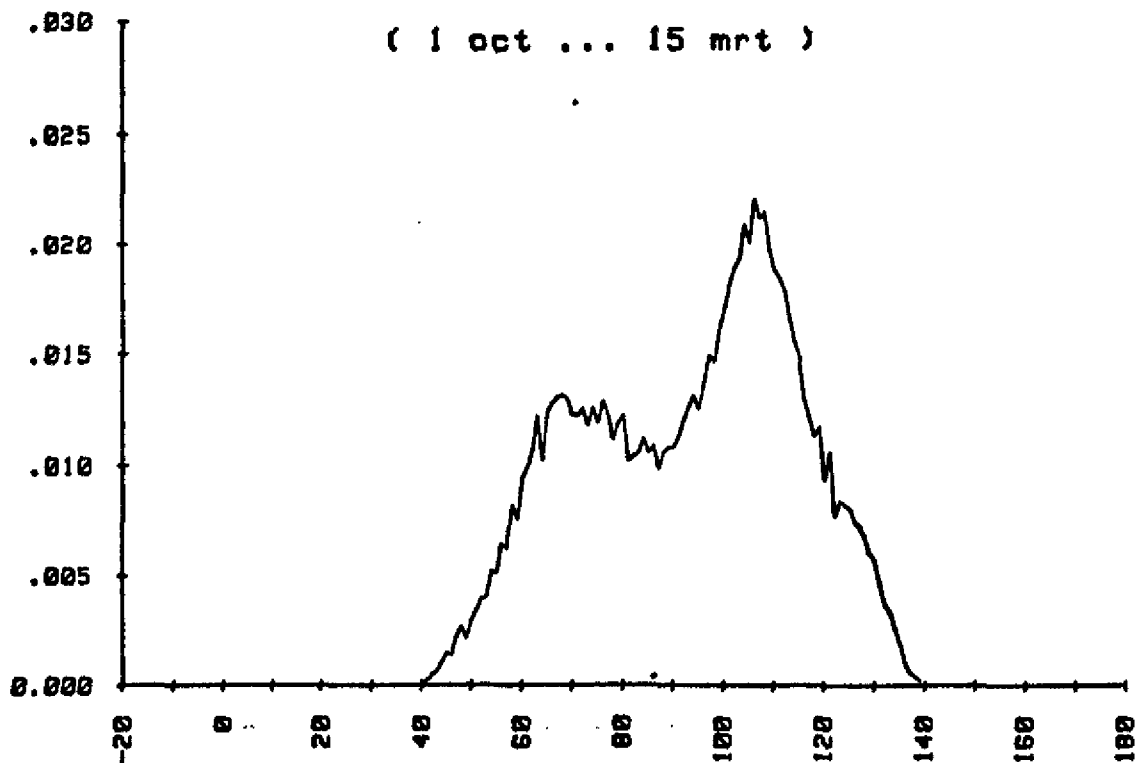
Astronomisch HW in stormseizoen
(1 oct ... 15 mrt)



Figuur 3.3

DEN HELDER

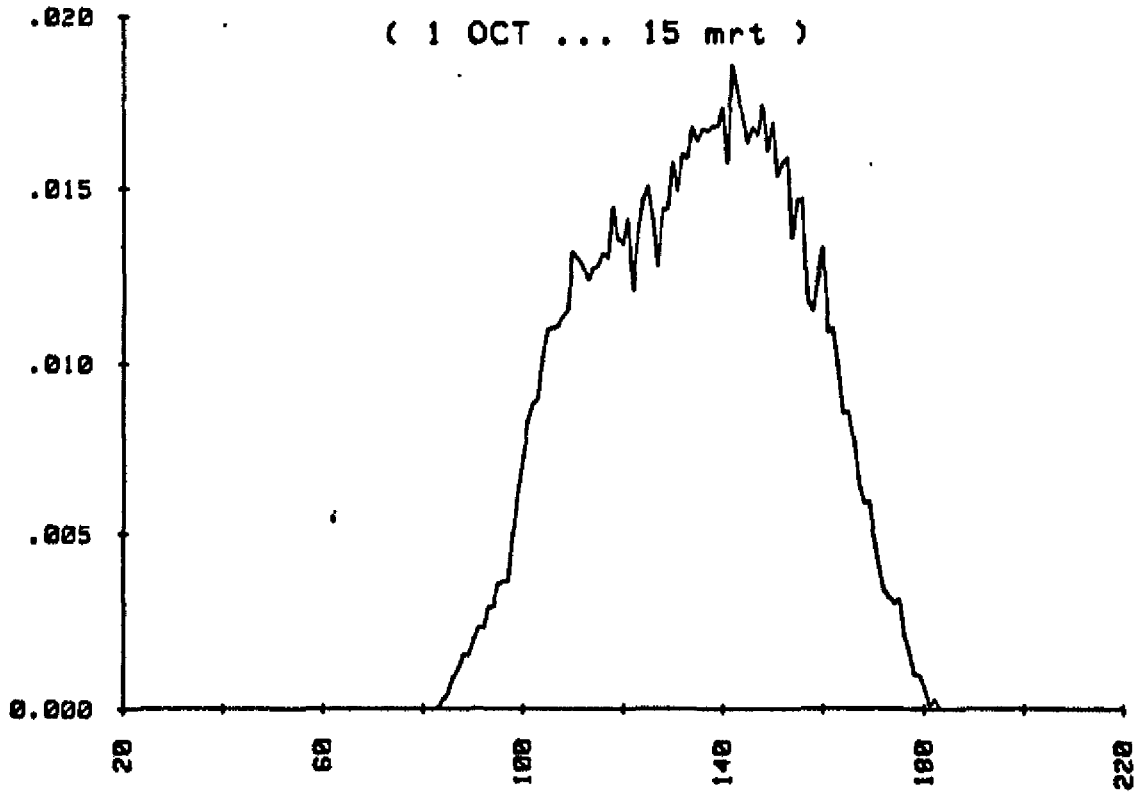
Astronomisch HW in stormseizoen
(1 oct ... 15 mrt)



Figuur 3.4

HARLINGEN

Astronomisch HW in stormseizoen
(1 OCT ... 15 mrt)



Figuur 3.5 DELFZIJL : HW in cm + NAP

3.5 Toetsing van de onafhankelijkheid van astronomisch HW en HW-opzet

3.5.1. Inleiding

Bij het onderzoek naar de overschrijdingskansen van extreme waterstanden heeft het gebruik van de HW-opzet als basiswaarneming voor de statistische analyse de voorkeur (par. 3.2). De HW-opzet is daarbij gedefiniëerd als het verschil tussen de opgetreden hoogwaterstand en de voor het betrokken hoogwater voorspelde astronomische hoogwaterstand, ongeacht een eventuele vervroeging of verlatening van het tijdstip van het werkelijke hoogwater t.o.v. dat van het astronomische hoogwater. Omdat het uiteindelijk gaat om de overschrijdingskansen van hoogwaterstanden, zullen de resultaten van de statistische analyse van de HW-opzetten weer "vertaald" moeten worden naar de som van astronomische HW-standen en HW-opzetten.

Voorwaarde voor deze werkwijze is dat de HW-opzetten onafhankelijk zijn van de astronomische hoogwaterstanden. Deze paragraaf beschrijft de statistische toetsing van de aanname van onafhankelijkheid.

3.5.2. De Spearman toets voor de hypothese van onafhankelijkheid.

Een toets die zich goed leent voor het toetsen op onafhankelijkheid van twee reeksen waarnemingen, gegeven in de vorm van N getallenparen, is de Spearman toets (lit. 5). De Spearman toets is een niet-parametrische of verdelingsvrije toets; de waarnemingen worden vervangen door hun rangnummer in de van klein naar groot gesorteerde reeksen. De toets werkt verder geheel met een functie van de rangnummers. Het voordeel van deze niet-parametrische toets is, dat hij geldig is zonder de gebruikelijke aanname van normaliteit van de verdelingen van de parameters, dit in tegenstelling tot de klassieke toets van de significantie van de steekproef-correlatiecoëfficiënt.

Gegeven zijn de getallenparen $(X_1, Y_1), \dots, (X_N, Y_N)$, waarbij de variabelen X_1 resp. Y_1 onafhankelijk en gelijk verdeeld zijn.

De nulhypothese H_0 is dat X en Y onafhankelijk van elkaar zijn.

Laat nu $R = (R_1, \dots, R_N)$ de vector zijn van de rangnummers corresponderend met $X = (X_1, \dots, X_N)$, en $Q = (Q_1, \dots, Q_N)$ de vector van de rangnummers corresponderend met $Y = (Y_1, \dots, Y_N)$.

De maat die wordt gebruikt bij het toetsen van de nulhypothese van onafhankelijkheid volgens Spearman is

$$S = \sum_{i=1}^N R_i Q_i$$

S is een (niet-genormeerde) maat voor de correlatie op de rangnummers i.p.v. op de variabelen zelf.

Afgeleid kan worden dat onder H_0

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

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

De verdeling van S onder H_0 is symmetrisch en bij benadering normaal voor grote N .

3.5.3. Groepen van gelijke waarnemingen (ties).

Een complicatie in de bovenbeschreven procedure wordt gevormd door het voorkomen van groepen van gelijke waarnemingen (bijvoorbeeld als gevolg van afrondingen), gewoonlijk aangeduid met de korte Engelse benaming "ties". Het is dan niet zondermeer duidelijk hoe de rangnummers en dus ook de scores moeten worden toegekend binnen de ties.

Een methode waarbij de volgorde binnen een tie ad random bepaald wordt heeft als nadeel dat de toets niet exact gelijk verloopt bij herhaling of uitvoering door een ander. Daarom wordt meestal, ook hier, de methode toegepast waarbij aan elke waarneming van een tie het over de groep gemiddelde rangnummer wordt toegekend.

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 :

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

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

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 is weer $S = \sum_{i=1}^N R_i Q_i$

waarbij voor R_i en Q_i bij ties de gemiddelde rangnummers moeten worden ingevuld.

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

3.5.4. Toepassing op de astronomische HW's en HW-opzetten.

De toetsing van de onafhankelijkheid van astronomische HW's en HW-opzetten is uitgevoerd voor de stations Vlissingen, Hoek van Holland en Delfzijl.

De getallenparen bestaan voor elk station uit de geselecteerde HW-opzetten en de bijbehorende astronomische HW's. Door de toegepaste selectiemethode (stormseizoen 1 okt...15 mrt en Daan-S-4-selectie) is voor de HW-opzetten voldaan aan de eis van onafhankelijkheid en gelijke verdeling (zie respectievelijk hoofdstuk 4 en hoofdstuk 5).

Aangenomen mag worden dat stormen geen voorkeur hebben voor een bepaalde getijconstellatie (lit. 3, bijlage 4). We kunnen de bij de geselecteerde HW-

opzetten behorende astronomische HW's dus beschouwen als willekeurige trekkingen uit de populatie van astronomische HW's. Voor de als stochastische grootheid opgevatte astronomische HW's mag er dus eveneens van worden uitgegaan dat ze onafhankelijk en gelijk verdeeld zijn.

Als toetsingsgrootheid is gekozen

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

Onder H_0 heeft \underline{t} voor grote N de standaard-normale verdeling. De resultaten van de berekeningen van \underline{t} zijn samengevat in onderstaande tabel.

| drempel (cm) | Vlissingen 1882...1885 | | Hoek van Holland 1888...1985 | | Delfzijl 1878...1985 | |
|-----------------|---------------------------|-----------------|---------------------------------|-----------------|-------------------------|-----------------|
| | N | \underline{t} | N | \underline{t} | N | \underline{t} |
| 30 | 1660 | +0,14 | 1576 | +0,22 | 1828 | +1,56 |
| 50 | 921 | -0,98 | 977 | +0,32 | 1281 | +1,08 |
| 70 | 447 | -1,21 | 562 | +1,14 | 857 | +0,68 |
| 90 | 223 | -1,50 | 315 | -0,52 | 601 | +0,09 |
| 110 | 99 | -0,54 | 162 | +0,19 | 400 | -0,42 |

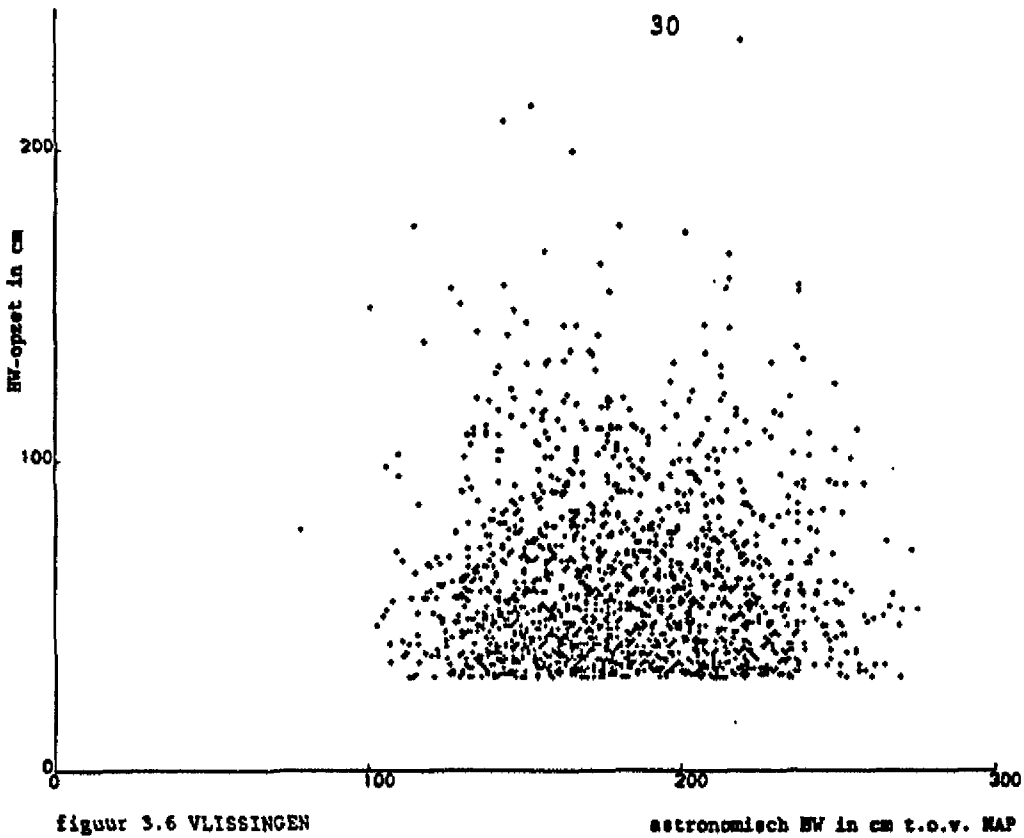
opm.: een stormseizoen wordt aangeduid met het jaar waarin de maand januari valt.

De kritieke gebieden voor \underline{t} bij tweezijdige toetsing en onbetrouwbaarheidsdrempel d van 5% zijn:

- a. $\underline{t} \geq 1,96$
- b. $\underline{t} \leq -1,96$

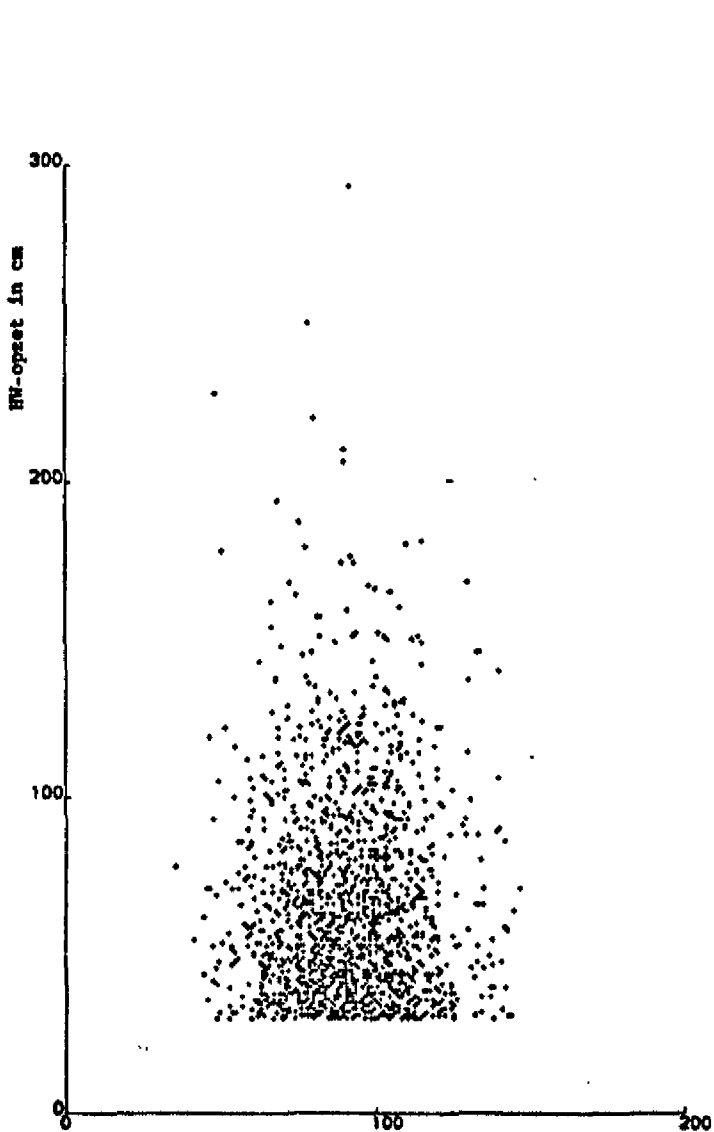
Conclusie: Aangezien in geen van de beschouwde gevallen \underline{t} in het kritieke gebied valt en dus de nulhypothese niet wordt verworpen, mag voor de verdere statistische bewerkingen worden uitgegaan van de onafhankelijkheid van astronomische HW's en HW-opzetten.

Ter illustratie zijn in de figuren 3.6 , 3.7 and 3.8 de getallenparen voor de drie beschouwde stations uitgezet bij een drempelniveau van 30 cm.



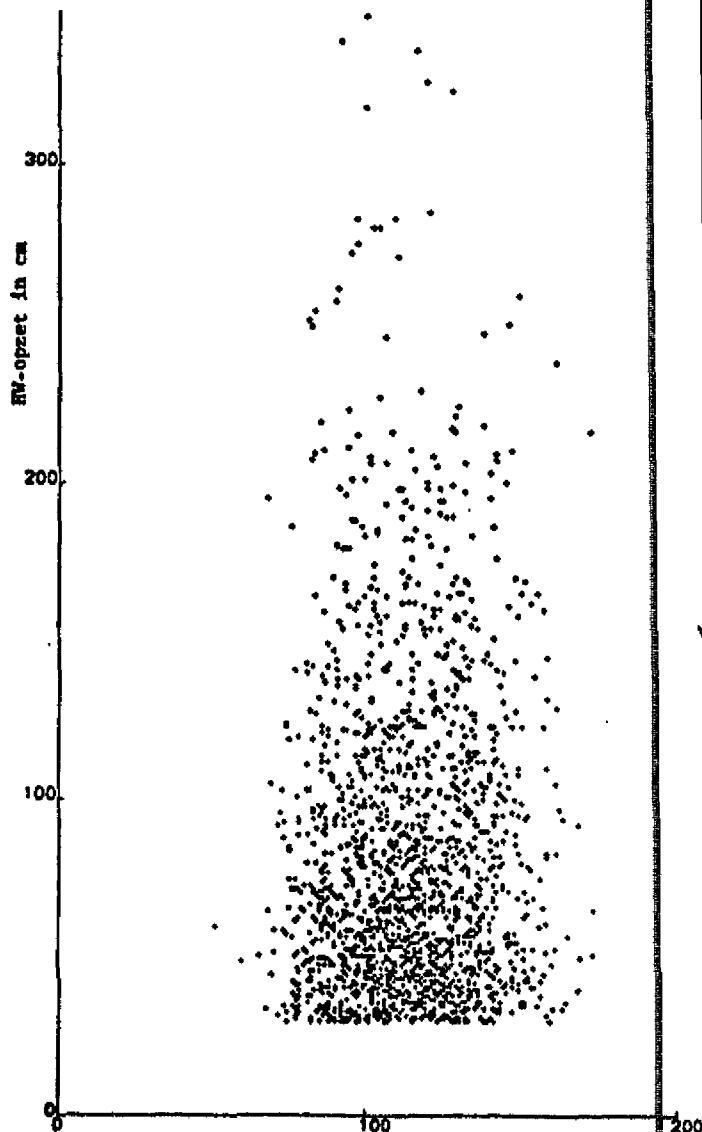
figuur 3.6 VLISSINGEN

astronomisch HW in cm t.o.v. MAP



figuur 3.7 HOEK VAN HOLLAND

astronomisch HW
in cm t.o.v. MAP



figuur 3.8 DELFZIJL

astronomisch HW
in cm t.o.v. MAP

4. WAARNEMINGSPERIODE EN STORMSEIZOEN

4.1 Waarnemingsperiode

Het is van groot belang voor de statistische bewerking en voor het verkrijgen van betrouwbare frequentiekrommen om een zo lang mogelijk reeks gegevens te gebruiken.

De data, sedert welke bruikbare gegevens beschikbaar zijn, zijn voor de verschillende stations als volgt:

| | | |
|------------------|---------|------|
| Vlissingen | 1 juli | 1881 |
| Hoek van Holland | 1 aug. | 1887 |
| Den Helder | 1 juni | 1932 |
| Harlingen | 1 juni | 1932 |
| Delfzijl | 1 maart | 1881 |

Het einde van de gebruikte waarnemingsperiode betreft uiteraard de meest recente winter waarover definitief vastgestelde gegevens beschikbaar zijn. Bij het samenstellen van dit rapport was dit de winter 1984/1985.

4.2 Het stormseizoen.

4.2.1 Inleiding.

Voor een goede extrapolatieprocedure t.a.v. de kans dat bepaalde extreme HW-opzetten worden overschreden is het van belang om uit te gaan van zo homogeen mogelijk waarnemingsmateriaal. Dat houdt in dat getracht wordt alleen die waarnemingen mee te nemen, waarvan we mogen aannemen dat ze onderling onafhankelijk zijn en dezelfde kansverdeling bezitten. Om die reden werd reeds in het Deltarapport onderscheid gemaakt tussen een zomer- en een winterperiode. Als winterperiode werden enigszins arbitrair de maanden november, december en januari gekozen; de waarnemingen uit de zomerperiode bleven buiten beschouwing. Vergroting of verkleining van deze winterperiode, waarin ook de meeste jaarmaxima vallen, bleek de schattingen van de overschrijdingskansen van extreme waterstanden echter niet significant te beïnvloeden.

Aanvankelijk werd bij onderhavige studie hetzelfde stormseizoen genomen als in het Deltarapport met als toevoeging 31 oktober en 1 februari (lit. 13). Toepassing van het uitgebreide stormseizoen van 1 oktober t/m 1 maart bleek echter van niet geringe invloed op de breedte van de betrouwbaarheidsintervallen van de puntschattingen voor het 10^{-4} -kwantiel. Dit was reden om de keuze van het stormseizoen opnieuw te bezien.

4.2.2 Criteria voor de keuze van een stormseizoen.

Het stormseizoen moet zodanig worden gekozen dat alle HW-opzetten bij benadering beschouwd mogen worden als onafhankelijke waarnemingen met dezelfde kansverdeling. De onafhankelijkheid van de HW-opzetten wordt verkregen door het toepassen van een adequate selectiemethode (zie hoofdstuk 5). Het gebruik van de HW-opzet als stochastische variabele waarborgt, behoudens klimaat-schommelingen, gelijke verdeling over de jaren, omdat trends en discontinuïteiten in de hoogwaters zelf geen invloed hebben op deze verdeling. De lengte van het stormseizoen dient dan nog zodanig te worden bepaald, dat de verdelingen van de opgetreden hoogwateropzetten in verschillende kalenderperiodes van bijvoorbeeld een maand (jan., febr., etc. of 15 jan. t/m 14 febr., etc.) bij benadering als gelijk kunnen worden beschouwd.

Het gaat dus voor een goede extrapolatieprocedure niet om het aantal onafhankelijke HW-opzetten per kalenderperiode, maar om het gelijk zijn van de voorwaardelijke verdelingen voor de HW-opzetten boven een bepaalde drempel. Het gemiddelde aantal stormen in bijvoorbeeld de maand oktober mag dan lager zijn dan in december of januari, dat wil nog niet zeggen dat de maand oktober buiten het stormseizoen gehouden moet worden; dat wordt bepaald door de verdeling van de opzetten.

Bij sommige toepassingen is het wél noodzakelijk dat niet alleen de opzetten een gelijke kansverdeling hebben, maar dat ook het aantal waarnemingen per tijdsinterval stationair is. In dat geval zullen de aantallen onafhankelijke HW-opzetten in evenlange kalenderperioden bij benadering gelijk moeten zijn. Deze additionele eis leidt tot een korter stormseizoen welke in het vervolg zal worden aangeduid als het "korte stormseizoen". Paragraaf 5.3 geeft een toepassing van het korte stormseizoen.

4.2.3 Boxplotten.

Voor de vergelijking van de voorwaardelijke verdelingen van de HW-opzetten boven een bepaalde drempel in de verschillende kalenderperioden is gebruik gemaakt van boxplotten (zie fig. 4.1 t/m 4.6). Zie ook lit. 1 en lit. 11. De eigenlijke box telt drie horizontale lijntjes, die de ligging aangeven van respectievelijk het 25%-kwantiel, de mediaan en het 75%-kwantiel. Noemen we de hoogte van het 25%-kwantiel h_1 en de hoogte van het 75%-kwantiel h_2 dan wordt de ondergrens van de streepjeslijn onder de box bepaald door de kleinste waarneming groter of gelijk aan $h_1 - 1.5 * (h_2 - h_1)$. Kleinere waarnemingen worden ieder afzonderlijk met een sterretje weergegeven. Overeenkomstig wordt de bovengrens van de streepjeslijn boven de box bepaald door de grootste waarneming kleiner of gelijk aan $h_2 + 1.5 * (h_2 - h_1)$ en worden grotere waarnemingen weer afzonderlijk met sterretjes weergegeven. Door de hier betrekkelijk willekeurig gekozen ondergrens van 70 cm voor de HW-opzetten bij het maken van de figuren 4.1 t/m 4.6 komen in deze figuren geen sterretjes voor onder de boxen.

4.2.4. Het beschouwde waarnemingsmateriaal van de HW-opzetten

Omdat een definitieve drempelwaarde nog niet kon worden bepaald, is hiervoor enigszins arbitrair de waarde van 70 cm aangehouden.

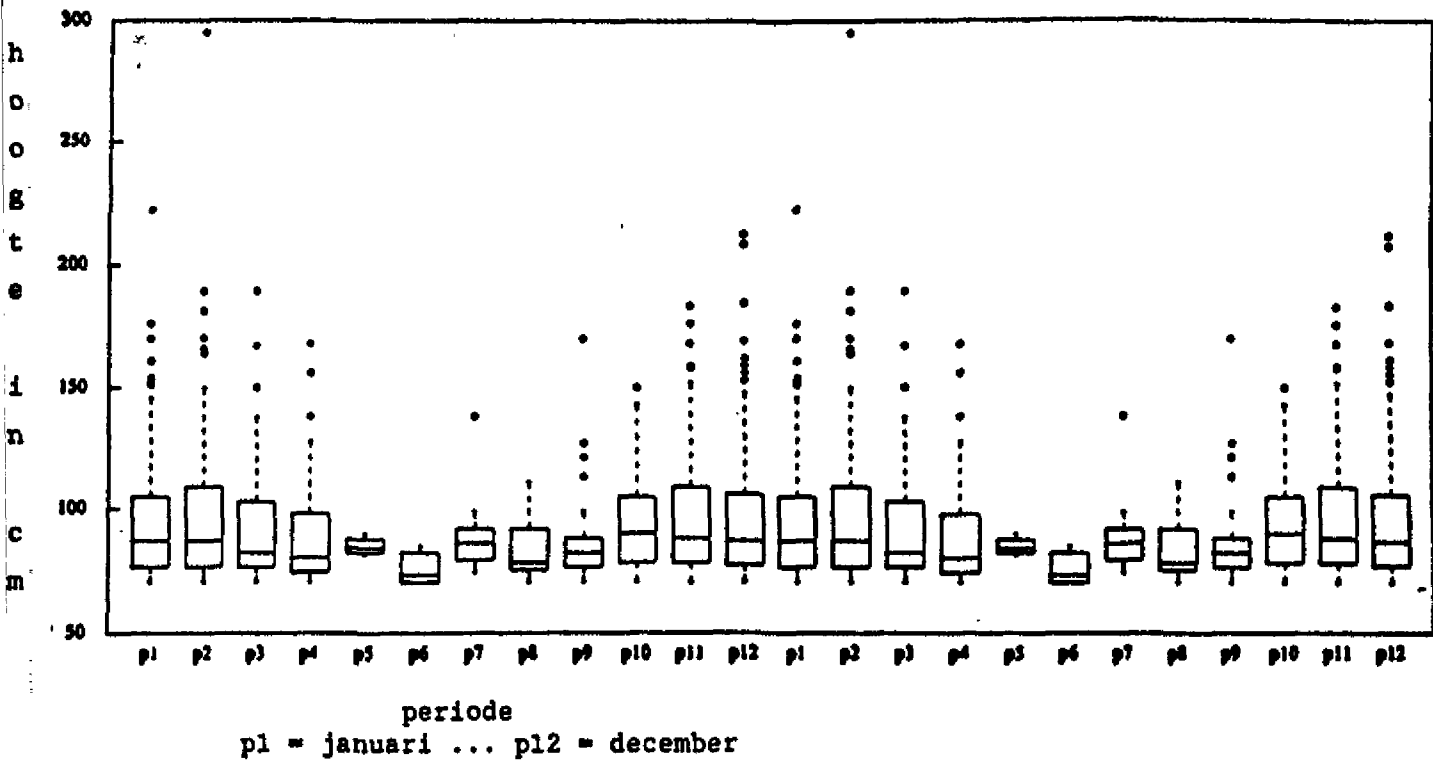
Beschouwd zijn de stations Hoek van Holland (1901...1980) met 940 HW-opzetten ≥ 70 cm (= 2% van alle waarnemingen) en Delfzijl (1902...1985) met 1837 HW-opzetten ≥ 70 cm (= 3% van alle HW-opzetten). De lengte van de tijdreeksen werd bepaald door wat op dat moment van dit deelonderzoek beschikbaar was.

Onderscheiden zijn de HW-opzetten in het interval [70,99] cm en de HW-opzetten ≥ 100 cm.

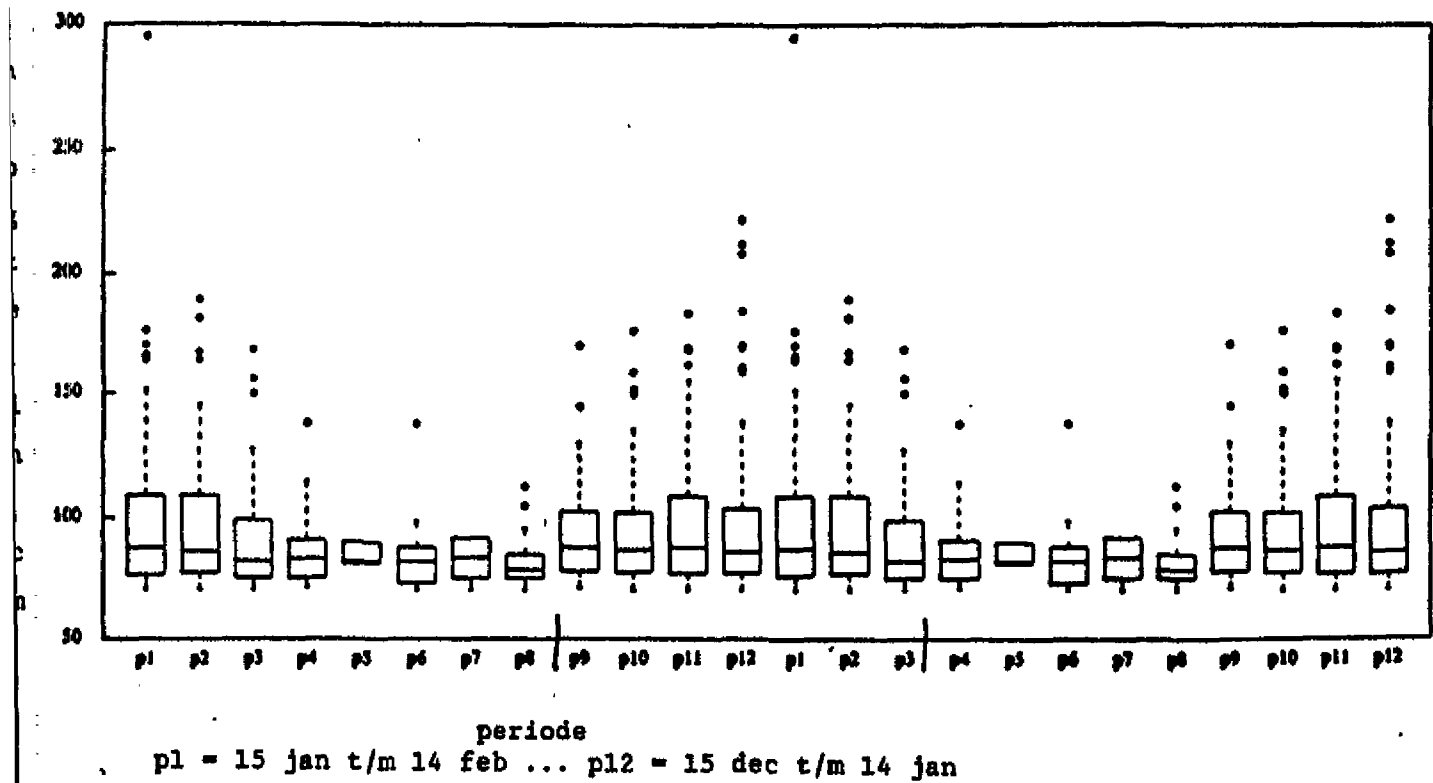
Er zijn twee soorten kalenderperioden beschouwd: de hele maanden (jan., febr., etc.) en de perioden van de 15^e van elke maand t/m de 14^e van de volgende maand. De figuren 4.1 t/m 4.4 geven een goede indruk van het verloop in het jaar van de verdelingen, bepaald per periode van een maand, van alle hoogwateropzetten ≥ 70 cm.

Om goed zicht te krijgen op het verloop van de zomer/winter-cyclus zijn in de figuren telkens twee (identieke) jaarcycli naast elkaar getekend.

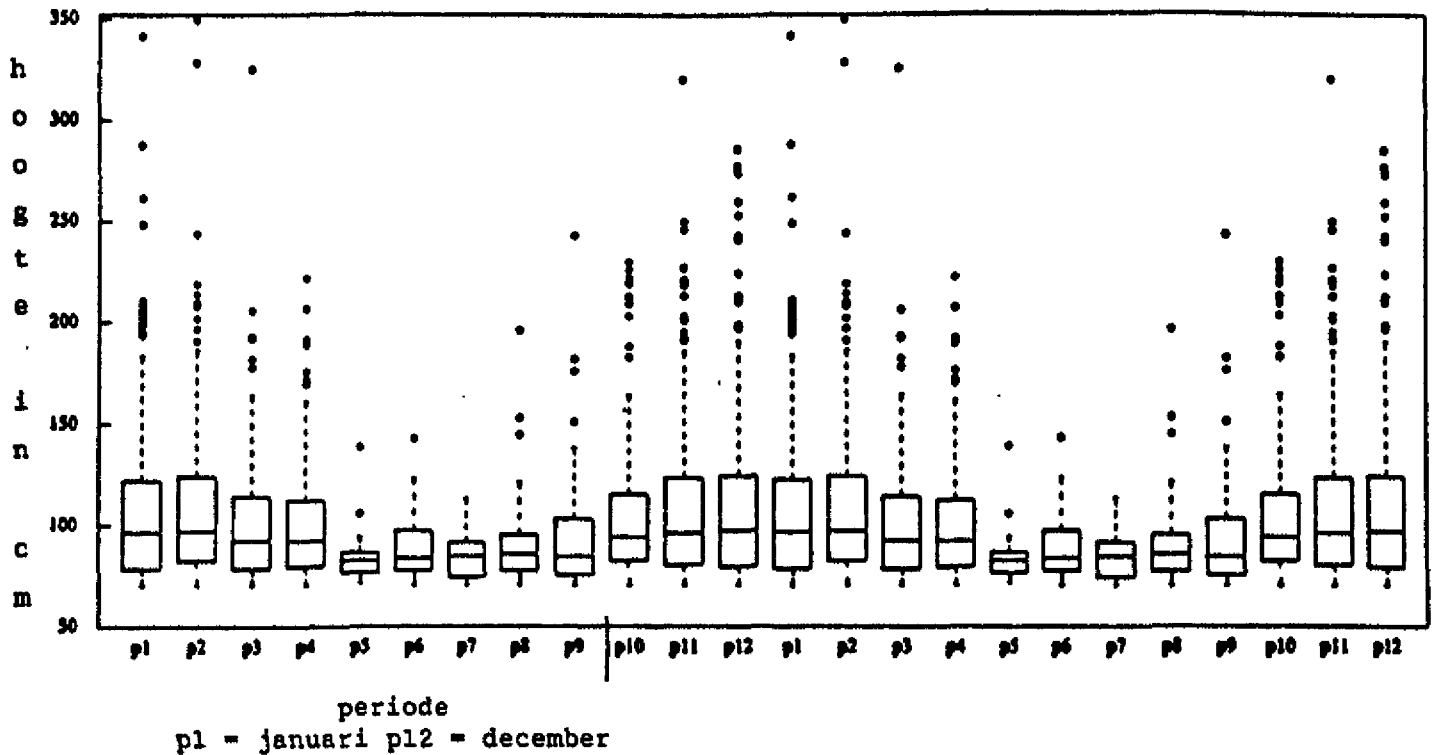
Figuur 4.1: 9 40 HW-opzetten ≥ 70 cm te Hoek van Holland 1901 ... 1980



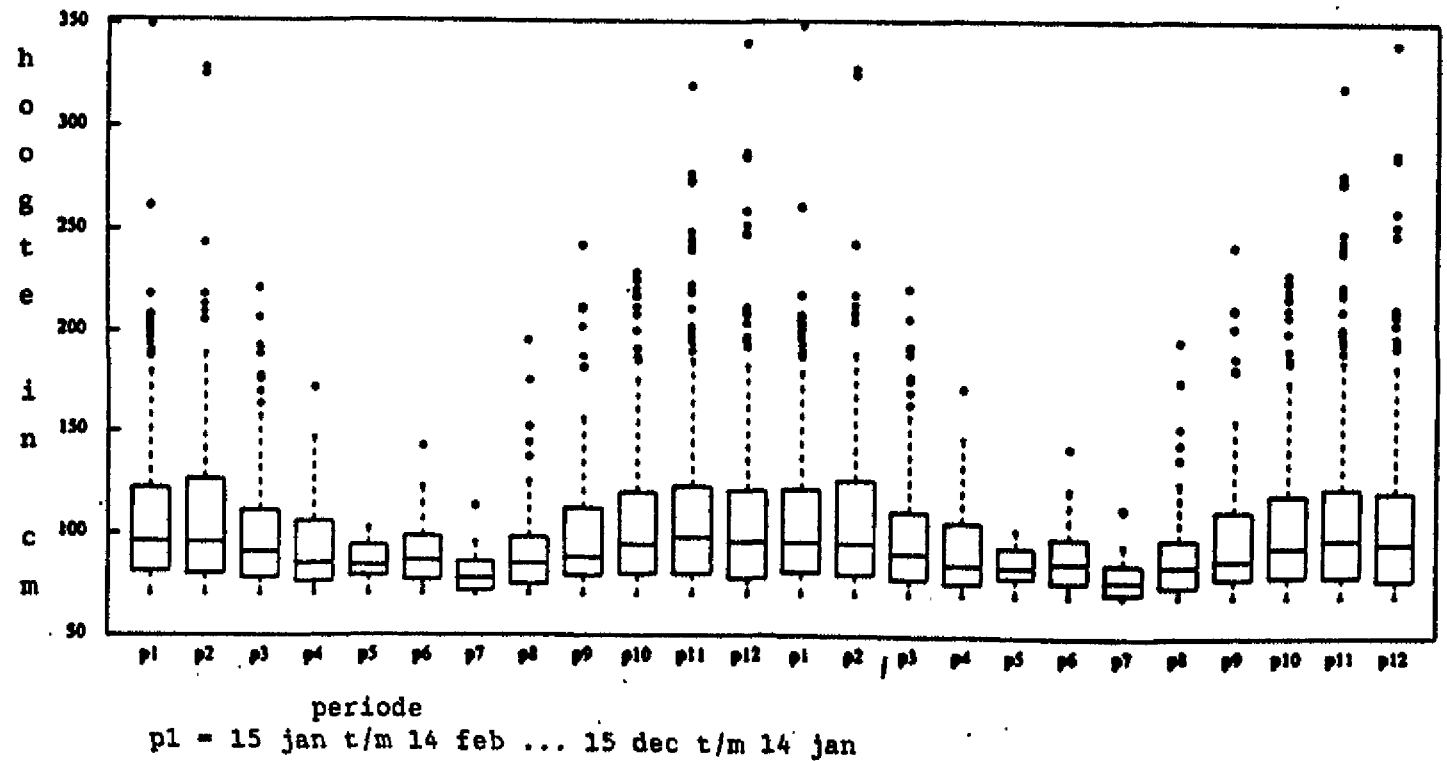
Figuur 4.2: 940 HW-opzetten ≥ 70 cm te Hoek van Holland 1901 ... 1980



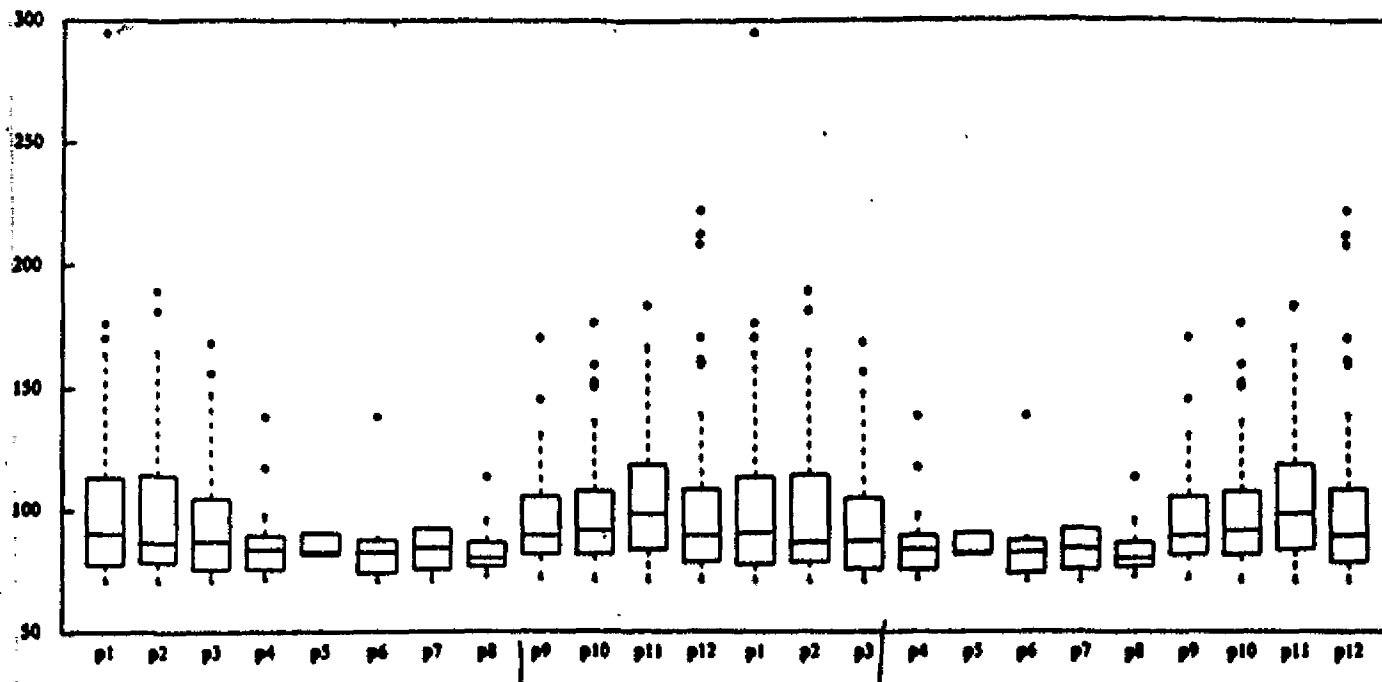
Figuur 4.3: 1837 HW-opzetten ≥ 70 cm. te DELFZIJL 1902 ... 1965



Figuur 4.4: HW-opzetten ≥ 70 cm. te DELFZIJL 1902 ... 1965



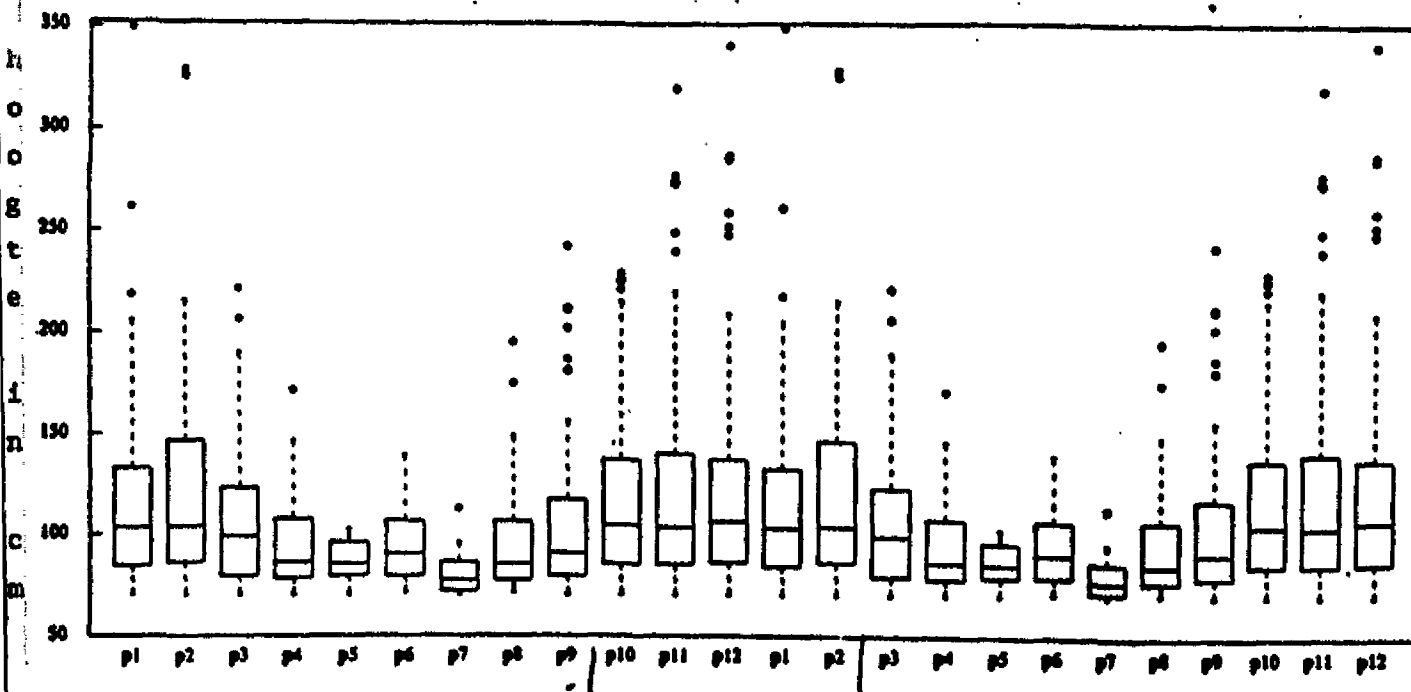
Figuur 4.5: 591 HW-opzetten ≥ 70 cm. te Hoek van Holland D-S-3 1901 ... 1980



periode

p1 = 15 jan t/m 14 feb ... p12 = 15 dec t/m 14 jan

Figuur 4.6: 983 HW-opzetten ≥ 70 cm te DELFZIJL D-S-3 1902 ... 1985



periode

p1 = 15 jan t/m 14 feb ... p12 = 15 dec t/m 14 jan

Omdat uitgegaan moet worden van onafhankelijke waarnemingen dienen de HW-opzetreksen te worden onderworpen aan een selectieprocedure. Omdat op het moment van dit deelonderzoek nog geen definitieve keuze was gedaan voor de selectiemethode, met name de i-waarde van de Daan-S-1 selectie, is gekeken of het gebruik van geselecteerd materiaal tot wezenlijk andere conclusies zou leiden. De figuren 4.5 en 4.6 hebben betrekking op volgens Daan-S-3 geselecteerd materiaal. Vergelijking van de figuren 4.2 en 4.5 respectievelijk 4.4 en 4.6 geeft aan dat, hoewel de boxplotten wel worden beïnvloed, een seizoenkeuze gebaseerd op niet-geselecteerd materiaal niet wezenlijk zal afwijken van de keuze gebaseerd op wel-geselecteerd materiaal. Deze conclusie wordt nog eens bevestigd door de figuren 4.7 en 4.8, waarop de gemiddelden van de HW-opzetten ≥ 70 cm per periode van een maand zijn weergegeven voor de gevallen met en zonder Daan-S-3-selectie. Tabel 4.1 geeft de numerieke waarden.

In de figuren 4.9 en 4.10 zijn de gemiddelden van de HW-opzetten ≥ 70 cm van het niet-geselecteerde materiaal weergegeven.

Figuur 4.11 geeft het verloop van de standaardafwijking van de HW-opzetten ≥ 70 cm (zie ook tabel 4.2). Deze zijn alleen berekend voor de geselecteerde HW-opzetten.

Figuur 4.12 geeft nog enige informatie over de vorm van de verdelingen.

Figuur 4.13 geeft tenslotte het verloop van het gemiddelde aantal HW-opzetten per periode van een maand, ten behoeve van de vaststelling van het "korte" stormseizoen (zie ook tabel 4.3).

4.2.5 De bepaling van de lengte van het stormseizoen.

Met behulp van het in de vorige paragraaf beschreven materiaal wordt nu getracht een periode te vinden, waarin de verdeling van de HW-opzetten bij benadering gelijk mag worden verondersteld. Er wordt daarbij uitgegaan van de visuele indruk die de genoemde figuren geven. Aan enige subjectiviteit valt daarbij niet te ontkomen.

Achtereenvolgens worden voor Hoek van Holland en Delfzijl, voor de verschillende figuren die daarvoor een aanwijzing geven, de gesuggereerde lengtes van het stormseizoen genoteerd. Met behulp van deze informatie wordt een keuze gemaakt voor het stormseizoen.

Hoek van Holland

| | |
|----------------------------------|------------------------|
| Fig. 4.1 en 4.2 (boxplotten) | : 15 sept. t/m 14 mrt. |
| Fig. 4.5 (boxplot, D-S-3) | : 15 sept. t/m 14 apr. |
| Fig. 4.9 (gemiddelde) | : 15 sept. t/m mrt. |
| Fig. 4.11 (standaardafw.) | : 15 sept. t/m 14 apr. |
| Fig. 4.12 ($\geq 100/[70,99]$) | : 15 sept. t/m mrt. |

Delfzijl

| | |
|----------------------------------|-----------------------|
| Fig. 4.3 en 4.4 (boxplotten) | : okt. t/m 14 mrt. |
| Fig. 4.6 (boxplot, D-S-3) | : 15 okt. t/m 14 mrt. |
| Fig. 4.10 (gemiddelde) | : okt. t/m 14 apr. |
| Fig. 4.11 (standaardafw.) | : okt. t/m 14 apr. |
| Fig. 4.12 ($\geq 100/[70,99]$) | : okt. t/m 14 mrt. |

De keuze voor het "korte" stormseizoen kan worden gemaakt aan de hand van figuur 4.13, waarin het jaarlijkse verloop van het gemiddelde aantal HW-opzetten per maand is geschetst.

Hoewel niet strikt noodzakelijk, is er de voorkeur aan gegeven voor alle basisstations hetzelfde stormseizoen aan te houden. Een tot op de dag nauwkeurig bepaalde lengte is daarbij niet mogelijk.

Voor de extrapolatieprocedure ter bepaling van de overschrijdingskansen van extreme hoogwaterstanden is gekozen voor een stormseizoen van 1 oktober t/m 15 maart.

Voor het voor de selectie gebruikte "korte" stormseizoen zijn de maanden november, december en januari gekozen.

tabel 4.1 Gemiddelde waarden van alle HW-opzetten ≥ 70 cm per periode van een maand [cm]

| Hoek van Holland (1901-1980) | | Delfzijl (1902-1985) | |
|------------------------------|-------------------------------|----------------------|-------------------------------|
| jan., febr., etc. | 15 jan. t/m 14 febr., etc. | jan., febr., etc. | 15 jan. t/m 14 febr., etc. |
| 94.0 | 95.8 | 107.6 | 108.4 |
| 97.4 | 96.5 | 111.9 | 109.9 |
| 92.3 | 90.4 | 101.7 | 120.2 |
| 89.9 | 86.2 | 102.1 | 94.8 |
| 84.8 | 84.3 | 85.2 | 86.2 |
| 76.0 | 86.0 | 89.1 | 89.6 |
| 90.4 | 82.8 | 84.0 | 80.0 |
| 83.8 | 81.8 | 90.8 | 81.4 |
| 86.8 | 92.9 | 92.7 | 99.5 |
| 93.5 | 92.4 | 104.2 | 105.6 |
| 95.2 | 96.2 | 106.8 | 109.4 |
| 95.4 | 95.0 | 108.9 | 107.9 |

tabel 4.2 Gemiddelde en standaardafwijking van de geselecteerde (Daan-S-3) HW-opzetten ≥ 70 cm

| Hoek van Holland (1901-1980) | | Delfzijl (1902-1985) | |
|------------------------------|--------|----------------------|--------|
| m [cm] | s [cm] | m [cm] | s [cm] |
| 99.4 | 31.8 | 116.5 | 43.5 |
| 99.3 | 30.2 | 120.8 | 50.6 |
| 94.4 | 24.4 | 106.8 | 34.0 |
| 87.8 | 18.8 | 97.1 | 25.3 |
| 84.3 | 4.9 | 87.3 | 10.1 |
| 96.4 | 21.1 | 94.5 | 18.9 |
| 82.7 | 8.9 | 88.4 | 10.3 |
| 82.6 | 9.9 | 96.8 | 27.9 |
| 96.2 | 21.4 | 103.9 | 35.6 |
| 96.9 | 21.7 | 116.6 | 41.5 |
| 103.5 | 25.7 | 119.6 | 47.2 |
| 97.8 | 28.5 | 117.2 | 45.6 |

tabel 4.3 De aantallen HW-opzetten ≥ 70 cm met de lengte van het bijbehorende tweezijdige (symmetrische) 95%-betrouwbaarheidsinterval

| Hoek van Holland (1901-1980) | | | | Delfzijl (1902-1985) | | | |
|------------------------------|--------|-------------------------------|--------|----------------------|--------|-------------------------------|--------|
| jan., febr., etc. | | 15 jan. t/m 14 febr., etc. | | jan., febr., etc. | | 15 jan. t/m 14 febr., etc. | |
| aantal | lengte | aantal | lengte | aantal | lengte | aantal | lengte |
| 176 | 78 | 160 | 80 | 338 | 142 | 293 | 115 |
| 119 | 66 | 91 | 58 | 207 | 98 | 162 | 94 |
| 73 | 50 | 61 | 46 | 159 | 96 | 142 | 87 |
| 40 | 36 | 16 | 20 | 105 | 68 | 47 | 34 |
| 4 | 8 | 3 | 6 | 17 | 20 | 13 | 16 |
| 5 | 10 | 10 | 14 | 29 | 30 | 36 | 34 |
| 9 | 14 | 6 | 12 | 25 | 28 | 29 | 28 |
| 13 | 20 | 24 | 24 | 45 | 38 | 79 | 54 |
| 33 | 16 | 67 | 46 | 107 | 56 | 139 | 64 |
| 113 | 56 | 144 | 72 | 185 | 84 | 234 | 96 |
| 175 | 84 | 200 | 88 | 317 | 118 | 337 | 136 |
| 180 | 86 | 158 | 70 | 303 | 130 | 326 | 119 |

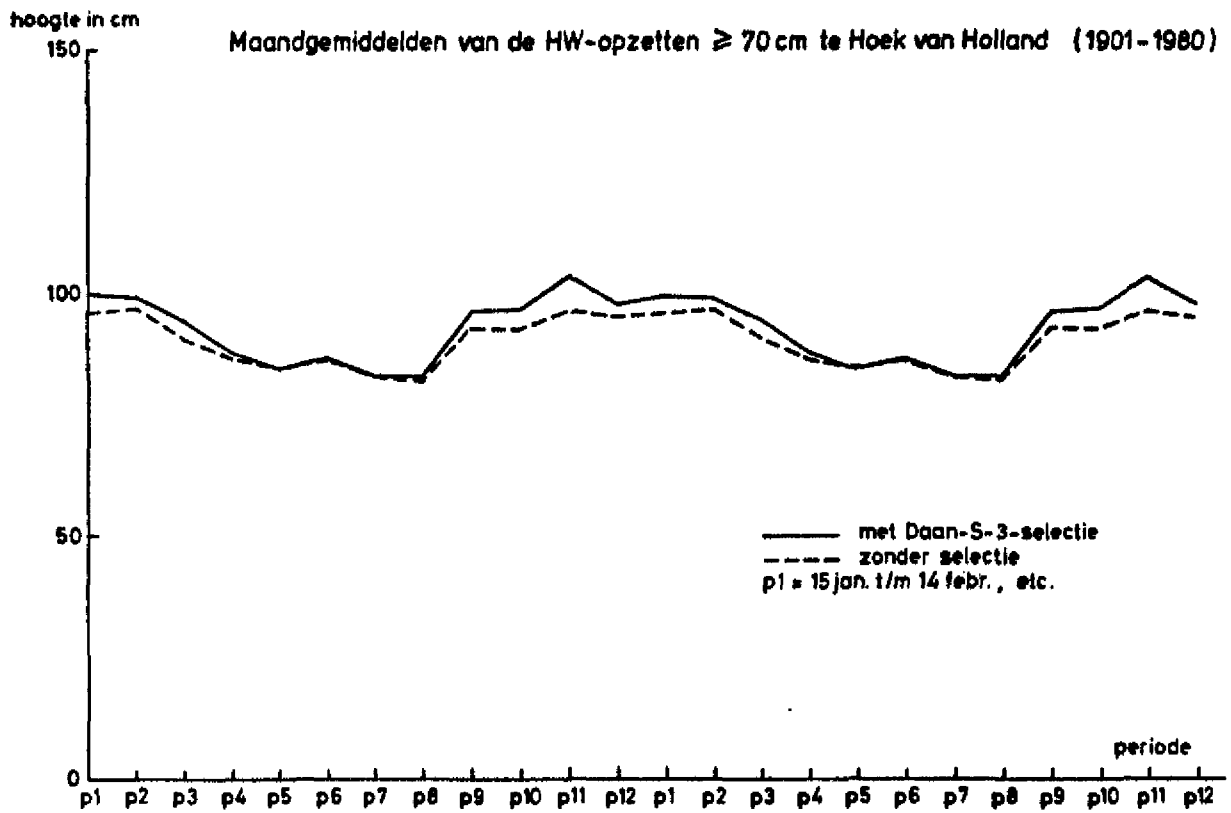


FIG 4.7

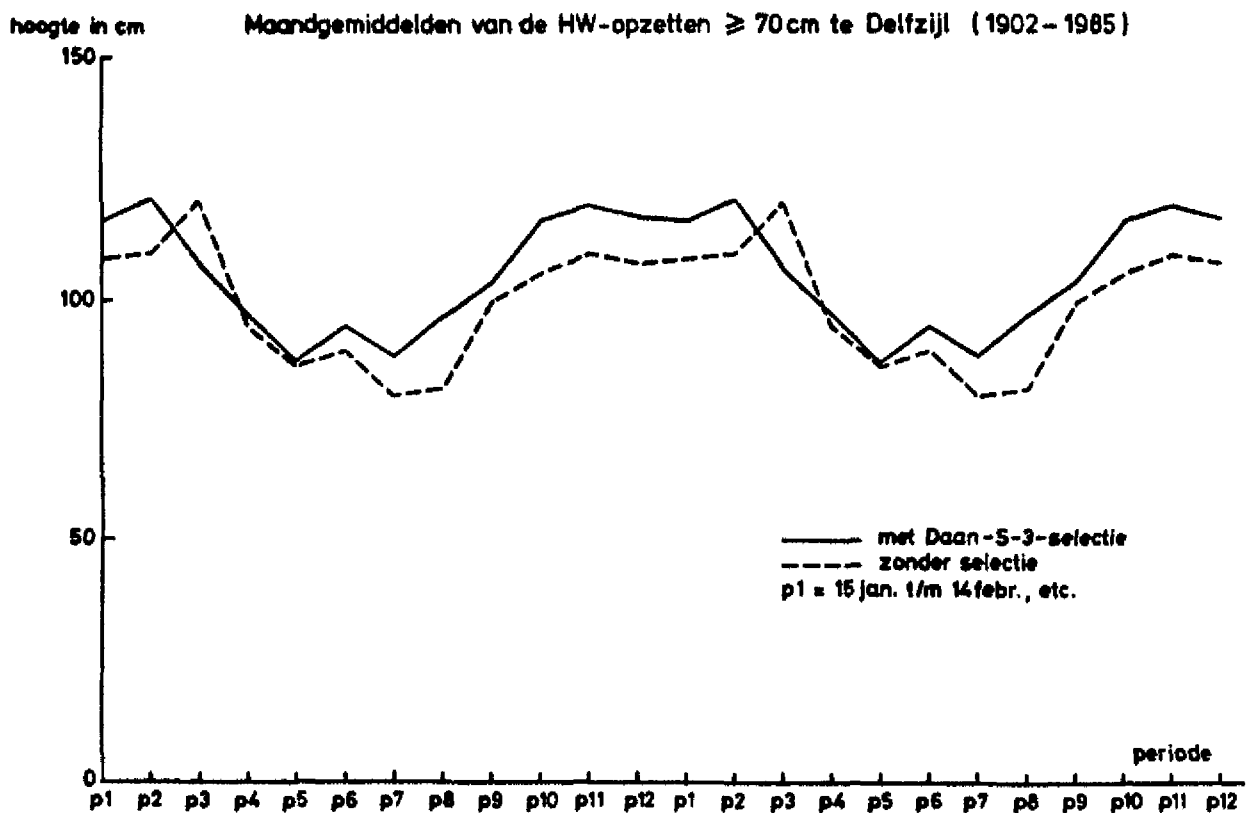


FIG 4.8

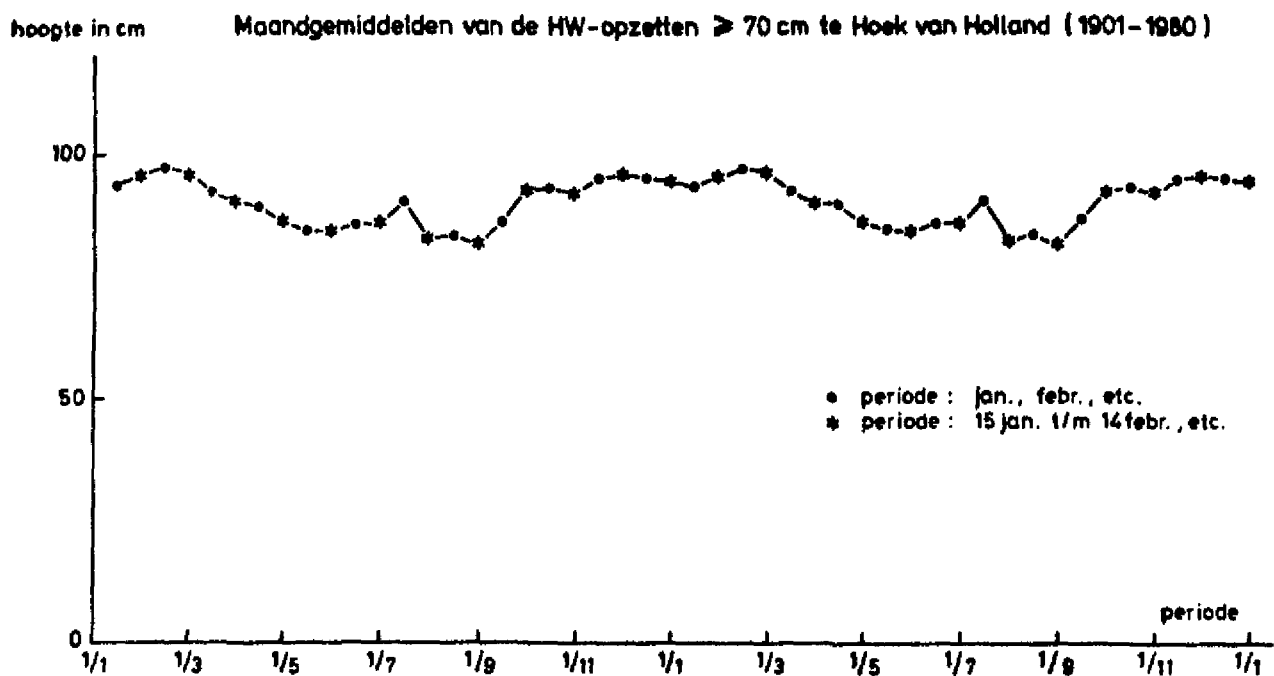


FIG 4.9

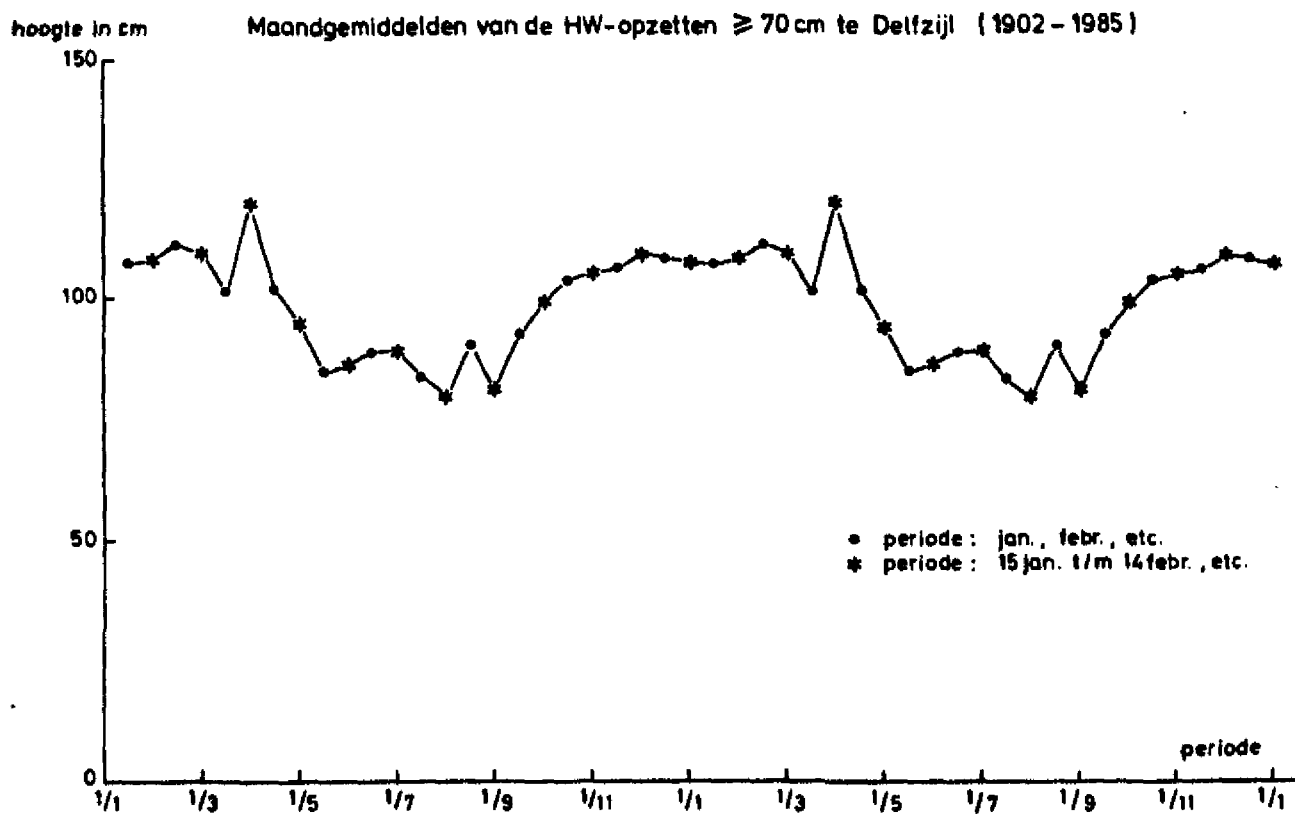


FIG 4.10

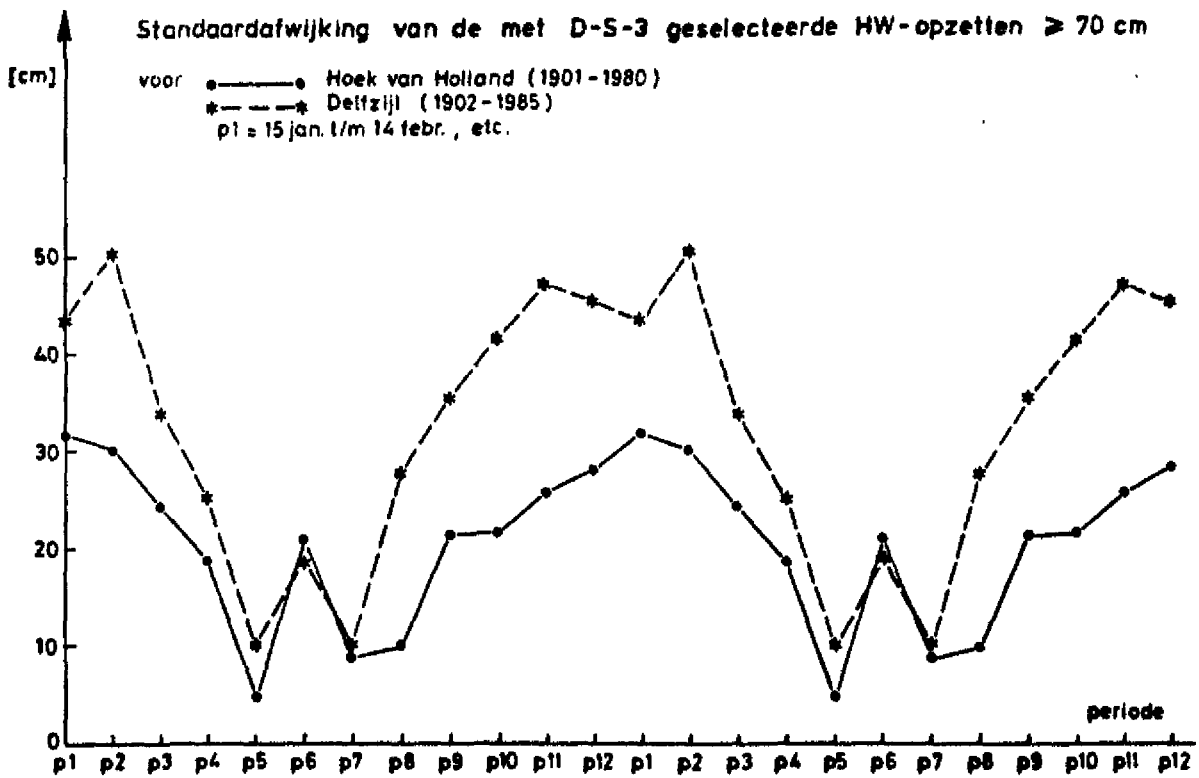


FIG 4.11

De verhouding van de aantallen HW-opzetten > 99 cm en $[70, 99]$ cm per maand.

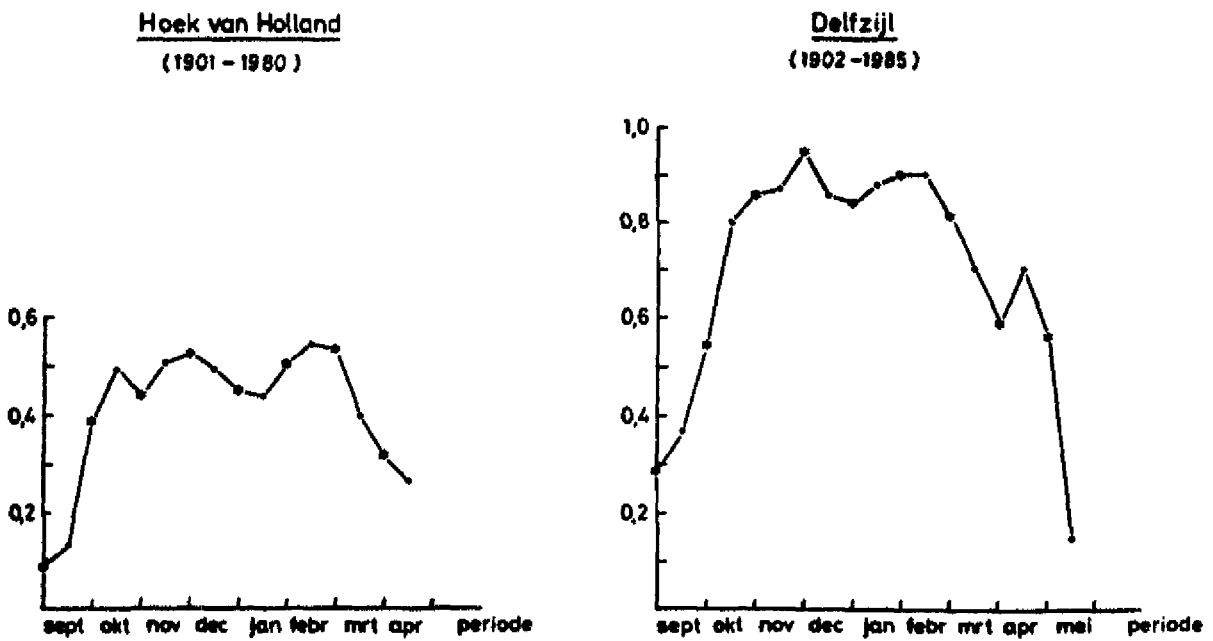


FIG 4.12

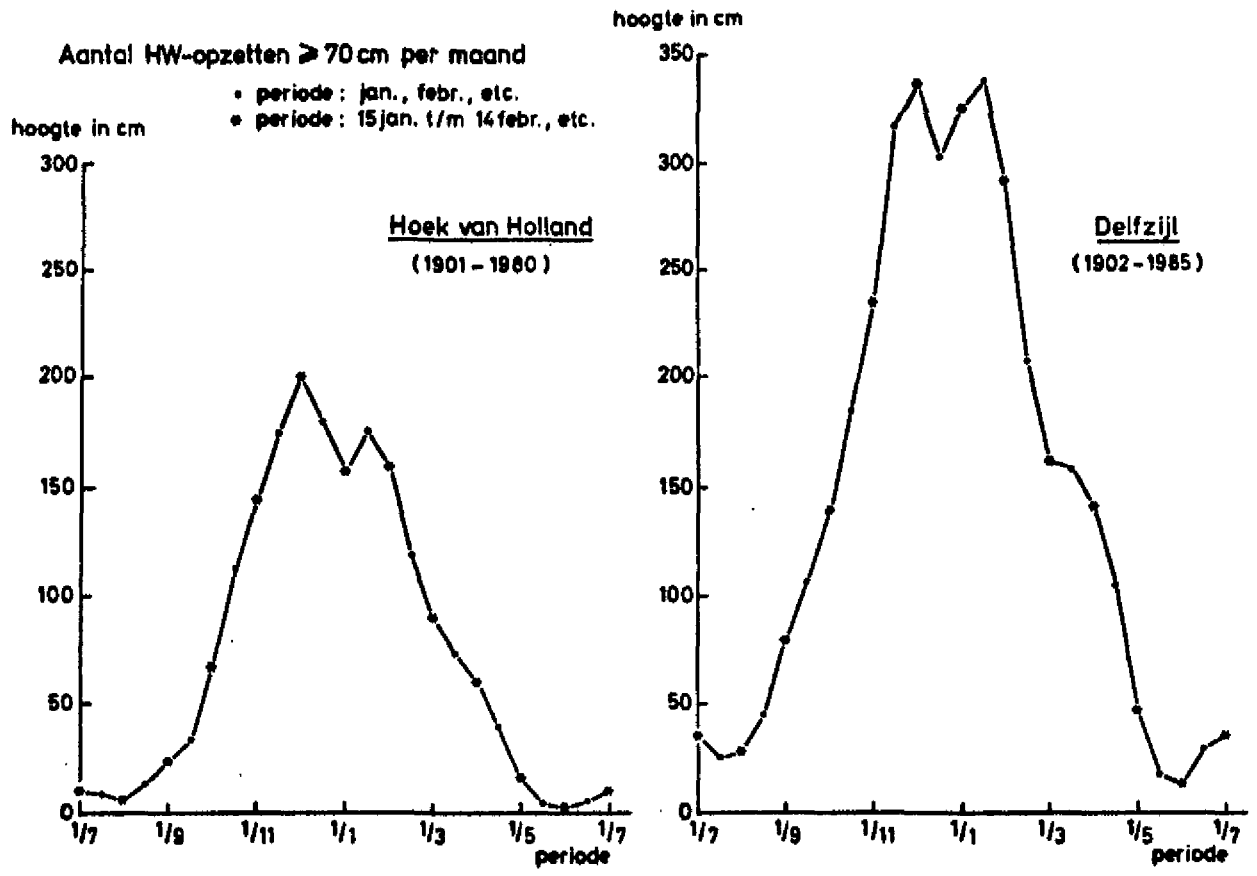


Fig. 4.13

5. SELECTIEMETHODE

5.1 Inleiding

Voor de samenstelling van een frequentiekromme van hoogwateropzetten is het gewenst dat de gegevensreeks uit elementen bestaat die

- onderling onafhankelijk zijn, en
- tot eenzelfde verdeling behoren.

Onderlinge afhankelijkheid tussen opeenvolgende gegevens leidt in het algemeen tot een andere benadering voor de verdeling van jaarmaximum (b.v. maximale HW-opzet in een gegeven stormseizoen), dan in het klassieke geval van onafhankelijkheid. Dit blijkt in te houden dat i.h.b. de extreme kwantielen van het jaarmaximum (b.v. 10^{-4} of 10^{-3} kwantiel) kleiner zullen uitvallen dan het geval is bij toepassing van de extreme waarden theorie voor maxima van onafhankelijke gegevens. De afhankelijkheid kan onderdrukt worden door toepassing van een geschikte selectiemethode in de tijd. Een aantal elementen uit de reeks blijft daarbij buiten beschouwing, de overblijvende worden geacht onderling onafhankelijk, of autocorrelatievrij te zijn. In par. 2.5 is reeds aangegeven welke methode daarvoor in aanmerking komt; slechts de bepaling van het getal i vereist nog de nodige aandacht.

In de volgende paragrafen wordt nader op deze zaak ingegaan. Par. 5.2 geeft een algemene beschouwing over autocorrelatie bij hoogwateropzetten, terwijl par. 5.3 het probleem volgens een bepaald statistisch criterium beziet. Uit één en ander zou een waarde voor i 4 volgen.

Het probleem van de autocorrelatie en de invloed hiervan op de resulterende verdeling van het maximum vormt het onderwerp van par. 5.4.

5.2 Enkele opmerkingen over autocorrelatie bij HW-opzetten

5.2.1 Probleemstelling

Ten behoeve van het Frequentielijnen Kust vindt statistische verwerking der HW-opzetten plaats. Het hiervoor beschikbare gegevensmateriaal beslaat voor de in eerste aanleg nader beschouwde stations (Vlissingen, Hoek van Holland en Delfzijl) ongeveer de periode 1885-1985. Ten einde homogeen materiaal te verkrijgen wordt hiervan per station een tijdreeks gevormd bestaande uit de desbetreffende gegevens van de opeenvolgende stormseizoenen. Daarmee worden echter nog geen stochastische reeksen verkregen; om deze moeilijkheid te ondervangen zijn verschillende selectiemethoden voorgesteld en toegepast ("Daan-1" en "Daan-S-3"); er is echter reden aan te nemen dat althans theoretisch bij elk van deze methoden materiaal ten onrechte wordt verwijderd en ander materiaal ten onrechte wordt gehandhaafd. Op grond hiervan leek de hierna volgende nadere beschouwing van de aard der verstoringen binnen een nog onaangepaste ("onbewerkte") tijdreeks van HW-opzetten en van hun draagwijdte van enig nut.

5.2.2 Afhankelijkheden binnen een tijdreeks

De opeenvolgende stormseizoenen zijn door tenminste 7 maanden (half maart tot half oktober) van elkaar gescheiden; men kan derhalve veilig stellen dat er van beïnvloeding van HW-opzetten in zo'n onbewerkte reeks alleen sprake kan zijn binnen de deelreeksen der stormseizoenen. Binnen zo'n deelreeks kan een

dergelijke beïnvloeding op uiteenlopende wijze tot stand zijn gekomen. Onder beperking tot aanmerkelijke HW-opzetten verdienen in dit verband vermelding;

- a de duur van het opwaaiingsverschijnsel;
- b het terugslingeringsverschijnsel op de Noordzee;
- c het tijdsbeslag van de invloed van een ongunstig meteorologisch circulatiepatroon.

ad a: Duur opwaaiingsverschijnsel

Belangrijke HW-opzetten, gevolg van stormen uit windstreken tussen West en Noord, vallen uiteraard binnen opwaaiingsperioden. Zo'n periode kan zo lang zijn (al is dit niet noodzakelijk!) dat daarin ettelijke hoogwaters vallen. Bij de stormvloed van 31 januari/1 februari 1953 b.v. besloeg de opwaaiing 50 uur en beïnvloedde 4 hoogwaters (lit. 12, fig. 2.12...2.21, blz. 111...120); het (verevende!) opwaaiingsverloop was toen bij benadering driehoekig met het maximum na 20 uur.

ad b: Terugslingeringsverschijnsel

Na het wegvallen van de wind "loopt het water weg" om na een zeker tijdsverloop terug te keren, waarmee het een dan optredend hoogwater kan beïnvloeden (lit. 13, par. 3.3, blz. 16). Dit tijdsverloop is max. 36 uur (d.w.z. zo'n 3 getijperioden) lang. Een complicerende factor is dat de snelheid waarmee de wind "wegvalt", van geval tot geval verschilt.

ad c: Invloedsduur ongunstig circulatiepatroon

In de maanden november...januari der periode 1949...1979 werden bij Hoek van Holland 53% der HW-opzetten 50 cm en 80% der HW-opzetten > 100 cm teweeggebracht bij 3 der 27 bij het KNMI onderscheiden circulatiepatronen; in 17,5% van het totale aantal beschouwde dagen trad een van deze 3 patronen of een overgang daartussen op. De gemiddelde duur van zo'n "optreden" beliep 2,4 dag, 2 x zo lang als bij aselectie zou behoren (par. 2.2/2.3). Verder bestaat tussen opeenvolgende circulatiepatronen een vrij sterke autocorrelatie, die pas na ongeveer 10 dagen nul wordt (lit. 13, par. 3.3, blz. 16).

5.2.3 Beperking der afhankelijkheid

Het is duidelijk dat er een aantal bronnen voor autocorrelatie tussen de HW-opzetten t.p.v. een willekeurig station bestaat. In de verschillende stormseizoenen zullen deze bronnen in wisselende verhouding "optreden" terwijl bovendien ook hun uitwerking nog zal "schommelen". Over een gehele tijdreeks gezien is de autocorrelatie dus een uitermate samengesteld verschijnsel, waarmee over haar grootte en de wisseling daarvan overigens nog niets is gezegd.

Men zou nu kunnen overwegen elk der 3 factoren genoemd onder par. 5.2.2, afzonderlijk op hun invloed te onderzoeken; daarbij passen de volgende kanttekeningen:

ad a: Voor ieder geval van aanmerkelijke opwaaiing zou het verloop ervan getekend moeten worden. Inderdaad is er hiervan voor wat betreft de periode vanaf november 1954 materiaal aanwezig (afkomstig van de zittingen van de Stormvloedwaarschuwingsdienst), maar dat is onvolledig (zowel naar het aantal gevallen als naar de opwaaiingsduur), zou correctie behoeven in

verband met mogelijke vervroeging of verlating van het getij en de interpretatie zou subjectief zijn. Dit lijkt al met al geen begaanbare weg; wel is onderzoek gedaan naar de opwaaiingsduur te Hoek van Holland aan de hand van de gegevens van een 50 stormvloed. Deze duur bleek gemiddeld 30 uur te zijn met een standaardafwijking van 10 uur; daarbij werd geen verband tussen opwaaiingsduur en grootte van de HW-opzet gevonden. Volledigheidshalve: de opwaaiingsduur kan mede beïnvloed zijn door de factor c .

ad b: Het terugslingeringsverschijnsel leent zich tot statistisch onderzoek en dit is althans in voorlopige vorm voor Vlissingen (in febr. '87) ook verricht; daarbij bleek de autocorrelatie tussen de HW-opzet en die van 3 hoogwaters later zeer klein (autocorrelatie-coëfficiënt ongeveer 0,1).

ad c: Uit de gegevens van par. 2.2 blijkt dat de bedoelde circulatiepatronen zich in de beschouwde periode gedurende 499 dagen voordoen. In 52% van die gevallen was de opzet ≥ 50 cm, in slechts 12% was HW-opzet 100 cm. De kans op relatief grote opzetten is bij deze weersconstellaties dus weliswaar veel groter dan bij andere, maar dit brengt niet noodzakelijkerwijze een aanmerkelijke auto-correlatie mee. In dit verband lijken drie omstandigheden van belang; allereerst de gevoeligheid van de HW-opzet voor de windrichting, welke windrichting tijdens storm op haar beurt sterk afhankelijk is van de ligging van de baan van de (een) lagedruk-kern: een relatief kleine zuidelijke verschuiving van die baan van zeg 500 km kan i.p.v. een zeer hoge een kleine opwaaiing tot gevolg hebben enz. De tweede omstandigheid (samenhangend met de eerste) ligt in de ruimtelijke verdeling van de opwaaiing, die niet zelden een beperkt-regionale concentratie vertoont. De derde omstandigheid ligt in duur, aard en ligging (in de tijd) van het opwaaiingsverloop, dat heel wel een groot maximum kan hebben zonder de hoogwaterstanden sterk te verhogen.

5.2.4 Invloed selectiemethode

Om een indruk te krijgen van de invloed van de toegepaste selectiemethode op de grootte en samenstelling van de reeks is het volgende gedaan *).

Uitgangspunt is dat tussen twee opzetten van tenminste 50 cm die in de tijd gescheiden zijn door drie aansluitende HW's met een opzet < 50 cm geen correlatie bestaat. Dan kan men de tijdreeks van opzetten ≥ 50 cm verdelen in groepen die telkens door zo'n interval zijn gescheiden. Uit elke groep zal tenminste één geselecteerde opzet voortkomen.

*) Deze vergelijking is opgesteld alvorens definitief tot een keuze voor een selectiesysteem werd gekomen. De in de vergelijking betrokken selecties Daan-3 en Daan-5 wijken af van de uiteindelijke keuze, maar kunnen voor een vergelijking met de in het Deltarapport toegepaste selectie-methode wel als exemplarisch voor het nu voorgestelde selectiesysteem beschouwd worden.

De samenstelling van deze groepen is nader bestudeerd. In tabel 5.1 is een verdeling gegeven van de groepen naar de lengte. Onder lengte is verstaan het aantal (opeenvolgende) gevallen dat tot de groep behoort; dit kunnen zowel gevallen met een opzet ≥ 50 cm zijn als beneden 50 cm.

Tabel 5.1: Verdeling van opzetten over groepen.

| Groep lengte | Aantal opzetten ≥ 50 cm | Aantal groepen | DAAN-3 selectie | DAAN-5 selectie |
|--------------|------------------------------|----------------|-----------------|-----------------|
| 1 | 240 | 240 | 240 | 240 |
| 2 | 186 | 93 | 93 | 93 |
| 3 | 145 | 52 | 52 | 52 |
| 4 | 134 | 45 | 45 | 45 |
| 5 | 111 | 30 | 40 | 34 |
| 6 | 58 | 14 | 23 | 19 |
| 7 | 71 | 14 | 24 | 21 |
| 8 | 40 | 7 | 14 | 10 |
| 9 | 37 | 6 | 11 | 10 |
| 10 | 32 | 4 | 8 | 6 |
| 11 | 11 | 2 | 5 | 4 |
| 12 | 8 | 1 | 2 | 1 |
| 13 | 24 | 3 | 9 | 8 |
| 14 | 9 | 1 | 3 | 2 |
| 15 | 9 | 1 | 2 | 2 |
| | 1115 | 513 | 571 | 547 |

Voor elke groeplengte zijn achtereenvolgens vermeld:

- het aantal opzetten ≥ 50 cm dat in de groepen voorkomt.
- het aantal groepen.
- het aantal gevallen in de DAAN-3 selectie (zie par. 2.4).
- het aantal gevallen in de DAAN-5 selectie (zie par. 2.4).

Het verschil tussen deze twee selectiemethoden blijkt in de praktijk niet groter dan 5% : DAAN-3 omvat 571 opzetten, slechts 24 meer dan de DAAN-5 selectie.

Tenslotte is nog een overzicht gegeven van de verschillen tussen de van der Ham - selectie en de selecties DAAN-5 en DAAN-S-3 (zie tabel 5.2). Er zijn 3 vergelijkingen gemaakt:

- alle opzetten
- alle opzetten bij een absolute stand ≥ 170 cm
- alle opzetten ≥ 65 cm.

Tabel 5.2: Vergelijking van selecties over de periode 1901 - 1939 en 1945 - 1957.

Vergelijking v.d. Ham (VDH) *) en DAAN-5 (D-5)

| | alle opzetten | | HW \geq 170 cm | | HW-opzet \geq 65 cm | |
|-------------------------|---------------|----------|------------------|---------|-----------------------|---------|
| | VDH | D-5 | VDH | D-5 | VDH | D-5 |
| Identiek | 174 | 174 | 99 | 99 | 130 | 130 |
| Verschoven | 28 | 28 | 9 | 9 | 13 | 13 |
| Radius | 51 | | 29 | | 37 | |
| Herberekening | 25 | | | | | |
| Baanselectie feb/okt | | 106 4 | | 28 2 | | 54 2 |
| Totaal | 278 | 312 | 137 | 138 | 180 | 199 |

Vergelijking v.d. Ham (VDH) en DAAN-S-3 (D-S-3)

| | alle opzetten | | HW \geq 170 cm | | HW-opzet \geq 65 cm | |
|-------------------------|---------------|----------|------------------|---------|-----------------------|---------|
| | VDH | D-S-3 | VDH | D-S-3 | VDH | D-S-3 |
| Identiek | 201 | 201 | 115 | 115 | 152 | 152 |
| Verschoven | 24 | 24 | 10 | 10 | 15 | 15 |
| Radius | 28 | | 12 | | 13 | |
| Herberekening | 25 | | | | | |
| Baanselectie feb/okt | | 131 5 | | 33 2 | | 60 2 |
| Totaal | 278 | 361 | 137 | 160 | 180 | 229 |

Vergelijking DAAN-5 en DAAN-S-3

| | alle opzetten | | HW \geq 170 cm | | HW-opzet \geq 65 cm | |
|-------------|---------------|-------|------------------|-------|-----------------------|-------|
| | D-5 | D-S-3 | D-5 | D-S-3 | D-5 | D-S-3 |
| Identiek | 298 | 298 | 135 | 135 | 191 | 191 |
| Verschoven | 9 | 9 | 1 | 1 | 3 | 3 |
| Verschillen | 5 | 54 | 2 | 24 | 5 | 35 |
| Totaal | 312 | 361 | 138 | 160 | 199 | 229 |

) Zie het 1 Tussentijds Verslag (lit 13, par 3.4.).

Onderscheid is gemaakt naar de aard van de verschillen:

| | | |
|---------------|---|---|
| Verschoven | : | De geselecteerde gevallen zijn niet op hetzelfde tijdstip gekozen. De meest voorkomende oorzaak is dat van der Ham niet de hoogste opzet selecteerde, maar de hoogste stand bij een opzet van minstens 50 cm. |
| Radius | : | Van der Ham selecteert een geval dat zeer dicht bij een andere geselecteerde opzet ligt. |
| Herberekening | : | Van der Ham selecteert een opzet die na herberekening van de astronomische standen beneden 50 cm blijkt te blijven. |
| Baansselectie | : | Van der Ham verwerpt een opzet omdat de depressiebaan niet aan de eisen voldoet. Hieronder kunnen zich ook gevallen bevinden waarin de opzet volgens de oorspronkelijke berekening beneden 50 cm bleef. |
| feb/okt | : | Van der Ham heeft slechts 3 gevallen buiten de maanden november t/m januari geselecteerd: 1 feb. 1908, 31 okt. 1912 en 1 feb. 1953. De hier vermelde gevallen kwamen dus in zijn selectie niet voor. |

5.3 Selectie van HW-opzetten

5.3.1 Afhankelijkheid en clustering van hoge HW-opzetten

Afhankelijkheid van de HW-opzetten in een gegeven stormseizoen (1 november...31 januari is een groot probleem bij de statistische analyse. Door selectie van HW-opzetten willen we proberen de situatie van onafhankelijkheid te benaderen met het idee dat als waarnemingen in de tijd ver uit elkaar liggen ze ongeveer onafhankelijk zijn.

Laat X_1, \dots, X_N de N in de tijd opeenvolgende HW-opzetten in het stormseizoen (in een gegeven jaar) voorstellen. We zijn geïnteresseerd in de kansverdeling van de maxima $M_N = \max X_i$. In het klassieke geval dat de X_i 's onafhankelijk en gelijkverdeeld zijn, kan de kansverdeling van een passend genormeerde M_N voor grote waarden van N , benaderd worden door een extreme-waarden verdeling G_k met vormparameter k . Ook is bekend dat in dit geval het puntproces van de stochastische tijdstippen waarop hoge HW-opzetten optreden - d.w.z. HW-opzetten welke een hoge drempel $u = u_n$ overschrijden - voor grote waarden van N , benaderd kunnen worden door een Poisson proces. In het bijzonder betekent dat dit geen clustering zal optreden van de hoge HW-opzetten.

In het algemene geval dat de X_i 's afhankelijk zijn kan echter wel clustering van hoge HW-opzetten optreden. We bekijken de situatie van een stationaire rij X_1, \dots, X_N van N opeenvolgende HW-opzetten. Indien de HW-opzetten welke in de tijd dicht bij elkaar liggen afhankelijk zijn zal veelal clustering optreden. Ook de kansverdeling van het genormeerde maximum M_N (de normering is dezelfde als in het speciale geval van onafhankelijkheid) ziet er anders uit: voor grote waarden van N , wordt deze verdeling benaderd door G_k^θ , waarbij G_k de extreme-waarden verdeling met vormparameter k is en θ Leadbetter's extremaalindex ($0 \leq \theta \leq 1$) voorstelt. Een nadere beschrijving van deze problematiek is te vinden in par. 5.4.

De parameter θ kan geïnterpreteerd worden als de reciproke van de gemiddelde clusteromvang in het puntproces van hoge HW-opzetten. In het klassieke geval van onafhankelijkheid is $\theta=1$; in het algemeen geldt $0 \leq \theta \leq 1$.

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 (lit. 16)).

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

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 quotient 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

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

en

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

zodat we mogen verwachten dat

$$EV_r(u) = \frac{2 \sum_{j=2}^{r+1} P(Y_1=Y_j=1)}{P(Y_1=1)} \quad (4)$$

In het klassieke geval van onafhankelijkheid reduceert (4) tot:

$$EV_r(u) = \frac{2rP(Y_1=1)P(Y_2=1)}{P(Y_1=1)} = 2rP(Y_1=1). \quad (5)$$

Indien we definiëren

$$V_r^*(u_N) = NV_r(u_N) \quad (6)$$

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 $V_r^*(u)$ de benadering

$$EV_r^*(u) = 2r\tau \quad (7)$$

In het algemene geval dat het optreden van hoge HW-opzetten positief gecorreleerd is, vinden we, voor grote N , de ongelijkheid

$$EV_r^*(u) = \frac{2N^2}{\tau} \sum_{j=2}^{r+1} P(Y_1=Y_j=1) \geq \frac{2N^2}{\tau} \sum_{j=2}^{r+1} P(Y_1=1)P(Y_j=1) = 2r\tau \quad (8)$$

d.w.z. $EV_r^*(u)$ zal, in het geval van positieve afhankelijkheid, groter zijn dan $2r\tau$.

Een eenvoudige berekening leert verder dat in het geval van "m-afhankelijkheid" - d.w.z. (hoge) HW-opzetten welke meer dan m tijdseenheden van elkaar optreden zijn onafhankelijk - geldt

$$E V_r^*(u) = \begin{cases} \approx 2\tau r + 2N \sum_{i=2}^{r+1} \rho_i, & \text{voor } r \leq m \\ \approx 2\tau r + 2N \sum_{i=2}^{m+1} \rho_i, & \text{voor } r > m \end{cases} \quad (9)$$

waarbij $\rho_i = \rho_i(N)$ de correlatiecoëfficiënt tussen Y_i en Y_j voorstelt. Voor grote waarden van N vinden we de benadering:

$$\rho_i = \frac{N}{\tau} P(Y_i = Y_{i+1} = 1) - \frac{\tau}{N} \quad (10)$$

Ook leidt men eenvoudig de volgende ongelijkheid voor de extremaalindex θ van een m -afhankelijke stationaire rij af:

$$1 - \tau^{-1} \lim_{N \rightarrow \infty} N \sum_{i=2}^{m+1} P(Y_i = Y_{i+1} = 1) = 1 - \lim_{N \rightarrow \infty} \sum_{i=2}^{m+1} \rho_i(N) \leq \theta \leq 1 \quad (11)$$

Merk op dat

$P(Y_i = Y_{i+1} = 1) = P(X_i > u_N, X_{i+1} > u_N)$ en dat $u_N = u_N(\tau)$, aangezien $P(X_i > u_N) = \frac{\tau}{N}$. In het geval van onafhankelijkheid ($m=1$) reduceert (11) tot het gewenste resultaat: $\theta=1$. Bij m -afhankelijkheid (met $m>1$) kan de benedengrens in (11) echter in principe ieder waarde van het interval $[0,1]$ aannemen.

5.3.2. Resultaten

Voor het station Hoek van Holland is de functie $V_r^*(u)$ berekend voor verschillende selecties van de HW-opzetten uit de jaren 1887/88...1984/85 met als stormseizoen 1 november...31 januari. Dit is tevens gedaan voor Delfzijl 1885/86...1984/85. De beschouwde selecties zijn: Daan-S- i met $i=0,1,\dots,5$ en als drempels (u) zijn gekozen $u=30,40,\dots,90$ cm.

In tabel 5.3 staan de aantallen van de Daan-S- i selecties weergegeven voor de diverse drempels voor Hoek van Holland. In tabel 5.4 de aantallen voor de Daan-S- i selecties te Delfzijl. In de figuren 5.1...5.4 zijn de resultaten van de berekeningen voor $i=0$ en $i=4$ grafisch weergegeven: voor beide stations zijn voor de vermelde drempels het gemiddelde van $V_r^*(u)$ over de beschouwde stormseizoenen uitgezet tegen r , de afstanden tussen de waarnemingen in de tijd. Naarmate de uitgezette kromme meer op een rechte lijn lijkt, benadert de corresponderende selectie de situatie van onafhankelijkheid beter. Uit de figuren komen de volgende 2 punten duidelijk naar voren:

- de clustering is duidelijker bij hogere drempels ($u=70,80,90$ cm; waarbij opgemerkt dient te worden dat er zeer weinig waarnemingen per stormseizoen zijn voor $u=90$ cm zodat de stochastische fluctuaties duidelijker te zien zijn.
- de krommen benaderen meer een rechte lijn naarmate de i van de selectie groter is. Dit volgt uit formule (9).

In fig. 5.5 zijn voor de drempels $u=90$ cm, voor beide stations, de krommen voor de verschillende selecties bijeengebracht. Hiermee wordt het strekken van de kromme bij toenemende i nog eens duidelijk gedemonstreerd.

5.3.3. Conclusie

Op grond van het voorgaande lijkt selectie van HW-opzetten volgens "Daan-S-4" heel aantrekkelijk. Voor de keuze van deze selectiemethode kunnen ook enkele meteorologische argumenten worden aangereikt (zie hoofdstuk 2). Een keuze voor "Daan-S-1" of "Daan-S-2" is op grond van de hier gegeven statistische analyse niet goed mogelijk (par. 5.2).

Hoek van Holland

Waarnemingen : HW-opzetten
 Selectie : Daan - S - 0
 Stormseizoen : 1 november ... 31 januari (kort stormseizoen)
 Periode : 1887/88 ... 1984/85
 Uitgezet : functie $V_r'(u_N)$ tegen r voor verschillende drempels u_N

In de laatste figuur zijn alle 7 curves tegelijk uitgezet met van boven naar beneden resp. drempel $u = 30, 40, \dots, 90$ cm.

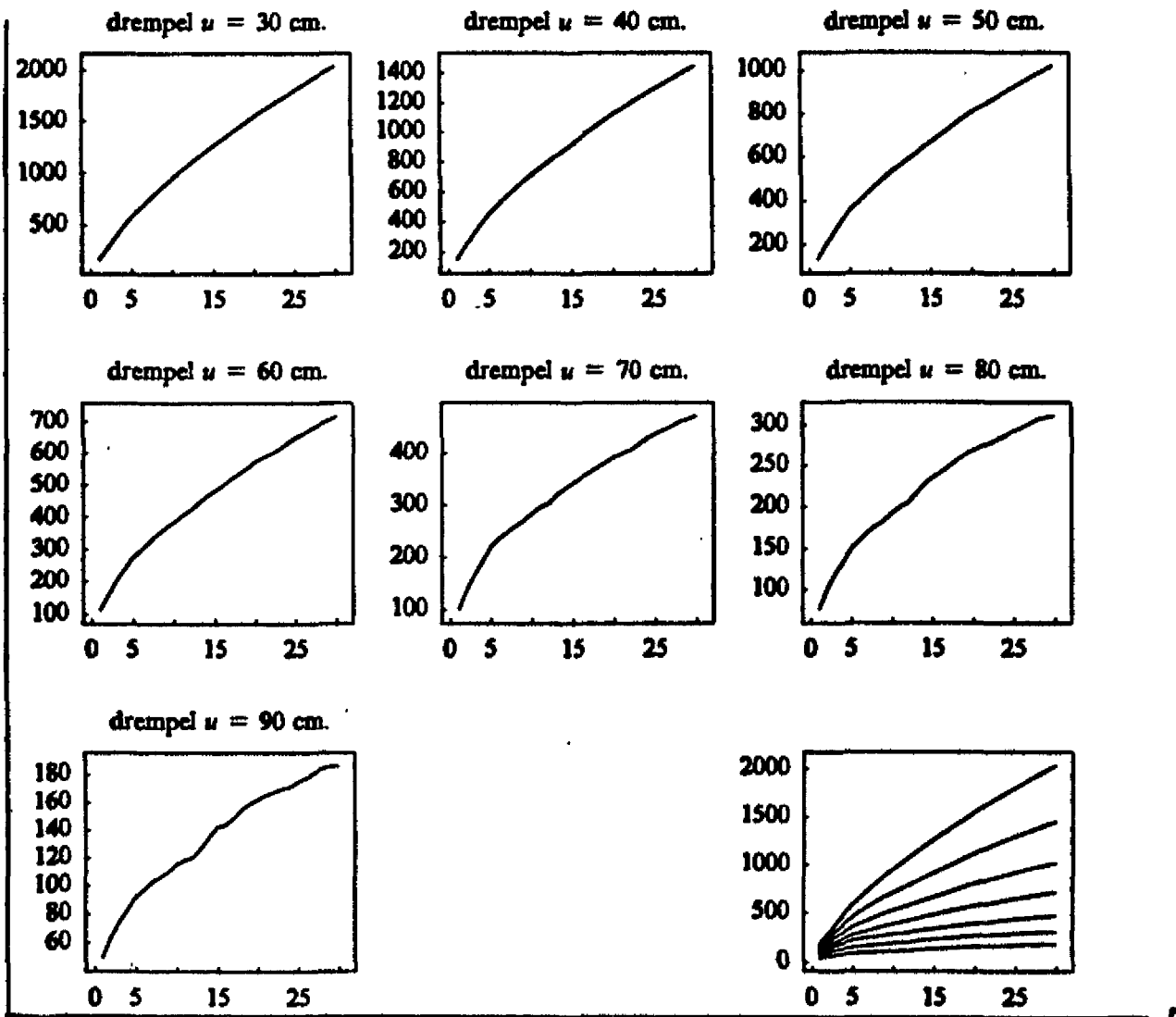
 functie $V_r'(u_N)$


Fig. 5.1. Verloop van $V_r'(u_N)$ als functie van het tijdsinterval r voor Hoek van Holland en Daan-S-0 bij verschillende drempelwaarden.

Hoek van Holland

Waarnemingen : HW-opzetten
 Selectie : Daan - S - 4
 Stormseizoen : 1 november ... 31 januari (kort stormseizoen)
 Periode : 1887/88 ... 1984/85
 Uitgezet : functie $V_r'(u_N)$ tegen r voor verschillende drempels u_N

In de laatste figuur zijn alle 7 curves tegelijk uitgezet met van boven naar beneden resp. drempel $u = 30, 40, \dots, 90$ cm.

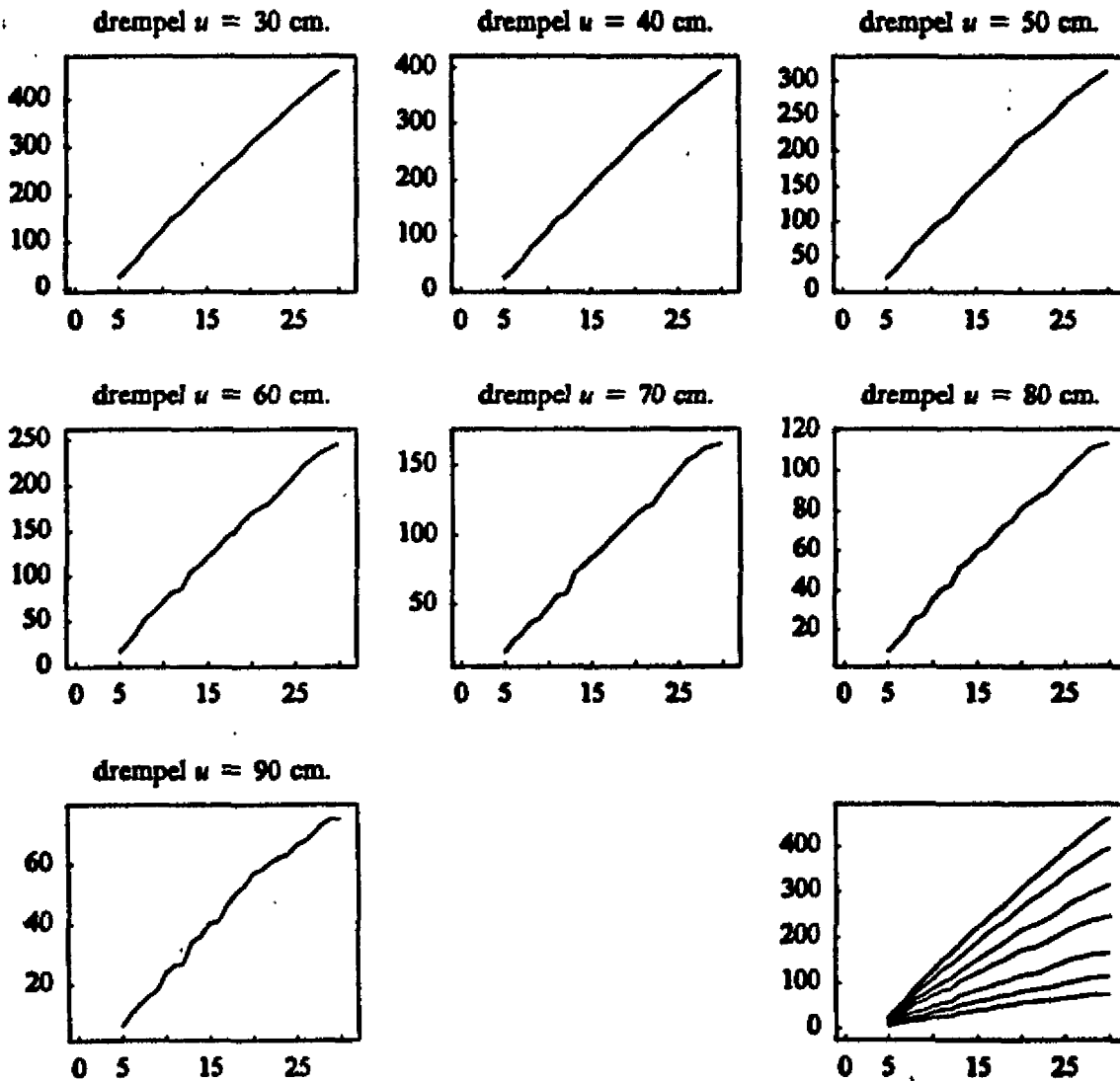
 functie $V_r'(u_N)$


Fig. 5.2. Verloop van $V_r'(u_N)$ als functie van het tijdsinterval r voor Hoek van Holland en Daan-S-4 bij verschillende drempelwaarden.

Delfzijl

Waarnemingen : HW-opzetten
 Selectie : Daan - S - 0
 Stormseizoen : 1 november ... 31 januari (kort stormseizoen)
 Periode : 1885/86 ... 1984/85
 Uitgezet : functie $V_r(u_N)$ tegen r voor verschillende drempels u_N

In de laatste figuur zijn alle 7 curves tegelijk uitgezet met van boven naar beneden resp. drempel $u = 30, 40, \dots, 90$ cm.

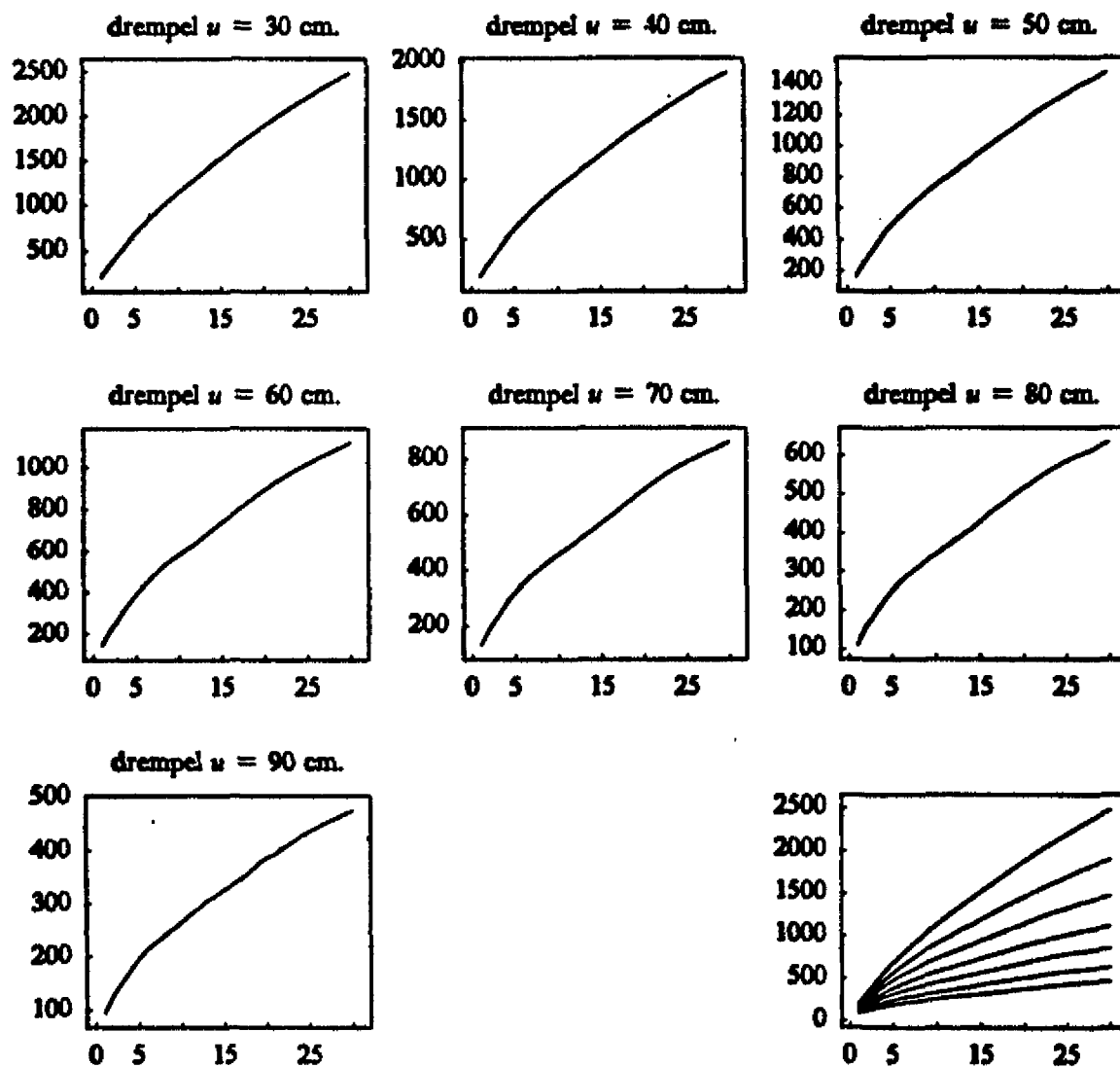
 functie $V_r(u_N)$


Fig. 5.3. Verloop $V_r(u_N)$ als functie van het tijdsinterval r voor Delfzijl en Daan-S-0 bij verschillende drempelwaarden.

Delfzijl

Waarnemingen : HW-opzetten
 Selectie : Daan - S - 4
 Stormseizoen : 1 november ... 31 januari (kort stormseizoen)
 Periode : 1885/86 ... 1984/85
 Uitgezet : functie $V_r^2(u_N)$ tegen r voor verschillende drempels u_N

In de laatste figuur zijn alle 7 curves tegelijk uitgezet met van boven naar beneden resp. drempel $u = 30, 40, \dots, 90$ cm.

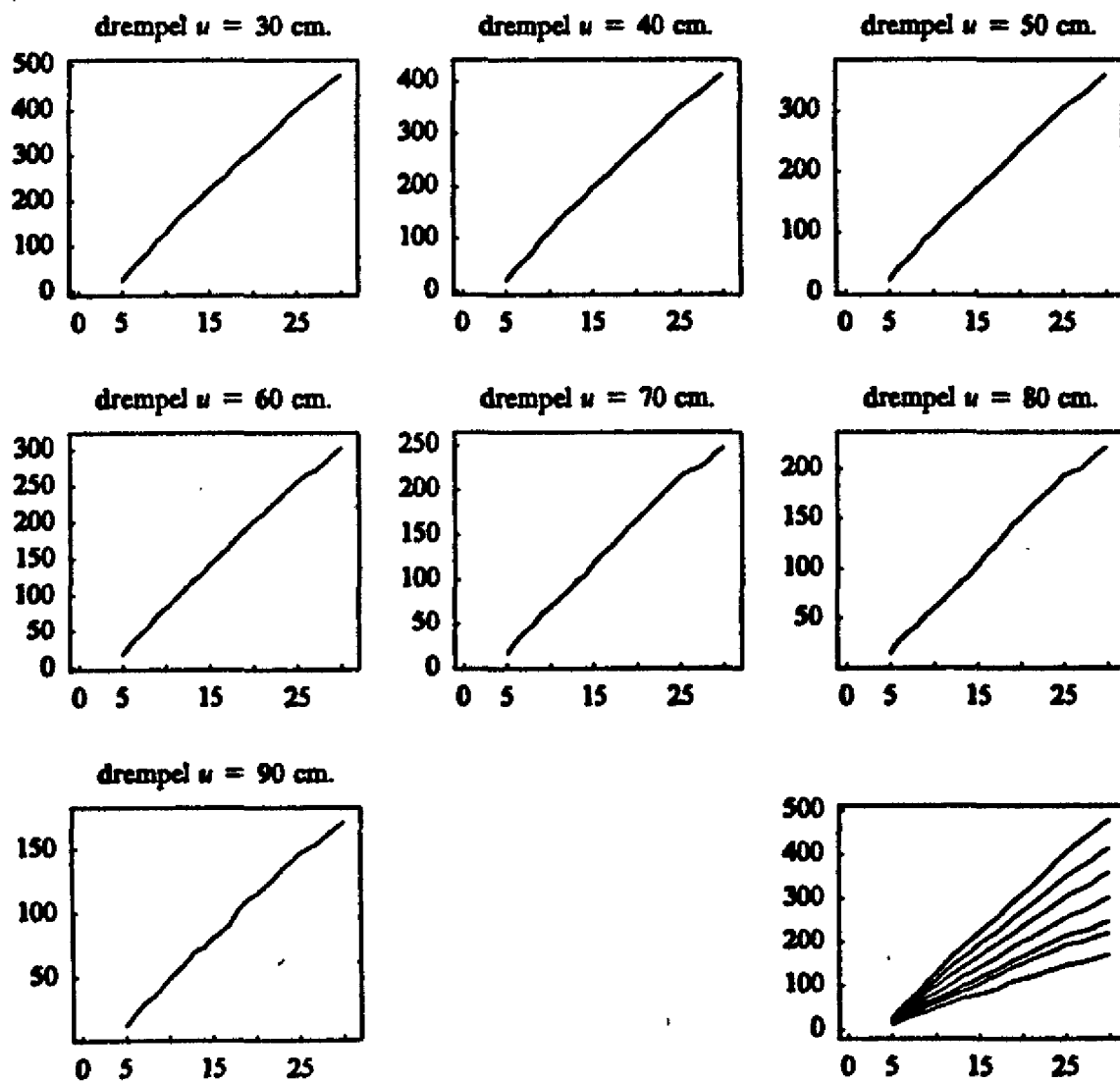
functie $V_r^2(u_N)$ 

Fig. 5.4. Verloop van $V_r^2(u_N)$ als functie van het tijdsinterval r voor Delfzijl en Daan-S-4 bij verschillende drempelwaarden.

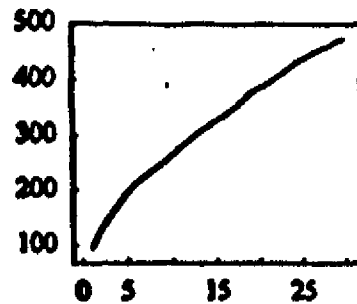
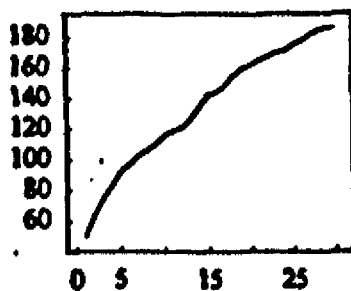
drempel $\mu = 90$ cm.

Hoek van Holland

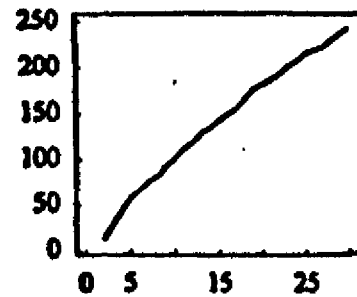
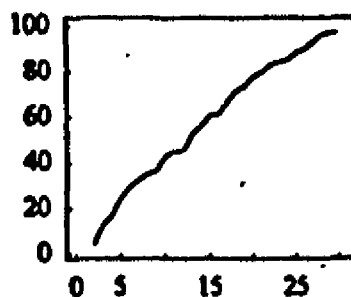
Delfzijl

functie $V_r(\mu_N)$

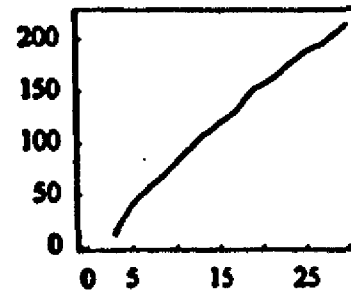
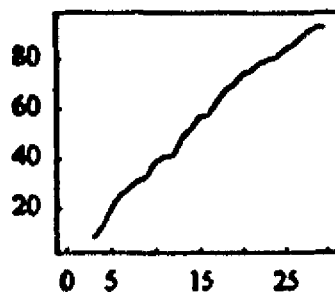
Daan - S - 0



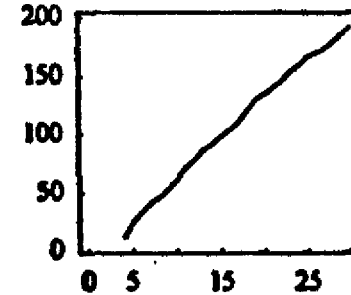
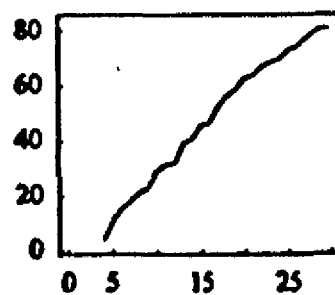
Daan - S - 1



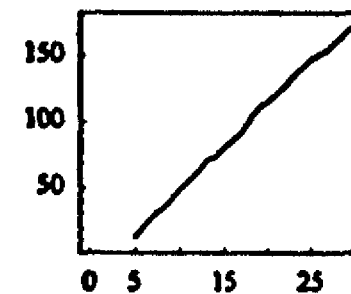
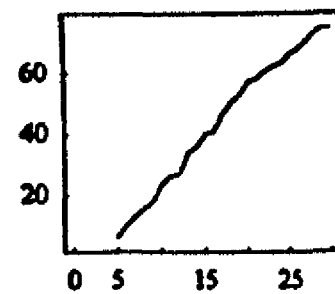
Daan - S - 2



Daan - S - 3



Daan - S - 4



Daan - S - 5

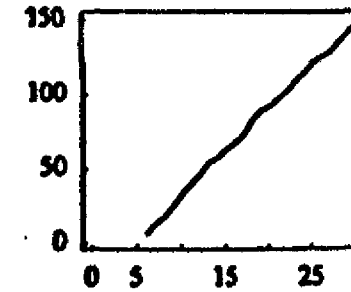
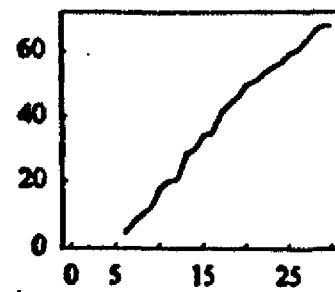


Fig. 5.5. Verloop van $V_r(\mu_N)$ als functie van het tijdsinterval r voor beide stations bij een drempelwaarde $\mu = 90$ cm.

Tabellen Weergegeven staan de aantallen geselecteerde HW-opzetten \geq drempel u in cm. van de verschillende D-S-i selecties.

Tabel 5.3 Hoek van Holland, 1887/88...1984/85 met stormseizoen 1 november ... 31 januari.

| u | Daan-S-0 | Daan-S-1 | Daan-S-2 | Daan-S-3 | Daan-S-4 | Daan-S-5 |
|----|----------|----------|----------|----------|----------|----------|
| 30 | 2935 | 1618 | 1357 | 1107 | 944 | 835 |
| 40 | 1996 | 1172 | 1017 | 866 | 765 | 684 |
| 50 | 1360 | 856 | 765 | 674 | 610 | 555 |
| 60 | 948 | 625 | 577 | 520 | 484 | 447 |
| 70 | 648 | 443 | 418 | 386 | 362 | 336 |
| 80 | 453 | 329 | 315 | 297 | 283 | 269 |
| 90 | 299 | 232 | 228 | 220 | 214 | 207 |

Tabel 5.4 Delfzijl, 1885/86...1984/85 met als stormseizoen 1 november ...31 januari.

| u | Daan-S-0 | Daan-S-1 | Daan-S-2 | Daan-S-3 | Daan-S-4 | Daan-S-5 |
|----|----------|----------|----------|----------|----------|----------|
| 30 | 3623 | 1743 | 1484 | 1221 | 1007 | 895 |
| 40 | 2660 | 1363 | 1181 | 1002 | 843 | 759 |
| 50 | 1992 | 1079 | 958 | 830 | 714 | 639 |
| 60 | 1478 | 852 | 772 | 682 | 604 | 546 |
| 70 | 1116 | 668 | 610 | 552 | 499 | 454 |
| 80 | 824 | 526 | 492 | 457 | 428 | 392 |
| 90 | 641 | 432 | 407 | 382 | 362 | 338 |

5.4 Extreme waarden theorie voor afhankelijke waarnemingen en de selectie methode Daan-1

In de tot dusver in het frequentielijnen onderzoek verrichte werkzaamheden speelt een selectiemethode, die is voorgesteld en nader uitgewerkt door H. Daan (KNMI), een belangrijke rol. Selectie van hoogwaterstanden (HW's) of HW-opzetten vindt plaats omdat achtereenvolgende HW's of opzetten niet beschouwd kunnen worden als onafhankelijke waarnemingen met dezelfde kansverdeling. In de klassieke extreme waarden theorie wordt echter van dit model uitgegaan. Wanneer we de resultaten van deze theorie willen toepassen zal eerst via selectie het waarnemingsmateriaal gereduceerd moeten worden tot onafhankelijke waarnemingen met onderling dezelfde verdelingsfunctie. Door beperking tot HW's of HW-opzetten in de wintermaanden worden -bij benadering- waarnemingen met dezelfde kansverdeling verkregen (homogeniteit). Vervolgens wordt er van de waarnemingen welke in de tijd dicht bij elkaar liggen slechts één in de selectie opgenomen. Op deze wijze wordt -hopelijk- onafhankelijkheid van de waarnemingen gegarandeerd. Op basis van de geselecteerde HW's en opzetten zijn via het in lit. 13 beschreven model o.a. schattingen gemaakt van het 10^{-4} kwantiel bij Hoek van Holland en Vlissingen.

Een andere -meer algemene- aanpak van ons probleem is denkbaar, waarbij het mogelijk lijkt de statistische analyse te baseren op alle HW's of opzetten in de winterperioden 1901-1980.x) In het tot dusver gehanteerde model wordt aangenomen dat de geselecteerde waarnemingen X_1, \dots, X_n beschouwd kunnen worden als een aselechte steekproef van omvang n uit een populatie met (onbekende) verdelingsfunctie F . (M.a.w. de geselecteerde X_i 's stellen in ons model onafhankelijke waarnemingen met gemeenschappelijke verdelingsfunctie F voor). We zullen nu in plaats hiervan uitgaan van de (veel omvangrijkere) oorspronkelijke collectie HW's of HW-opzetten in de winterperioden 1901-1980 en veronderstellen dat we te maken hebben met een stationaire rij X_1, \dots, X_N met noodzakelijk gemeenschappelijke marginale verdelingsfunctie F voor de afzonderlijke waarnemingen; m.a.w. : $P(X_i \leq x) = F(x)$ voor alle i en x , verder geldt dat de simultane verdelingsfunctie van de HW's (of opzetten): $(X_{i_1}, \dots, X_{i_n})$ geobserveerd op tijdstippen (i_1, \dots, i_n) , welke gegeven worden door

$$F_{i_1, \dots, i_n}(x_1, \dots, x_n) = P(X_{i_1} \leq x_1, \dots, X_{i_n} \leq x_n) \quad (1)$$

dezelfde blijft indien we i.p.v. HW's (of HW-opzetten) op de tijdstippen (i_1, \dots, i_n) de HW's (of opzetten) op de tijdstippen (i_1+a, \dots, i_n+a) , voor iedere $a \neq 0$ bekijken.

Leadbetter (lit. 7) heeft recentelijk een extreme waarden theorie voor bovenstaande situatie ontwikkeld. Zij (X_1, \dots, X_N) de collectie van alle HW's (of HW-opzetten) in Hoek van Holland of Vlissingen, in de periode 1901-1980, opgetreden in de winterperiode 31 oktober t/m 1 februari.

We zullen nu veronderstellen dat (X_1, \dots, X_N) N achtereenvolgende waarnemingen in een stationaire rij voorstellen. Verder nemen we aan dat de X_i 's die in de tijd ver uit elkaar liggen vrijwel onafhankelijk zijn. (Dit is voorwaarde (D) van Leadbetter; deze voorwaarde is zeer zwak en lijkt in ons geval géén probleem). Onder een kleine extra aanname - namelijk dat de stationaire rij $\{X_n\}$ een extremaal index $\theta > 0$ heeft, dit betekent dat indien $\{u_N\}$ een rij constanten is met de eigenschap $N(1-F(u_N)) \rightarrow \tau$, voor een zekere $0 \leq \tau < \infty$ - dan geldt

x) Dit onderzoek werd verricht vóór de te gebruiken waarnemingsperioden werden vastgesteld. Zie par. 4.1.

$$P(\max_{1 \leq i \leq N} X_i \leq u_N) \rightarrow e^{-\theta} \quad (2)$$

voor een zekere θ met $0 < \theta \leq 1$ blijkt te gelden dat

$$P\left(\frac{\max_{1 \leq i \leq N} X_i - b_N}{a_N} \leq y\right) \approx \left[P\left(\frac{\max_{1 \leq i \leq N} \hat{X}_i - b_N}{a_N} \leq y\right) \right]^\theta \quad (3)$$

voor alle y en N groot. De grootheden $\hat{X}_1, \hat{X}_2, \dots$ stellen fictieve onafhankelijke stochastische grootheden voor, met als gemeenschappelijke verdelingsfunctie F , de (marginale) verdelingsfunctie van de afzonderlijke grootheden X_i ($i = 1, 2, \dots$) in de stationaire rij $\{X_n\}$. (Men noemt $\{\hat{X}_n\}$ de met de stationaire rij $\{X_n\}$ geassocieerde onafhankelijke rij).

Merk op dat relatie (3) onmiddellijk impliceert dat het $(1-p)$ -de kwantiel van de verdeling van het genormaliseerde maximum gegeven wordt door $G_k^{-1}((1-p)^{\theta^{-1}})$,

waarbij G_k de extreme waarden verdeling (met parameter k) in het klassieke geval van onafhankelijkheid voorstelt. Omdat $0 < p < 1$ en $\theta^{-1} \geq 1$ volgt dat $(1-p)^{\theta^{-1}} \leq (1-p)$: de afhankelijkheid van de waarnemingen leidt dus tot lagere waarden voor het kwantiel, in vergelijking met de situatie van onafhankelijkheid.

Indien $\theta = 1$ dan zijn we terug in de situatie die beschreven wordt in de klassieke extreme waarden theorie: de kansverdeling van $\max_{1 \leq i \leq N} X_i$ kan bepaald worden alsof de X_i 's onafhankelijke waarnemingen met dezelfde verdelingsfunctie F voorstellen. Dit is het model dat in lit. 13 gehanteerd is voor de geselecteerde HW's en opzetten. Indien we echter uitgaan van alle HW's of opzetten in de winterperioden 1901-1980, dan gelden de relaties (2) en (3) voor een of andere $\theta \in (0, 1]$. (De aanname dat $\{X_n\}$ dan een stationaire rij is met een extremaal index $\theta > 0$ lijkt niet onaannemelijk, mits de HW's eerst gecorrigeerd zijn voor de invloed van de zeespiegelrijzing). Combineren we relatie (3) met bijvoorbeeld formule (2.5) (zie ook (2.6) uit lit. 13 dan vinden we de benadering (zie (3.1) van lit. 13):

$$P(\max_{1 \leq i \leq N} X_i \geq x) \approx 1 - \left\{ \exp\left[-N\left(1 - \frac{kx}{\sigma}\right)^{\frac{1}{k}}\right] \right\}^\theta \quad (4)$$

voor de kans dat het maximum van de N HW's (precieser: overschrijdingen van een drempel u met als marginale verdelingsfunctie een GPD-verdeling) een hoogte x overschrijdt. Hierbij stellen $\sigma > 0$ en $-\infty < k < \infty$ de onbekende schaal- en vormparameter van een GPD-verdeling voor; de parameter θ , de extremaal index, is eveneens onbekend en moet in principe op basis van de X_i 's geschat worden, evenals trouwens de parameters σ en k .

Hoe kan nu extremaal index θ geschat worden? Leadbetter (lit. 7) heeft laten zien dat θ geïnterpreteerd kan worden als de inverse van het gemiddelde aantal overschrijdingen van een (hoog) niveau u_N in een "cluster" van een aantal -zeg r_N - in tijd opeenvolgende HW's (of opzetten). Het aantal HW's in een cluster - r_N - moet zo gekozen worden dat indien alle r_N HW's in een cluster vervangen gedacht worden door één HW, -geobserveerd op (een of ander) tijdstip behorende tot het tijdsinterval waarop de tot de cluster behorende HW's (of opzetten) waargenomen zijn- de verzameling van aldus verkregen HW's -één HW per cluster, d.w.z. $k_N = \frac{r_N}{N}$ HW's- beschouwd kunnen worden als vrijwel onafhankelijk; d.w.z. voor r_N de k_N op deze wijze geselecteerde HW's (één per cluster) is de klassieke extreme waarden theorie toepasbaar en kan $\theta = 1$ genomen worden.

We merken nu op dat de hiervoor beschreven clustervorming van in de tijd dicht bij elkaar liggende HW's overeen lijkt te stemmen met een data-onafhankelijke versie voor de door H. Daan geïntroduceerde selectiemethode, met $r_N = 2i + 1$ in de Daan-i-selectie. Het voorgaande levert ons nu een methode om een beter inzicht te krijgen in de Daan-i-selecties. In Tabel 5.5 is voor

$i = 2, 3, \dots, 9$ de waarde van θ geschat voor het geval de Daan- i -selectie gebruikt is om de clusters te vormen. Een ruwe schatting van θ wordt eenvoudig verkregen door het aantal geselecteerde HW's in de Daan- i -selectie te delen door het totaal aantal HW's dat beschouwd wordt (d.w.z. de Daan-0 selectie). We vinden daarvoor voor θ van de orde $\approx 0.5-0.7$.

Ook werd deze procedure uitgevoerd voor het geval de Daan-5 selectie als basismateriaal wordt genomen en een Daan- i selectie, voor $i = 5, 6, 7, 8, 9$, wordt uitgevoerd. In tabel 5.6 zien we dat θ van de orde ≈ 0.9 à 1 is. Dit laatste resultaat suggereert dat de Daan-5 selectie -bij benadering- inderdaad de eigenschap heeft dat de tot de Daan-5 selectie behorende HW's vrijwel onafhankelijk zijn. Onafhankelijkheid van de geselecteerde HW's in de Daan-5 selectie lijkt bij benadering in orde te zijn. Anderzijds vinden we $\theta \approx 0.5$ à 0.7 wanneer alle HW's bekeken worden en de clustering m.b.v. Daan-5 uitgevoerd wordt. Ruwweg betekent dit dat we in principe het 10^{-4} -kwantiel ook zouden kunnen schatten met behulp van formule (4) met een $\theta \approx 0.5$ à 0.7 op basis van alle HW's in de winterperioden 1901-1980.

We concluderen dat Laedbetter's theorie voor extreme waarden bij afhankelijke waarnemingen een theoretische rechtvaardiging geeft voor de Daan- i selectie en, naar men mag verwachten, ook voor de Daan- S - i -selectie: Selectie van de grootste HW-opzet per "cluster" (storm) leidt tot geselecteerde HW-opzetten, welke opgevat kunnen worden als steekproef van onafhankelijke waarnemingen. De limietverdeling van het maximum van de op deze wijze geselecteerde waarnemingen is een extreme waarden verdeling met parameter k . Het aantal geselecteerde waarnemingen is bij benadering een fractie θ (de extremaalindex) van het oorspronkelijke aantal waarnemingen N .

Tabel 5.5. Berekening van

$$\theta = \frac{\text{aantal geselecteerde waarnemingen Daan-}i \text{ selectie}}{\text{aantal geselecteerde waarnemingen Daan-0 selectie}}$$

voor Hoek van Holland, periode 1901-1980, HW's ≥ 170 cm + NAP en opzet ≥ 50 cm.

| Tabel 1 | | θ | | | | | |
|---------|-------|----------|-------|-------|-------|-------|-------|
| mhw | 165 | 170 | 175 | 180 | 185 | 190 | 194 |
| daan-2 | 0.648 | 0.659 | 0.689 | 0.700 | 0.725 | 0.743 | 0.757 |
| daan-3 | 0.606 | 0.627 | 0.648 | 0.655 | 0.675 | 0.705 | 0.724 |
| daan-4 | 0.572 | 0.592 | 0.617 | 0.628 | 0.650 | 0.675 | 0.687 |
| daan-5 | 0.522 | 0.536 | 0.563 | 0.583 | 0.604 | 0.633 | 0.659 |
| daan-6 | 0.508 | 0.526 | 0.555 | 0.568 | 0.586 | 0.616 | 0.640 |
| daan-7 | 0.494 | 0.517 | 0.545 | 0.553 | 0.575 | 0.608 | 0.636 |
| daan-8 | 0.478 | 0.502 | 0.530 | 0.544 | 0.568 | 0.599 | 0.621 |
| daan-9 | 0.467 | 0.491 | 0.522 | 0.538 | 0.561 | 0.591 | 0.617 |

Tabel 5.6: Berekening van

$$\theta = \frac{\text{aantal geselecteerde waarnemingen Daan-1 selectie}}{\text{aantal geselecteerde waarnemingen Daan-5 selectie}}$$

voor Hoek van holland, periode 1901-1980, HW's ≥ 170 cm. + NAP en opzet ≥ 50 cm.

| Tabel 2 | | θ | | | | | |
|---------|-------|----------|-------|-------|-------|-------|-------|
| mhw | 165 | 170 | 175 | 180 | 185 | 190 | 194 |
| daan-5 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| daan-6 | 0.973 | 0.981 | 0.986 | 0.974 | 0.970 | 0.973 | 0.972 |
| daan-7 | 0.945 | 0.965 | 0.968 | 0.948 | 0.953 | 0.960 | 0.965 |
| daan-8 | 0.914 | 0.938 | 0.941 | 0.933 | 0.941 | 0.947 | 0.943 |
| daan-9 | 0.894 | 0.918 | 0.927 | 0.923 | 0.929 | 0.933 | 0.936 |

6. INVLOED VAN DE DREMPELWAARDE

6.1 Inleiding

De volgens de in hoofdstuk 5 geselecteerde elementen (hoogwateropzetten) kunnen tot verschillende verdelingen behoren die ieder in een bepaald bereik domineren. Indien de opzetten dan toch tot een en dezelfde populatie (met een daarvoor afgeleide frequentieverdeling) worden gerekend, kan een foutieve extrapolatie naar zeldzame waarden het gevolg zijn. Omdat, uiteraard, juist de hogere opzetten van belang zijn is het wenselijk de populatie met de lagere opzetten te negeren. De vraag is echter waar de grens ligt, die als drempelwaarde voor verdere selectie kan fungeren. Hieraan is par. 6.2 gewijd.

Par. 6.3 bevat een beschouwing over het effect van de drempelhoogte op de naar een frequentie 10^{-4} /jaar geëxtrapolerde waarde. Indien de opzetten in feite volgens GPV zijn verdeeld (zie lit. 13, hoofdstuk 4), maar als hierop een exponentiele verdeling wordt toegepast dan neemt de 10^{-4} waarde af bij oplopende drempelwaarde. Indien echter op de overblijvende opzetten weer een GPV-verdeling wordt toegepast dan lijkt de afhankelijkheid van de 10^{-4} waarde voor de drempelhoogte minder groot of zelfs afwezig. Voor een definitieve conclusie is nog nader onderzoek gewenst.

6.2 Het drempelniveau voor de HW-opzet

Het drempelniveau is de ondergrens van de (geselecteerde) waarnemingen met behulp waarvan de extreme waarden worden geschat. De waarde ervan heeft nogal veel invloed op deze schattingen. Een duidelijk criterium aan de hand waarvan een keuze gemaakt kan worden ontbreekt nog.

Er zijn twee redenen te noemen voor het hanteren van een drempel:

1. Voor de selectie van onafhankelijke HW-opzetten (Daan-selectie). Bij de v.d. Ham-selectie werden alleen depressies met een opzet van minstens 50 cm geselecteerd, enerzijds omdat lichte depressies moeilijk te herkennen en van elkaar te onderscheiden zijn, anderzijds omdat daardoor een aanzienlijke werkbesparing werd verkregen. Bovendien werden lage opzetten niet van belang geacht voor de extrapolatie (zie rapport Deltacommissie deel 3, blz. 31). Vanwege de huidige geautomatiseerde selectiemethode vervalt het argument van de werkbesparing. Er wordt daardoor echter niet meer gekeken naar meteo-omstandigheden, maar puur numeriek geselecteerd op basis van een rij getallen (HW-opzetten). Vanwege diverse versturende invloeden hebben de berekende HW-opzetten een "ruis"-component. De belangrijkste zijn:
 - a. de invloed van luchtdrukvariatiën op de hoogwaterstand (1 mbar = 1 cm waterkolom). Het effect hiervan kan 20 à 30 cm bedragen. De standaardafwijking σ_a wordt geschat op 10 cm.
 - b. Onnauwkeurigheden in de berekende astronomische hoogwaters. Het effect hiervan kan ook 20 à 30 cm bedragen. De standaardafwijking σ_b wordt eveneens geschat op 10 cm.
 Combinatie van de beide foutenbronnen geeft een ruis op de opzetten met een standaardafwijking van $\sqrt{10^2 + 10^2} = 14$ cm. Hierdoor is het mogelijk dat er opzetten berekend zijn zonder dat er in werkelijkheid sprake is van opwaaiing. Toepassing van 30 cm ($\approx 2\sigma$) als minimale waarde voor de drempel geeft een vrij grote zekerheid dat we met werkelijke opzetten te maken hebben.

2. De verdeling van de lage opzetten zegt mogelijk niets over de verdeling van de extreem hoge opzetten. Dat is met name het geval wanneer er beneden een bepaald niveau duidelijk sprake is van een afwijkende verdeling. Om dit na te gaan zijn voor de stations Vlissingen, Hoek van Holland en Delfzijl zowel de geselecteerde (Daan-S-4) als de niet geselecteerde HW-opzetten uitgezet op logaritmisch papier (fig. 6.1 t/m 6.3). Een sterke afbuiging of knik in het verloop van de uitgezette punten duidt op mogelijke afwijkende verdelingen voor hoge en lage opzetten. Uit de figuren volgt echter geen duidelijke aanwijzing voor een te hanteren drempelniveau boven de onder 1. genoemde minimale waarde.

Conclusie:

1. Bij de selectiemethode een drempel hanteren van 30 cm.
2. Bij de verdere analyses verschillende drempels toepassen met een minimum van 30 cm. De uiteindelijke keuze, die overigens per station kan verschillen, zal moeten worden gemaakt op grond van de stabiliteit van de schatters als functie van het drempelniveau en van de onderlinge vergelijking van de verschillende schattingsmethodes van de kwantielen.

VLISSINGEN 1881-1985

HW-opzet tijdens stormseizoen (1 okt - 15 mrt)

HW-opzet in cm

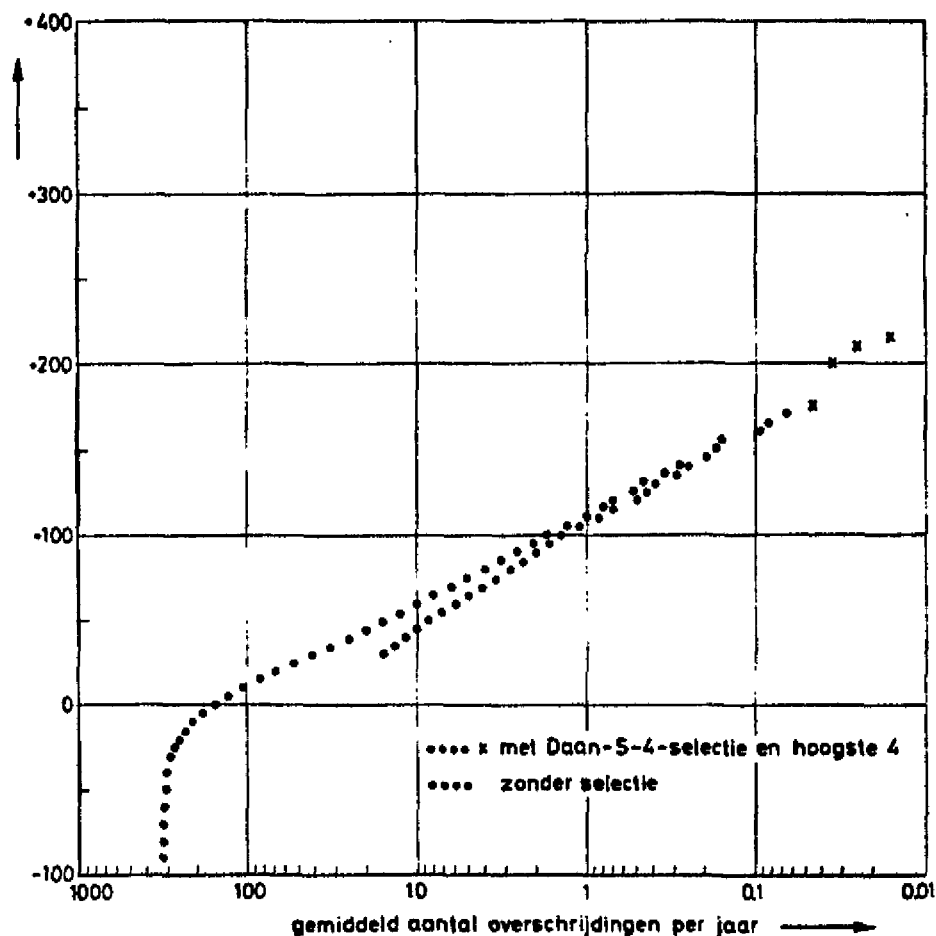


Fig. 6.1

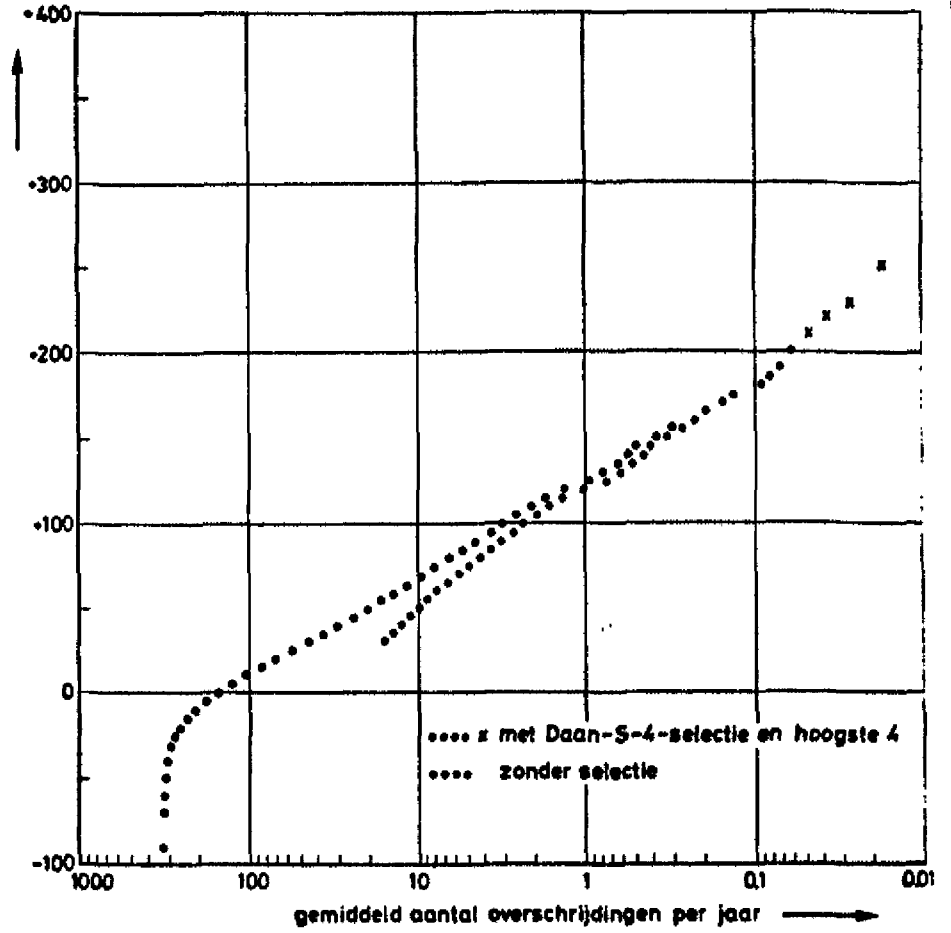
HW-opzet in cm

Fig. 6.2

HOEK VAN HOLLAND

1887-1985

HW-opzet tijdens
stormselzoen (1 okt - 15 mrt)

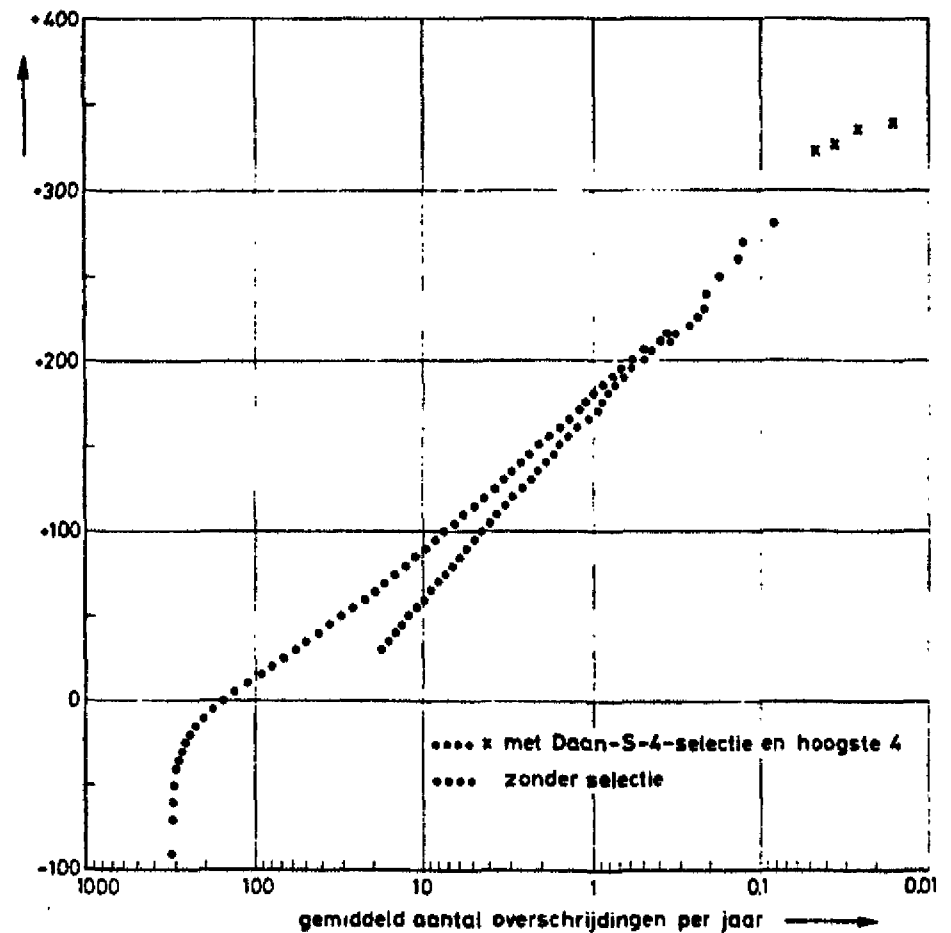


HW-opzet in cm

Fig. 6.3

DELFSJUL 1881-1985

HW-opzet tijdens
stormselzoen (1 okt - 15 mrt)



6.3 Onderzoek naar het verloop van het 10^{-4} -kwantiel bij oplopende drempel.

6.3.1. Vroegere resultaten

In de bijlagen 6 en 8 van het interimrapport (lit. 13) zijn de puntschattingen van het 10^{-4} -kwantiel voor HW's en HW-opzetten voor zowel het Gumbel - als het EV-model grafisch weergegeven voor de verschillende selectiemethoden. De puntschattingen zijn uitgezet tegen de drempelwaarden. In deze notitie bekijken we het verloop van de puntschattingen van het 10^{-4} -kwantiel voor HW's en HW-opzetten voor zowel het Gumbel - als het EV-model als functie voor de drempelwaarde.

6.3.2 Het geschatte 10^{-4} -kwantiel als functie van de drempelwaarde in het Gumbelmodel

In deze paragraaf laten we, voor het geval van het Gumbelmodel, zowel door een theoretische analyse als ook door simulaties, zien dat het dalende verloop van de puntschattingen van het 10^{-4} -kwantiel voor HW's en HW-opzetten bij oplopende drempelwaarde goed verklaard kan worden door aan te nemen dat de HW's (of HW-opzetten), welke boven een drempel u uitsteken, verminderd met de waarde u , een GPV verdeling bezitten.

Gegeven is een aselechte steekproef X_1, \dots, X_n van omvang n uit een populatie met verdelingsfunctie F . Stel een drempelwaarde u vast en zij Y_1, \dots, Y_N de waarnemingen die boven de drempel u uitsteken, verminderd met de drempelwaarde u . N is het (stochastische) aantal overschrijdingen van de drempel u in de steekproef X_1, \dots, X_n van omvang n . Als N vast gedacht wordt, zijn de Y_i 's onafhankelijk en hebben ze allen dezelfde kansverdeling, welke eenvoudig uit F en u bepaald kan worden. Deze kansverdeling kunnen we benaderen door $H_k(x)$ (zie 4.2) van lit. 13, de GPV verdeling. Voor het speciale geval $k=0$ reduceert H_k tot de exponentiele verdeling en kan het 10^{-4} -kwantiel voor HW's en HW-opzetten geschat worden d.m.v.

$$\hat{h}_n(p) = -\frac{\sum_{i=1}^N Y_i}{N} \ln \left[\frac{mp}{N} \right] + u \quad (2.1)$$

waarbij $p=10^{-4}$, m het aantal jaren voorstelt en $N=N(u)$ = aantal overschrijdingen van drempel u . Merk op dat (2.1) de formule is welke in het Deltarapport wordt gebruikt; relatie (2.1) volgt ook uit (4.4) van (lit. 13) als we $k \rightarrow 0$ laten gaan en $\ln(1-p)$ met $-p$ benaderen. De Gumbelgrafieken uit de bijlagen 6 en 7 van lit. 13 zijn m.b.v. (2.1) bepaald.

Veronderstel nu dat de Y_i 's, in feite GPV verdeeld zijn, maar dat -evenals in het Deltarapport en in de bijlage 6 van lit. 13 - formule (2.1) gebruikt wordt voor de schatting van het 10^{-4} -kwantiel van de HW's (basispeil) of HW-opzetten. We weten nu dat

$$EY_i = E(X_i - u | X_i > u) = \frac{\sigma}{k+1} \quad (2.2)$$

en vinden eenvoudig m.b.v. (2.1) dat

$$E\hat{h}_n(p) = E \left[-\frac{\sigma}{k+1} \ln \left[\frac{mp}{N(u)} \right] \right] + u \quad (2.3)$$

waarbij

$$E \frac{N(u)}{n} = 1 - F(u) \quad (2.4)$$

Een Taylorbenadering geeft

$$\ln \left[\frac{mp}{N(u)} \right] = \ln \left[\frac{mp}{n} \right] - \ln \left[\frac{N(u)}{n} \right] = \ln \left[\frac{mp}{n} \right] - \ln(1-F(u)) - \left[\frac{N(u)}{n} - (1-F(u)) \right] \cdot \frac{1}{1-F(u)} \quad (2.5)$$

+ lagere orde termen

We weten nu dat $N(u)/n$, voor grote n , met grote kans waarden dichtbij $1-F(u)$ zal aannemen. M.a.w.: we mogen verwachten dat indien $N(u)$ in (2.3) vervangen wordt door $n(1-F(u))$ een fout gemaakt wordt die voor voldoende grote n willekeurig klein wordt.

We vinden dus:

$$E\hat{h}_n(p) = -\frac{\sigma}{k+1} \ln \left[\frac{mp}{n(1-F(u))} \right] + u \quad (2.6)$$

Dit suggereert dat ook geldt

$$\frac{d}{du} E\hat{h}_n(p) = -\frac{\sigma}{k+1} \frac{f(u)}{1-F(u)} + 1 \quad (2.7)$$

waarbij $r(u) = f(u)/(1-F(u))$ de "failure rate" van F voorstelt.

Indien $k=0$ is, is zowel F als ook de gemeenschappelijke verdeling van de y_i 's exponentieel verdeeld en de failure rate r constant. We vinden dan $r(u) = \sigma^{-1}$ en vinden m.b.v. (2.7) $\frac{d}{du} E\hat{h}_n(p) = 0$. M.a.w. het gemiddelde geschatte 10^{-4} -kwantiel van HW-opzetter is constant, d.w.z. onafhankelijk van de drempelwaarde u .

Indien echter $k > 0$ is, dan geldt (onder een kleine extra aanname - (zie Thm 1.6.1 van lit. 8)); we gebruiken hier een eenvoudige voldoende voorwaarde voor F , opdat het genormeerde maximum van een aselechte steekproef met F voor grote waarden van n benaderd wordt door een extreme waarden verdeling met vormparameter $k > 0$)

$$\frac{f(u)}{1-F(u)} = \frac{1}{k(x_f - u)} \quad (2.8)$$

voor drempels u in de buurt van x_f , waarbij x_f de bovengrens van de support van F voorstelt.

Combineren we (2.7) en (2.8) dan vinden we dat

$$\frac{d}{du} (E\hat{h}_n(p)) = -\frac{\sigma}{k(k+1)(x_f - u)} + 1 \quad (2.9)$$

zodat als u voldoende dicht bij x_f gekozen wordt

$$\frac{d}{du} (E\hat{h}_n(p)) \quad (2.10)$$

in eerste benadering negatief wordt.

We kunnen concluderen dat, als de HW's (of HW-opzetten) in feite GPV verdeeld zijn met vormparameter $k > 0$, het gemiddelde geschatte basispeil $Eh_n(p)$, voor voldoende grote u , een dalende functie van de drempel u is. Een grafische weergave voor de Daan-5-selectie wordt gegeven in de figuren 6.1 en 6.2.

Een en ander wordt ook bevestigd door de uitgevoerde simulaties (zie paragraaf 4).

6.3.3 Het geschatte 10^{-4} -kwantiel als functie van de drempelwaarde in het EV-model

We beschouwen dezelfde situatie als in paragraaf 2, maar i.p.v. met formule (2.1), schatten we het 10^{-4} -kwantiel van de HW's (of HW-opzetten) met

$$\hat{h}_n(p) = \frac{\hat{\sigma}_N}{\hat{k}_N} \cdot \left\{ 1 - \left[\frac{-m \log(1-p)}{N} \right]^k \right\} + u \quad (3.1)$$

waarbij $\hat{k}_N = \hat{k}_N(Y_1, \dots, Y_N)$ en $\hat{\sigma}_N = \hat{\sigma}_N(Y_1, \dots, Y_N)$ de meest aannemelijke schatter (MLE) van de parameters k en σ van de GPV-verdeling voorstellen; $N = N(u)$ is het aantal overschrijdingen van drempel u . Inplaats van (3.1) kunnen we ook schrijven

$$\hat{h}_n(p) = \psi(\hat{k}_N, \hat{\sigma}_N) \quad (3.2)$$

waarbij de functie ψ gegeven is door

$$\psi(k, \sigma) = \frac{\sigma}{k} \left\{ 1 - \left[\frac{-m \log(1-p)}{N} \right]^k \right\} + u \quad (3.3)$$

Hoewel het onmogelijk is om -op analoge wijze als in paragraaf 2- een benadering te geven voor $\frac{d}{du} E\hat{h}_n(p)$ (voor het geval $\frac{d}{du} k > 0$) blijkt deze uitdrukking (welke hier niet gegeven wordt) van weinig nut voor de beschrijving van de grafieken van het 10^{-4} -kwantiel, gebaseerd op het EV-model, als functie van de drempelwaarde.

Hoe is dit te verklaren? In fig. 6.3 is naast een grafiek van het geschatte 10^{-4} -kwantiel versus drempelwaarde, ook een "theoretische" kromme getekend (de stippellijn), welke bepaald is met behulp van de relatie

$$\psi(\hat{k}_2, u_2) = \psi(\hat{k}_1, u_1) + (\hat{k}_2 - \hat{k}_1) \frac{\partial \psi}{\partial k} \Big|_1 + (u_2 - u_1) \frac{\partial \psi}{\partial u} \Big|_1 \quad (3.4)$$

Deze theoretische kromme beschrijft vrij behoorlijk het verloop van het 10^{-4} -kwantiel als functie van de drempel. We concluderen dat de fluctuaties van de schattingen van de parameter k het verloop van het 10^{-4} -kwantiel in belangrijke mate bepalen. Merk overigens op dat vanaf $u=190$ de schattingen van k afnemen bij toenemende drempel (vermoedelijk vanwege de steeds grotere invloed van de grootste HW-stand op deze schattingen).

In fig. 6.4 zijn de HW-opzetten bij het 10^{-4} -kwantiel uitgezet. Ook hier vinden we dat het verloop in de schattingen goed beschreven wordt met behulp

van formule (3.4), voor drempelwaarden < 89 cm. Voor hogere drempel is dit niet meer het geval; dit is ongetwijfeld het gevolg van de tekenwisseling in de schattingen van k welke optreden bij $u=89$ en $u=92$ (zie Tabel 1).

6.3.4 De simulaties

Het Gumbelmodel

Er zijn 50 steekproeven genomen van 485 waarnemingen uit de GPV-verdeling met parameterwaarden $k=0.0787$, $u=50$ en $\sigma = 36.6$. Dit zijn de waarden gevonden bij Hoek van Holland voor alle HW-opzetten ≥ 50 cm, geselecteerd volgens het Daan-5 criterium. Voor elke steekproef is nu bij iedere drempel $u(u=50,51,\dots,99)$ het 10^{-4} -kwantiel bepaald onder aanname van het Gumbelmodel. In fig. 6.5 is het gemiddelde 10^{-4} -kwantiel van 50 steekproeven uitgezet tegen oplopende drempel. Ook zijn, benaderende betrouwbaarheidsintervallen bepaald, m.b.v. de formule $(\hat{\mu}-2\hat{\sigma}, \hat{\mu}+2\hat{\sigma})$, waarbij $\hat{\mu}$ het gemiddelde 10^{-4} -kwantiel en $\hat{\sigma}^2$ de steekproefvariantie van de 50 schattingen van de 10^{-4} -kwantiel, bij gegeven drempel u , voorstellen. We zien duidelijk een dalend verloop van de puntschattingen van het 10^{-4} -kwantiel. Verder zijn nog toegevoegd de figuren 6.6, 6.7 en 6.8. In fig. 6.6 is, nu voor 20 gesimuleerde steekproeven, het gemiddelde 10^{-4} -kwantiel gebaseerd op het Gumbelmodel uitgezet tegen de drempel, maar nu zijn de 20 gesimuleerde steekproeven getrokken uit een exponentiële verdeling ($k=0$). In overeenstemming met de theorie zien we dat het verloop van schattingen constant blijft bij toenemende drempel. Fig. 6.7 geeft het verloop van het 10^{-4} -kwantiel tegen de drempelwaarde, voor het geval de 20 gesimuleerde steekproeven getrokken zijn uit een GPV-verdeling met $k=0.0787$. Omdat de k -waarde nu negatief is blijkt het verloop stijgend te zijn. In fig. 6.8 zien we de door simulatie verkregen schattingen van het 10^{-4} -kwantiel, met daarbij (de kromme met sterren) de door Daan-5 selectie bepaalde kromme. In de figuren 6.9, 6.10 en 6.11 zijn de 10^{-4} -kwantielen te Delfzijl uitgezet voor opzetten ≥ 50 cm, voor zowel de Daan S-0 selectie als de Daan S-3 selectie, berekend volgens het Gumbelmodel. Ook hier valt duidelijk het dalende verloop van het 10^{-4} -kwantiel op.

Het EV-model

De simulaties zijn als volgt uitgevoerd. Er zijn 100 steekproeven genomen van 485 waarnemingen met de GPV-verdeling met parameterwaarden $k=0.0787$, $u=50$ en $\sigma = 36.6$. Voor elke steekproef is nu bij iedere drempel $u(u=50,51,\dots,99)$ het 10^{-4} -kwantiel bepaald onder aanname van het EV-model. In fig. 6.12 is het gemiddelde kwantiel van de 100 steekproeven uitgezet tegen oplopende drempel. Duidelijk is dat het kwantiel vrijwel constant blijft, hoewel vanaf $u=90$ er een stijgende trend lijkt te bestaan. In fig. 6.13 is dezelfde kromme uitgezet met per drempelwaarde het gemiddelde betrouwbaarheidsinterval en daar zien we dat het gemiddelde kwantiel (de middelste kromme zonder sterren) inderdaad constant is bij oplopende drempel. In fig. 6.13 is tevens het kwantiel uitgezet berekend op basis van de reële data (de kromme met sterren). Ook bij deze kromme mag men concluderen dat het stijgende verloop zeer wel meevalt wanneer we de 95% betrouwbaarheidsband in acht nemen. Het stijgende verloop kan best op toeval berusten. Om hier meer inzicht in te verkrijgen, zijn dezelfde berekeningen ook uitgevoerd voor de opzetten ≥ 50 cm te Delfzijl over het stormseizoen 31 oktober t/m 1 februari van de jaren 1900-1980. In fig. 6.14 zijn de basispeilen te Delfzijl uitgezet voor opzetten ≥ 50 cm, voor zowel de Daan S-0 selectie (getrokken kromme) als de Daan S-3 selectie (kromme met sterren). Het 10^{-4} -kwantiel bij de Daan S-0 selectie vertoont een dalend verloop. De Daan S-3 selectie is echter vrijwel constant en schommelt globaal gesproken tussen 450 en 500 cm.

In fig. 6.15 is het 10^{-4} -kwantiel van de Daan S-3 selectie alleen uitgezet. De schaal loopt van 440-520 cm. In fig. 6.16 is datzelfde kwantiel weer uitgezet maar nu met als drempelwaarde het 95% betrouwbaarheidsinterval. De geringe schommelingen in de schattingen voor het kwantiel corresponderen goed met die in de schattingen voor k (zie Tabel 2 en formule (3.4) op p.67).

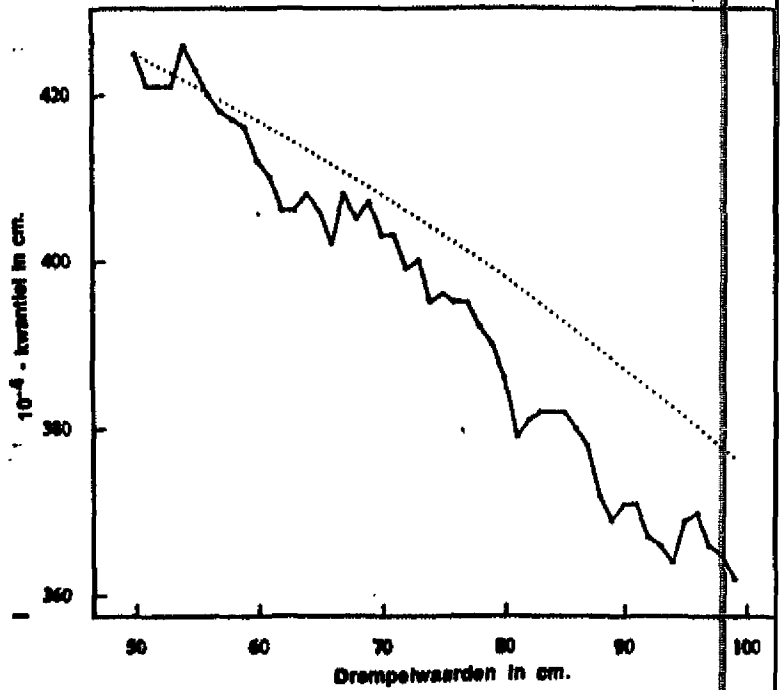
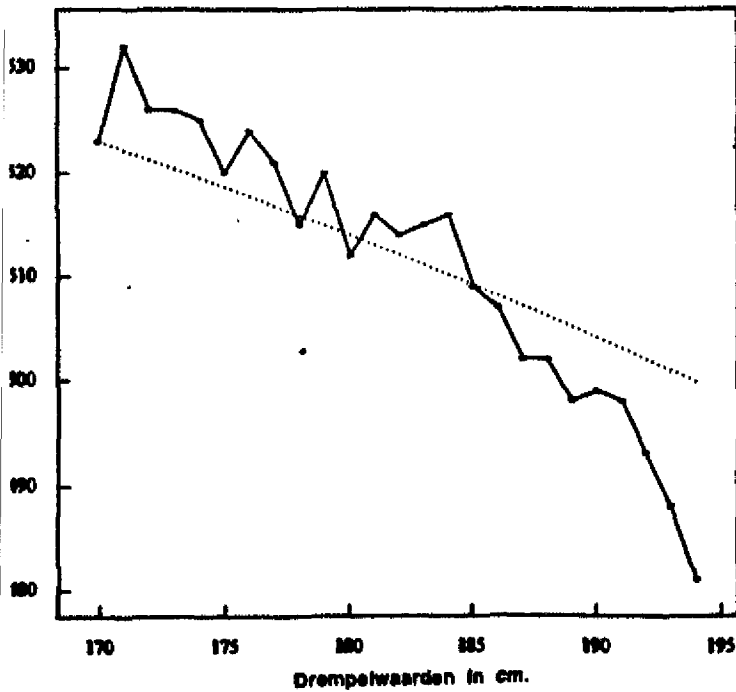
In Tabel 6.1 vinden we voor de HW-opzetten op basis van respectievelijk de Daan-5 selectie bij Hoek van Holland en de Daan-S-3 selectie te Delfzijl, schattingen voor het 10^{-4} -kwantiel en de vormparameter k . We zien dat voor Delfzijl bij oplopende drempel de schattingen van k geen dalend verloop vertonen, in tegenstelling tot wat we bij Hoek van Holland zien. Dit ondersteunt het vermoeden dat de stijgende trend bij Hoek van Holland het gevolg is van het dalen van k voor hoge drempels. Deze systematische afname van de schatting van k bij drempels $u > 90$ cm is het gevolg van de zeer hoge HW-opzet van 1 februari 1953. Dat deze HW-opzet echt extreem hoog is wordt nog eens geïllustreerd door de histogrammen 1 t/m 4 in fig. 6.17 van zowel de Daan S-0 als de Daan S-3 selectie te Delfzijl en Hoek van Holland.

6.3.5. Conclusie

In deze § hebben we laten zien dat, voor het geval van het Gumbel-model, het dalende verloop van de puntschattingen van het 10^{-4} -kwantiel voor HW's en HW-opzetten bij oplopende drempelwaarde, verklaard kan worden door aan te nemen dat de verdeling van de HW's (of HW-opzetten), welke boven een drempel u uitsteken, bij benadering een GPV-verdeling is, met vormparameter $k > 0$. Het verloop van het geschatte 10^{-4} -kwantiel als functie van de drempelwaarde in het EV-model wordt daarentegen grotendeels bepaald door de fluctuaties in de schattingen van k .

fig. 6.1 : HW's

fig. 6.2 : HW-opzetten



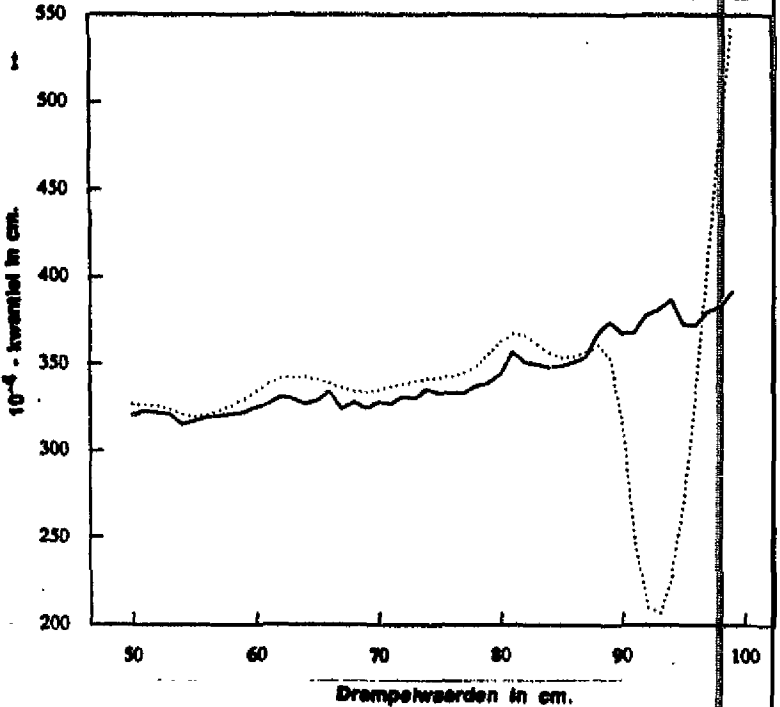
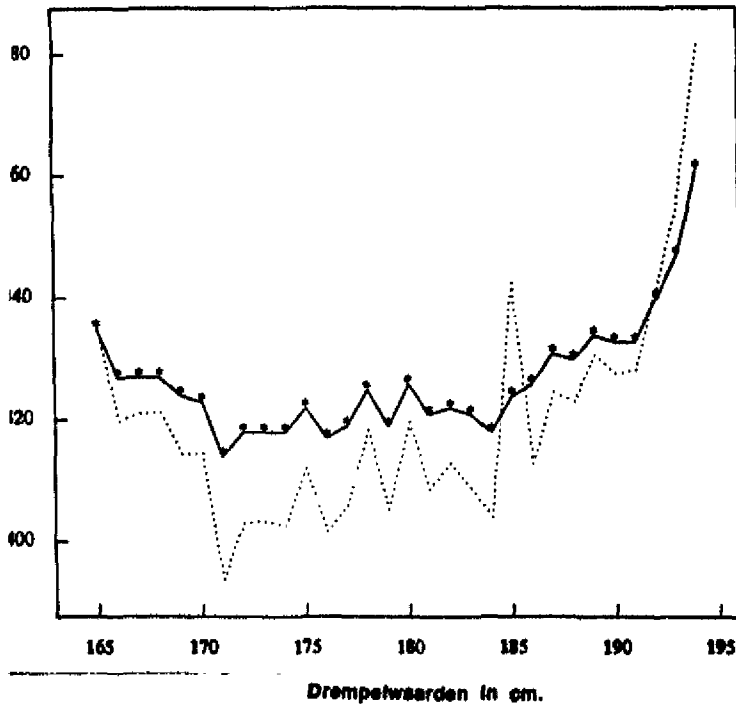
Uitleg figuren 6.1 en 6.2

Kromme met blokjes : Geschat 10^{-4} - kwantiel volgens het Gumbel - model voor HW's (Grafiek 1) resp. HW - opzetten (Grafiek 2) tegen oplopende drempelwaarde (Daan - 5 selectie te Hoek van Holland van 1900...1980 met als stormseizoen 31 oktober t/m 1 februari).

Gestippelde kromme : Verwachte schatting van het 10^{-4} - kwantiel (op basis van formule (2.9) pag.66) volgens het Gumbel - model indien de overschrijdingen (van drempel u) van de geselecteerde HW - opzetten resp. HW's GPV verdeeld (met geschatte k en σ) zijn; x_F is geschat met de grootste waarneming d.w.z. voor HW - opzetten is dit 293 cm. en voor de HW's 385 cm.

fig. 6.3 : HW's

fig. 6.4 : HW-opzetten



Uitleg figuren 6.3 en 6.4

Getrokken kromme : Idem als de kromme met blokjes in fig. 6.1 en 6.2 maar nu zijn de kwantielen bepaald volgens het FV - model.

Gestippelde kromme : Verwachte schatting 10^{-4} - kwantiel op basis van formule (3.4) pag 67.

HOEK VAN HOLLAND

0.0001-kwantielen Gumbel-verdeling van gesimuleerde waarnemingen uit de GPD-verdeling ($k = 0.0787, u = 50, \sigma = 36.6, N = 485, \# \text{sim.} = 50$)

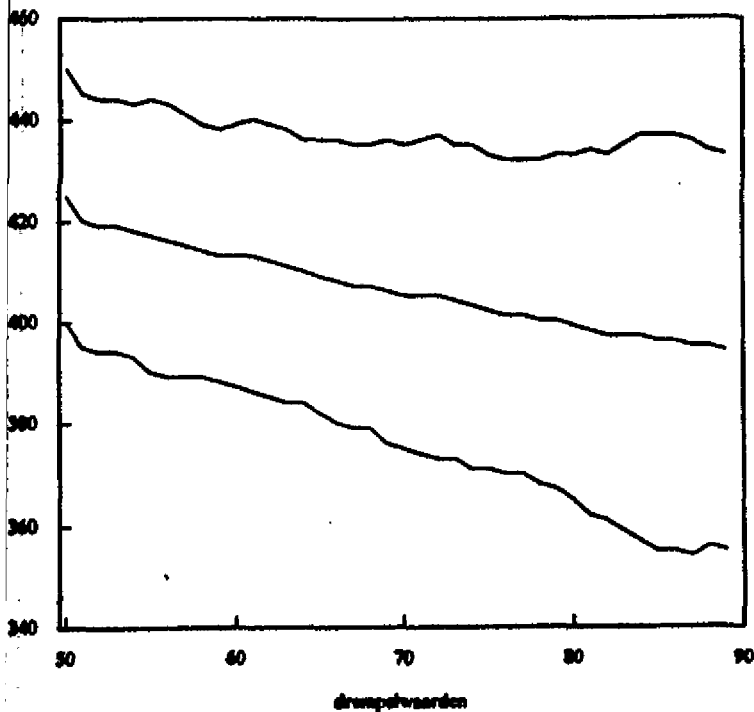


Fig. 6.5

0.0001-kwantielen Gumbel-verdeling van gesimuleerde waarnemingen uit de GPD-verdeling ($k = 0, u = 50, \sigma = 36.6, N = 485, \# \text{sim.} = 20$)

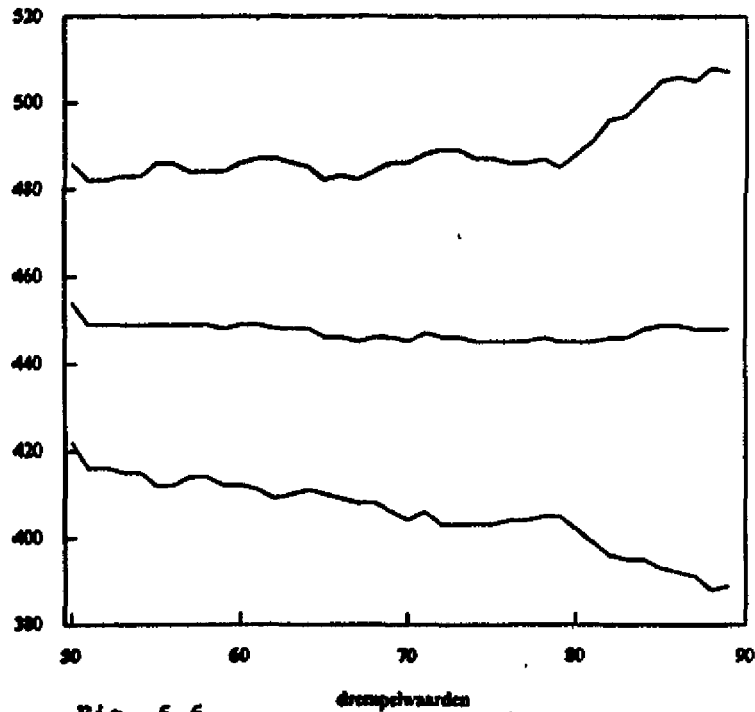


Fig. 6.6

0.0001-kwantielen Gumbel-verdeling van gesimuleerde waarnemingen uit de GPD-verdeling ($k = -0.0787, u = 50, \sigma = 36.6, N = 485, \# \text{sim.} = 20$)

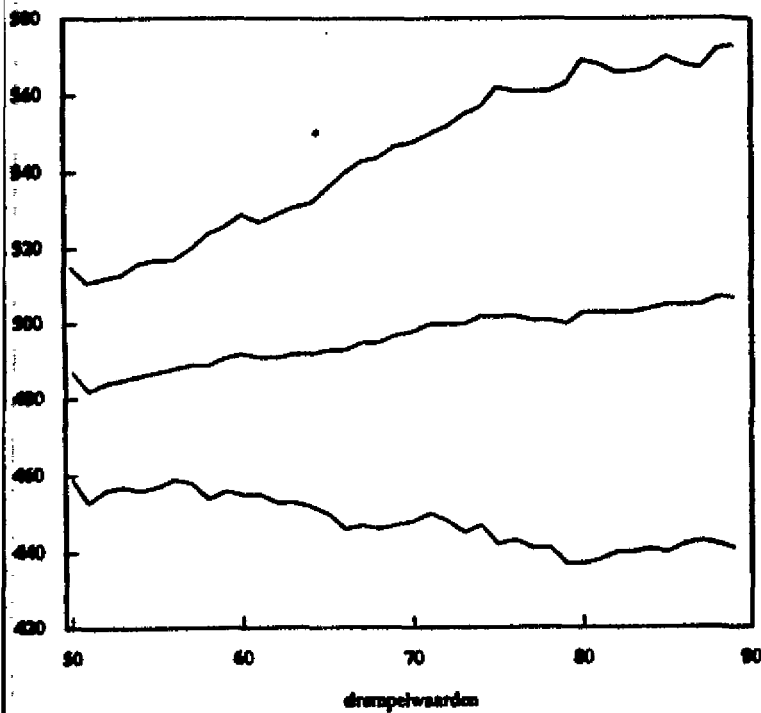


Fig. 6.7.

0.0001-kwantielen Gumbel-verdeling van gesimuleerde waarnemingen uit de GPD-verdeling ($k = 0.0787, u = 50, \sigma = 36.6, N = 485, \# \text{sim.} = 20$)

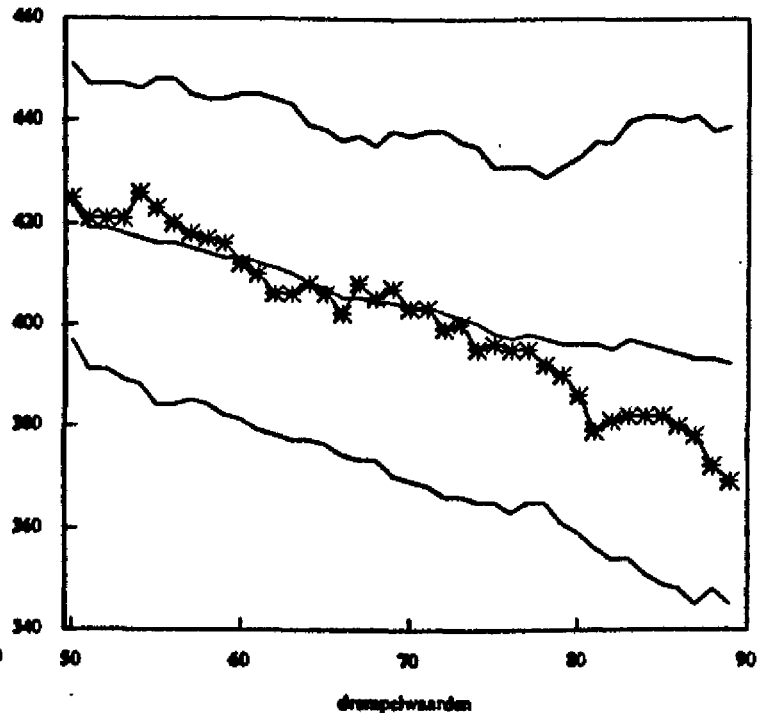


Fig. 6.8

DELFTIJL

Fig. 6.9

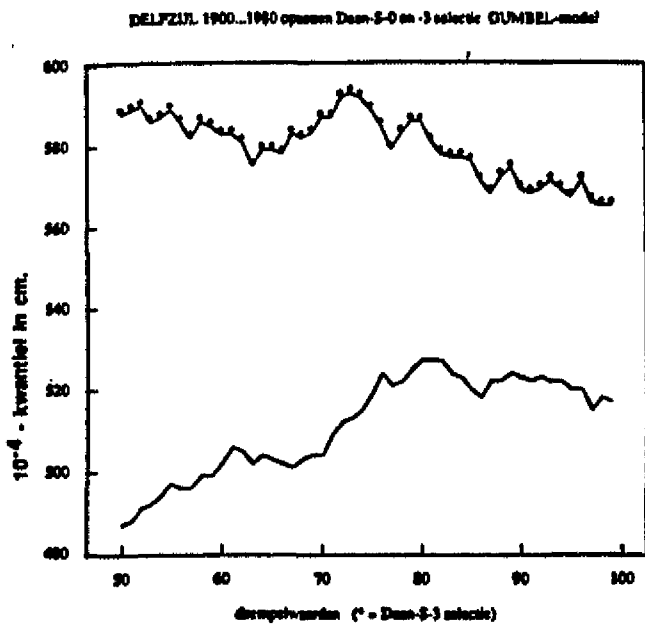


Fig. 6.10

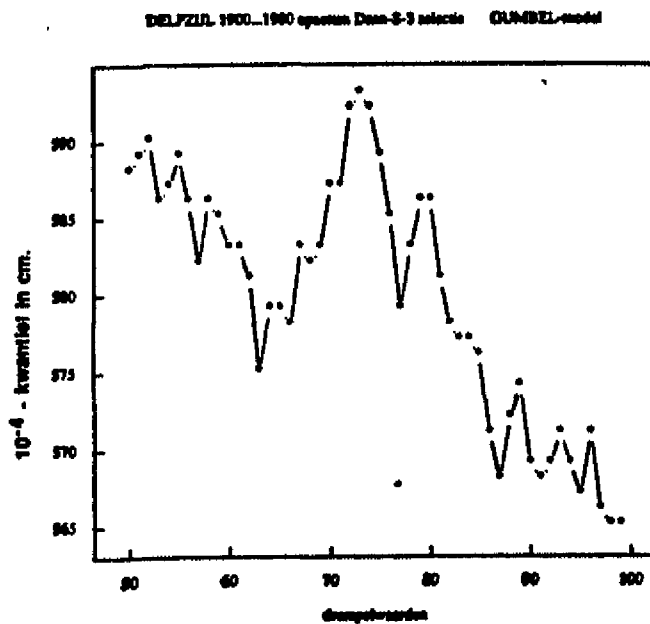
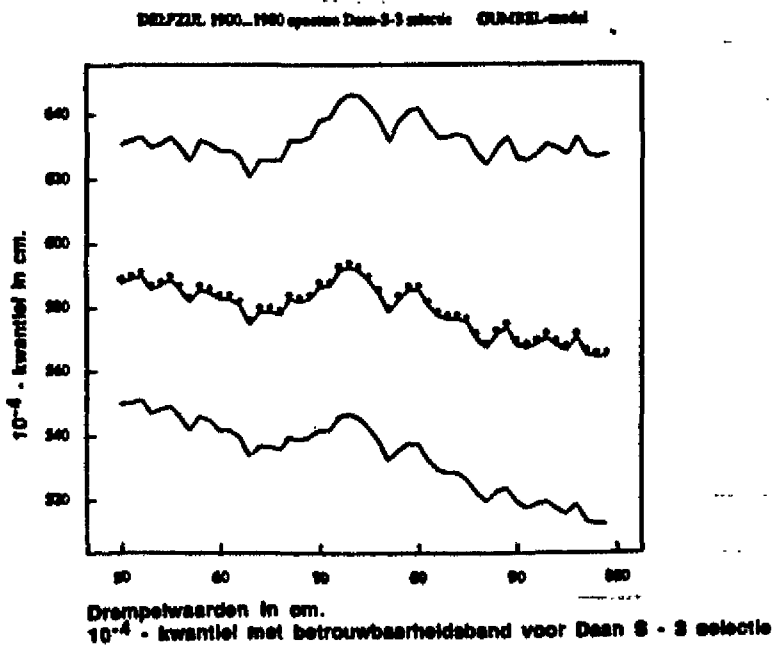


Fig. 6.11



HOEK VAN HOLLAND

0.0001-kwantielen EV-verdeling van gesimuleerde waarnemingen uit de
GPD-verdeling ($k = 0.0787$, $u = 50$, $\sigma = 36.6$, $N = 485$, #sim. = 100)

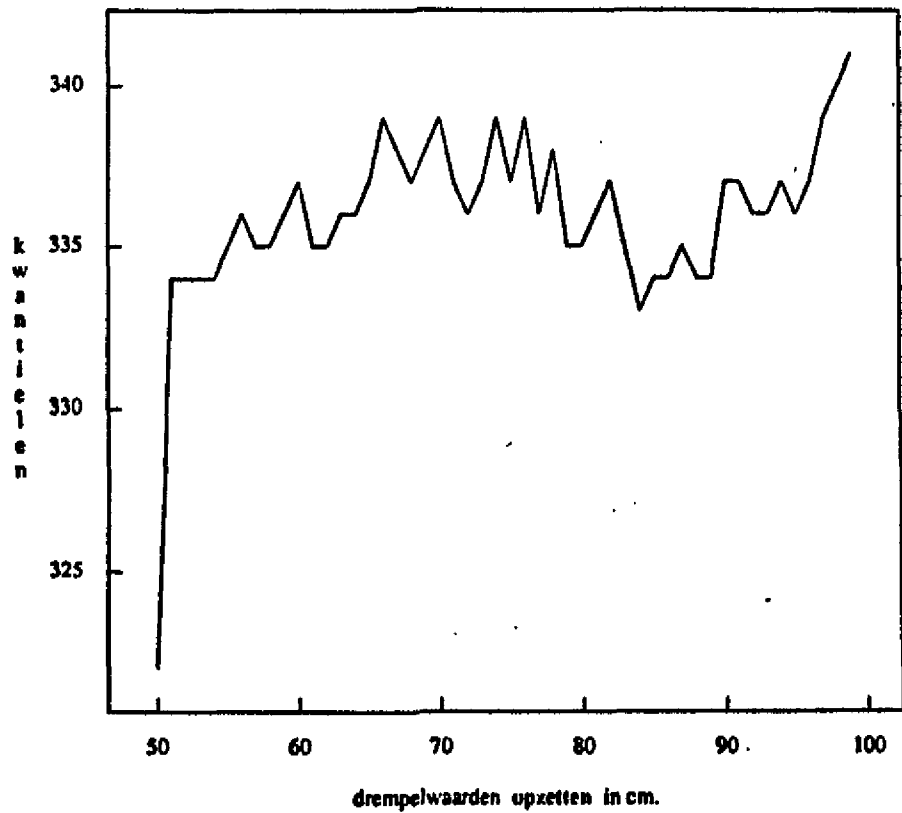


Fig. 6.12

0.0001-kwantielen EV-verdeling van gesimuleerde waarnemingen uit de
GPD-verdeling ($k = 0.0787$, $u = 50$, $\sigma = 36.6$, $N = 485$, #sim. = 100)

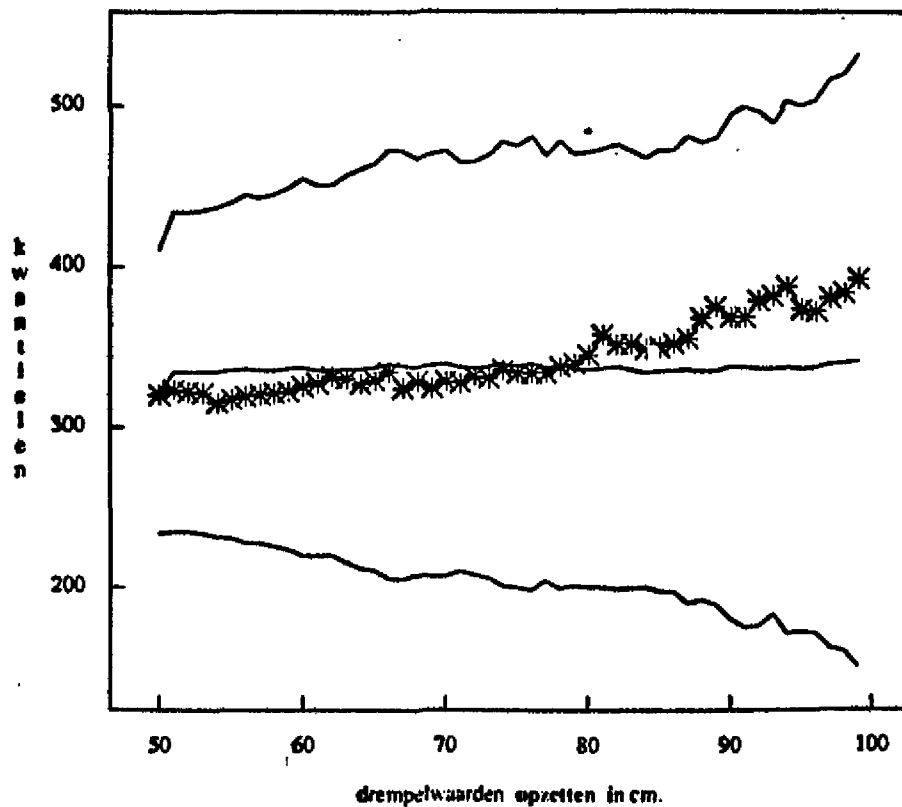


Fig. 6.13

DELFIJL

DELFIJL 1900...1980 opzetten Deen-S-0 en -3 selectie EV-model

Fig. 6.14

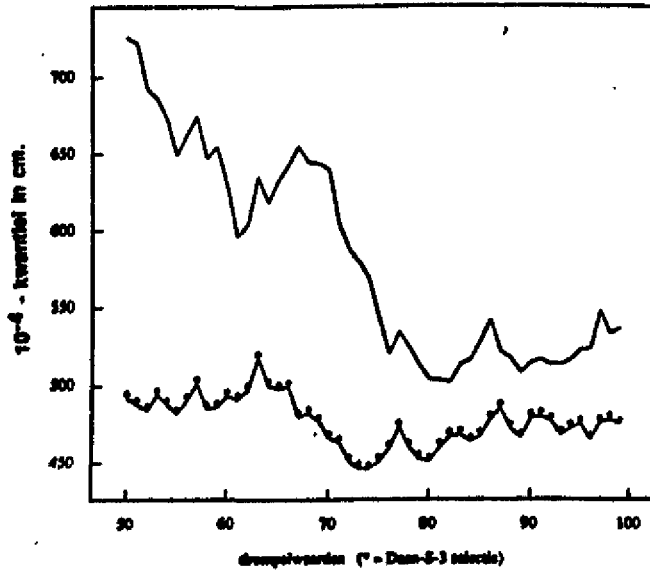


Fig. 6.15

DELFIJL 1900...1980 opzetten Deen-S-3 selectie EV-model

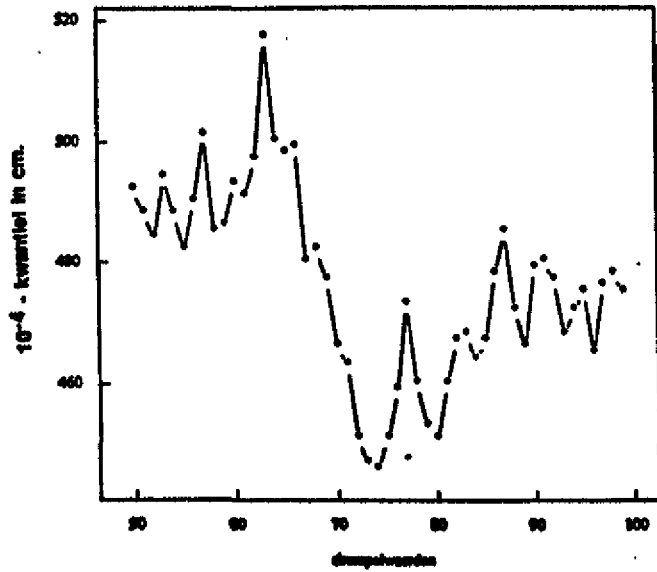
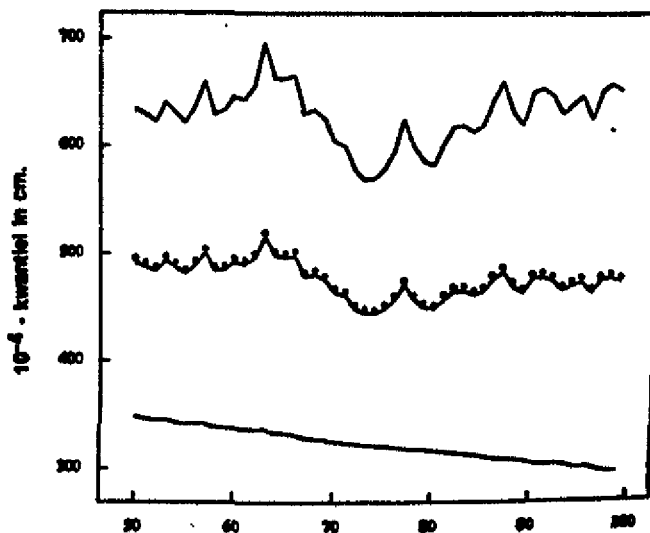


Fig. 6.16

DELFIJL 1900...1980 opzetten Deen-S-3 selectie EV-model

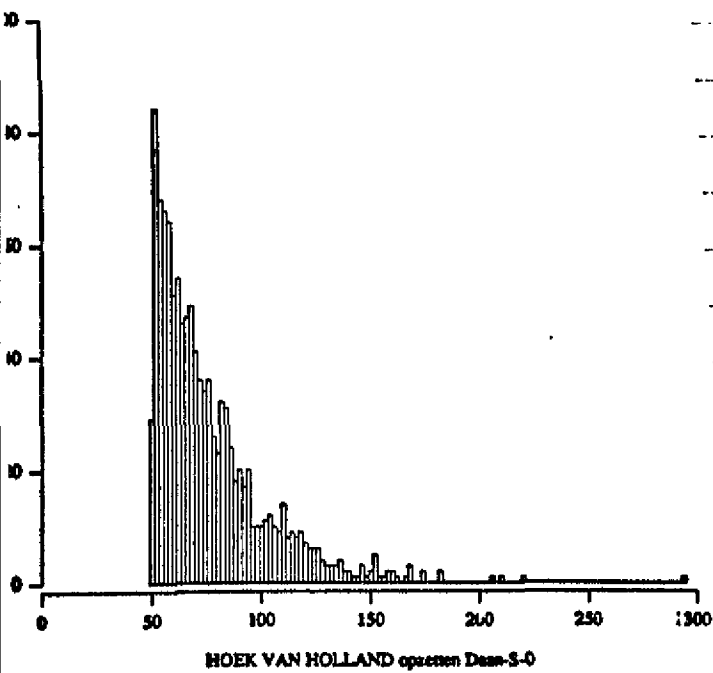


Drempelwaarden in cm.
 10^{-4} - kwantiel met betrouwbaarheidsband voor Deen S - 3 selectie

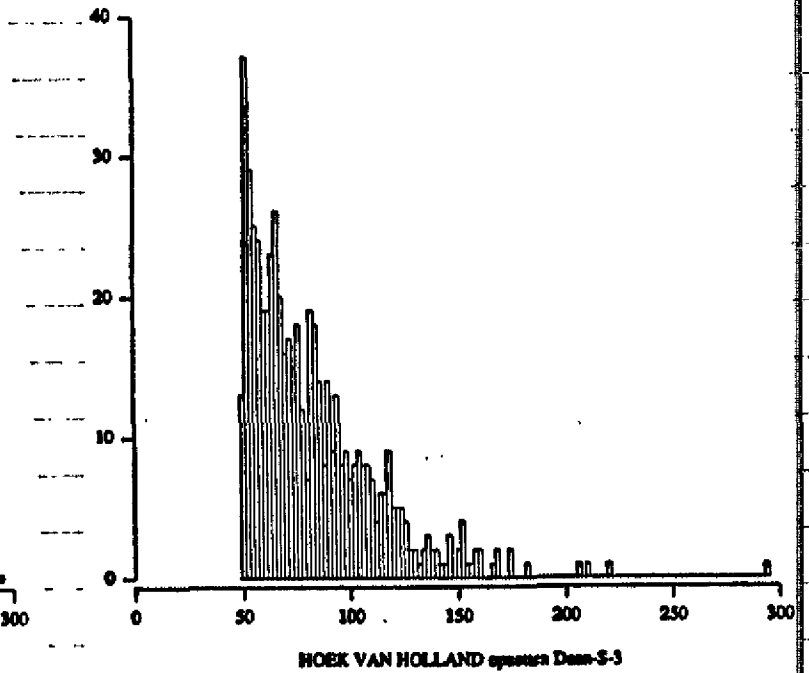
Tabel 6.1 Schattingen voor het 10^{-4} -kwantiel van de HW-opzetten (h) en van de vormparameter (k) onder het EV-model als functie van de drempel (u) [cm].
 $N(u)$ = het aantal overschrijdingen van drempel u.

| Hoek van Holland 1900...1980 Daan-5 selectie | | | | Delfzijl 1900...1980 Daan-S-3 selectie | | | |
|---|------|-----|--------|---|------|-----|-------|
| u | N(u) | h | k | u | N(u) | h | k |
| 50 | 485 | 320 | 0.079 | 50 | 690 | 492 | 0.044 |
| 51 | 476 | 323 | 0.073 | 51 | 675 | 488 | 0.046 |
| 52 | 462 | 322 | 0.075 | 52 | 660 | 484 | 0.049 |
| 53 | 449 | 321 | 0.076 | 53 | 651 | 484 | 0.042 |
| 54 | 430 | 315 | 0.087 | 54 | 636 | 488 | 0.046 |
| 55 | 421 | 317 | 0.083 | 55 | 621 | 482 | 0.051 |
| 56 | 413 | 319 | 0.079 | 56 | 612 | 490 | 0.045 |
| 57 | 403 | 320 | 0.077 | 57 | 604 | 501 | 0.037 |
| 58 | 393 | 321 | 0.075 | 58 | 586 | 485 | 0.048 |
| 59 | 383 | 322 | 0.074 | 59 | 575 | 486 | 0.047 |
| 60 | 376 | 325 | 0.068 | 60 | 566 | 493 | 0.043 |
| 61 | 367 | 327 | 0.066 | 61 | 554 | 491 | 0.044 |
| 62 | 360 | 331 | 0.059 | 62 | 545 | 497 | 0.040 |
| 63 | 349 | 330 | 0.061 | 63 | 540 | 517 | 0.027 |
| 64 | 337 | 327 | 0.065 | 64 | 524 | 500 | 0.038 |
| 65 | 329 | 329 | 0.062 | 65 | 513 | 498 | 0.039 |
| 66 | 323 | 334 | 0.055 | 66 | 503 | 499 | 0.038 |
| 67 | 307 | 324 | 0.070 | 67 | 487 | 480 | 0.052 |
| 68 | 301 | 328 | 0.064 | 68 | 478 | 482 | 0.050 |
| 69 | 290 | 324 | 0.070 | 69 | 467 | 477 | 0.054 |
| 70 | 285 | 328 | 0.053 | 70 | 454 | 466 | 0.063 |
| 71 | 276 | 327 | 0.064 | 71 | 444 | 463 | 0.066 |
| 72 | 271 | 331 | 0.057 | 72 | 431 | 451 | 0.076 |
| 73 | 262 | 330 | 0.060 | 73 | 421 | 447 | 0.080 |
| 74 | 258 | 335 | 0.051 | 74 | 413 | 446 | 0.081 |
| 75 | 249 | 333 | 0.055 | 75 | 407 | 451 | 0.076 |
| 76 | 242 | 333 | 0.054 | 76 | 402 | 459 | 0.069 |
| 77 | 235 | 333 | 0.053 | 77 | 399 | 473 | 0.056 |
| 78 | 230 | 337 | 0.047 | 78 | 387 | 460 | 0.067 |
| 79 | 224 | 339 | 0.045 | 79 | 377 | 453 | 0.074 |
| 80 | 220 | 344 | 0.035 | 80 | 369 | 451 | 0.076 |
| 81 | 218 | 357 | 0.018 | 81 | 365 | 460 | 0.067 |
| 82 | 209 | 351 | 0.026 | 82 | 360 | 467 | 0.061 |
| 83 | 202 | 350 | 0.027 | 83 | 353 | 468 | 0.060 |
| 84 | 195 | 348 | 0.029 | 84 | 345 | 464 | 0.063 |
| 85 | 189 | 349 | 0.029 | 85 | 339 | 467 | 0.061 |
| 86 | 184 | 351 | 0.025 | 86 | 335 | 478 | 0.052 |
| 87 | 179 | 354 | 0.021 | 87 | 330 | 485 | 0.045 |
| 88 | 177 | 367 | 0.004 | 88 | 320 | 472 | 0.057 |
| 89 | 173 | 374 | -0.006 | 89 | 312 | 466 | 0.062 |
| 90 | 166 | 368 | 0.001 | 90 | 309 | 479 | 0.050 |
| 91 | 160 | 368 | 0.002 | 91 | 303 | 480 | 0.049 |
| 92 | 157 | 378 | -0.012 | 92 | 296 | 477 | 0.052 |
| 93 | 152 | 381 | -0.014 | 93 | 288 | 468 | 0.060 |
| 94 | 148 | 387 | -0.022 | 94 | 283 | 472 | 0.057 |
| 95 | 140 | 373 | -0.005 | 95 | 278 | 475 | 0.054 |
| 96 | 135 | 372 | -0.004 | 96 | 270 | 465 | 0.063 |
| 97 | 132 | 380 | -0.014 | 97 | 267 | 476 | 0.053 |
| 98 | 128 | 383 | -0.018 | 98 | 262 | 478 | 0.051 |
| 99 | 125 | 392 | -0.029 | 99 | 256 | 475 | 0.053 |

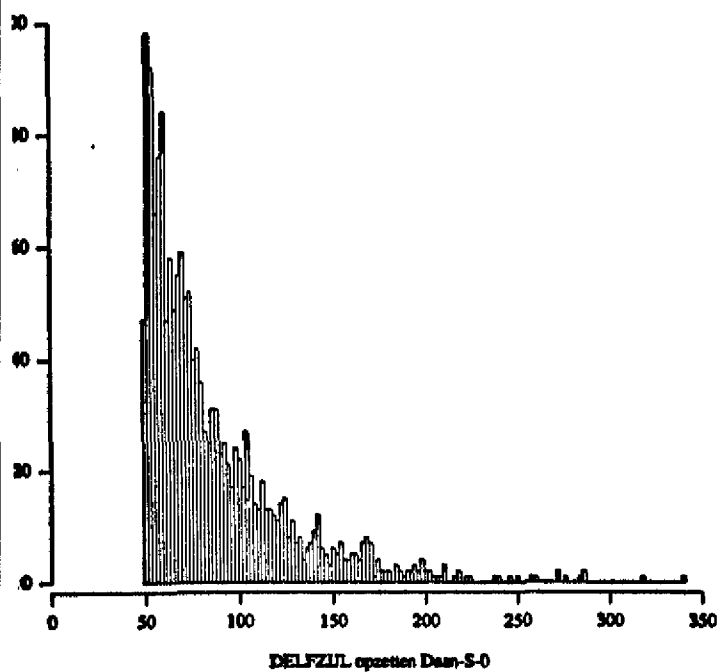
Histogram 1



Histogram 2



Histogram 3



Histogram 4

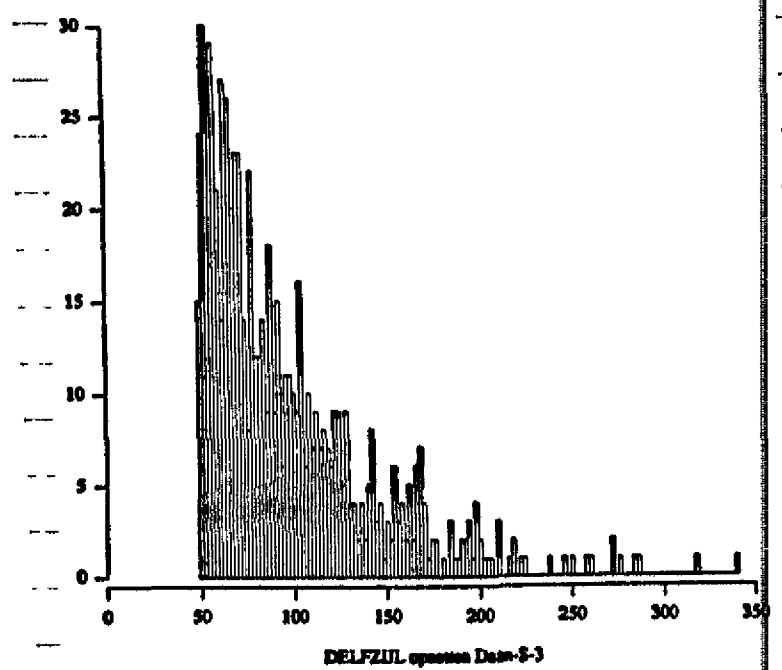


fig. 6.17 Histogrammen HW-opzetten bij selectie Daan-S-0 en Daan-S-3

7. CONCLUSIES

1. Om een reeks onafhankelijke waarden voor HW-opzetten te verkrijgen wordt toepassing van de Daan S-i selectiemethode aanbevolen (zie par. 2.5).
2. Voor het hierboven genoemde getal i is gekozen de waarde $i=4$ (zie par. 5.3.3).
3. Het gebruik van de hoogwateropzet als variabele voor het frequentie-onderzoek verdient de voorkeur boven het gebruik van de hoogwaterstand (zie par. 3.2.4).
4. Voor de onderzochte stations zijn de volgende waarnemingsperioden bruikbaar:

| | |
|------------------|--------------------------|
| Vlissingen | 1881.07.01....1985.12.31 |
| Hoek van Holland | 1887.08.01....1985.12.31 |
| Den Helder | 1932.06.01....1985.12.13 |
| Harlingen | 1932.06.01....1985.12.31 |
| Delfzijl | 1881.03.01....1985.12.31 |

 (zie par. 4.1).
5. Als lang stormseizoen is voor alle stations gekozen de periode 1 oktober...15 maart (zie par. 4.2); als kort stormseizoen, gebruikt voor de bepaling van de selectiemethode, de periode november-december-januari.
6. Een minimumwaarde voor de drempel is vastgesteld op 30 cm. De uiteindelijk toe te passen drempelwaarde volgt uit het gedrag van de kwantiel-schattingen volgens de verschillende toe te passen methoden als functie van de drempel. Deze moet nog nader worden vastgesteld.
7. Er dient nog onderzoek te worden verricht naar de wijze van samenvoeging van de frequentieverdelingen van de ongestoorde (=astronomische) standen en die van de opzetten tot een frequentiekromme van de hoogwaterstanden (zie par. 3.3).

8. LITERATUUR

1. J.M. Chambers, W.S. Cleveland, B. Kleiner en P.A. Tukey (1983): Graphical Methods for Data Analysis. Wadsworth.
2. D. van Dantzig, J. Hemelrijk (1960): Extrapolatie van de overschrijdingslijn van de hoogwaterstanden te Hoek van Holland met behulp van geselecteerde stormen. Rapport Deltacommissie, deel 3, bijdrage II.1.
3. J. Doekes (1968): Zijn HW-opzet en astronomisch hoogwater onafhankelijk? Rijkswaterstaat, Dienst Getijdewateren, nota GWIO-86.0113.
4. 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) pp. 51-60.
5. J. Hájek (1969): A Course in Nonparametric Statistics. Holden Day, San Francisco.
6. S. Kruizinga (1978): Objectieve classificatie van dagelijkse 500 mbarpatronen. KNMI, de Bilt, Wetenschappelijk Rapport WR-78-8.
7. M.R. Leadbetter (1983): Extremes and Local Dependence in Stationaire Sequences, Z. Wahrscheinlichkeitstheorie verw. Gebiete 65, 291-306.
8. M.R. Leadbetter, G. Lindgren, H. Rootzen (1983): Extremes and Related Properties of Random Sequences and Processes, Springer Series in Statistics.
9. J. van Malde (1986): Relatieve rijzing van gemiddelde zeeniveaus. In: verslag symposium "Zee-in-zicht", (TH-Delft), uitgave Gezelschap "Practische Studie".
10. J. van Malde (1987): Enkele opmerkingen over autocorrelatie bij HW-opzetten. Notitie 19/20 augustus 1987.
11. McGill, Tukey en Larsen (1978): Variations of Box Plots. The American Statistician, Vol. 32, No. 1, pp 12-16
12. Rijkswaterstaat en KNMI (1961): Verslag over de stormvloed van 1953. Staatsdrukkerij- en uitgeverijbedrijf.
13. Rijkswaterstaat, Dienst Getijdewateren (1986): Onderzoek Basispeilen. Tussentijds verslag juni 1986, nota GWIO-86.008.
14. J.G. de Ronde (1985): Wisselwerking tussen opzet en verticaal getij. Rijkswaterstaat, Dienst Getijdewateren, nota GWIO-85.003.
15. A. van Urk en J.G. de Ronde (1982): Getijtafels voor Nederland vanaf 1980. Rijkswaterstaat-serie nr. 37.
16. P.J. Diggle (1983), Statistical Analysis of spatial point patterns, Academic Press.

BIJLAGE

STAAT I: KENMERKENDE WATERSTANDEN, SLOTGEMIDDELDEN 1981.0

| Naam station | Grens peil G | Gemiddelde hoogwaterst. | | | G-HWM | HWS-HWM | HWM-HWD | HWM in % van G |
|---------------------|--------------------|----------------------------|-----|-----|--------|---------|---------|----------------------|
| | | HWS | HWM | HWD | | | | |
| Vlissingen | 340 | 239 | 202 | 153 | 140 cm | 37 cm | 49 cm | 59,4% |
| Hoek van Holland | 255 | 121 | 105 | 82 | 150 cm | 16 cm | 23 cm | 41,2% |
| IJmuiden | 245 | 106 | 91 | 70 | 155 cm | 15 cm | 21 cm | 37,1% |
| Den Helder | 220 | 62 | 54 | 44 | 165 cm | 8 cm | 10 cm | 24,5% |
| Harlingen | 290 | 103 | 91 | 74 | 200 cm | 12 cm | 17 cm | 31,4% |
| Delfzijl | 340 | 144 | 131 | 112 | 210 cm | 13 cm | 19 cm | 38,5% |
| in cm t.o.v. N.A.P. | | | | | | | | |

HWS = gem. HW springtij
HWM = gem. HW gemidd. tij
HWD = gem. HW doodtij

STAAT II: RIJZING ZEENIVEAUS - ENIGE GEGEVENS (in cm/eeuw)

| Station | Periode | Z | HWS | HWM | HWD | LWS | LWM | LWS | TVS | TVM | TVD |
|------------|-------------|----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Vlissingen | 1900...1984 | 22 | 39 | 33 | 26 | 19 | 19 | 19 | 20 | 14 | 7 |
| Den Helder | 1933...1984 | 19 | 19 | 19 | 19 | 10 | 10 | 10 | 10 | 10 | 10 |
| Harlingen | 1933...1984 | veran- derd | | 31 | | | 11 | | | 20 | |
| Delfzijl | 1900...1960 | 15 | | 17 | | | 16 | | | 1 | |

De rijzingen zijn ontleend aan lit. 9 en zijn berekend voor de jaarwaarden van de gemiddelde zeespiegel Z, de gemiddelde hoog- en laagwaterstand en het gemiddelde tijverschil TV met onderscheid naar springtij, gemiddeld tij en doodtij.

STAAT III: ASTR. DAGELIJKSE ONGELIJKHEID ("DO") voor HW,
SLOTGEMIDDELDEN 1981.0

| Station | DO | HW | 1981.0 |
|------------------|----|-------|--------|
| Vlissingen | | 11 cm | |
| Hoek van Holland | | 15 cm | |
| IJmuiden | | 13 cm | |
| Den Helder | | 13 cm | |
| Harlingen | | 26 cm | |
| Delfzijl | | 19 cm | |