

MATHEMATISCH CENTRUM

2e BOERHAAVESTRAAT 49

AMSTERDAM

STATISTISCHE AFDELING

Leiding: Prof. Dr D. van Dantzig

Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

1954-11(2)

Statistische analyse van waterstanden,
methoden en resultaten

III. Splitsing van het jaar in groepen van maanden

1954

1. Notatie ¹⁾

In Hoek van Holland blijken de hogere H.W.standen, cumulatief uitgezet op halflogarithmisch papier, voor lange perioden ongeveer op een rechte lijn te liggen ²⁾. Stelt $X(h)$ het gemiddelde aantal overschrijdingen van de hoogte h per jaar voor, dan heeft een dergelijke rechte lijn de vorm

$$(1) \quad X(h) = ck^{-ch}, \quad (k > 0)$$

waarin één der parameters c , k en c' nog willekeurig gekozen kan worden.

Daar het nuttig is, de parameters in een dergelijke vergelijking een anschouwelijke betekenis te geven, stelt Prof. VAN DANTZIG voor de volgende parameters in te voeren:

1) a_k = het hoogteverschil, waardoor het gemiddelde aantal overschrijdingen k maal zo klein wordt. Deze grootte wordt dus gedefinieerd door

$$(2) \quad X(h+a_k) = \frac{1}{k} X(h)$$

voor iedere h en $h+a_k$ in het lineaire gebied van de lijn.

Uit (1) en (2) volgt

$$(3) \quad c' = \frac{1}{a_k},$$

dus

$$(4) \quad X(h) = ck^{-\frac{h}{a_k}}.$$

2) h_k = hoogte, die gemiddeld $\frac{N}{k}$ maal per jaar wordt overschreden, waarbij N het aantal H.W.'s per jaar voorstelt. Deze grootte wordt dus gedefinieerd door

$$(5) \quad X(h_k) = \frac{N}{k}$$

en uit (4) en (5) volgt

$$(6) \quad c = N \cdot k^{\frac{h_k}{a_k} - 1}.$$

Vullen wij (6) in (4) in, dan verkrijgen wij

$$(7) \quad X(h) = Nk^{-\frac{h-h_k}{a_k} - 1} \quad (k > 0).$$

In deze formule, waarin k dus nog willekeurig gekozen kan worden, hebben nu alle parameters een anschouwelijke betekenis

1) Dit rapport is het resultaat van gemeenschappelijk onderzoek van een aantal leden van de Statistische Afdeling van het Mathematisch Centrum.

2) Vgl. Ir P.J. Wemelsfelder, Wetmatigheden in het optreden van stormvloeden, "De Ingenieur", 3 Maart 1939.

en bij iedere gewenste waarde van k kunnen zij gemakkelijk uit de grafiek van de lijn bepaald worden.

Werkt men niet met het gemiddelde aantal overschrijdingen van h per jaar, maar met de kans op overschrijding bij één waarneming, dan vervalt de factor N in het rechterlid.

Het grondtal k kan nu in (7) nog willekeurig gekozen worden. Voor twee verschillende grondtallen k en g geldt nu volgens (7):

$$(8) \quad k^{-\frac{h-h_k}{a_k}-1} = g^{-\frac{h-h_g}{a_g}-1}$$

voor iedere h in het lineariteitsgebied.

Vullen wij hierin $h = h_k - a_k$ in, dan volgt gemakkelijk

$$(9) \quad h_k - h_g = a_k - a_g.$$

Invullen van $h = h_k$ resp. $h = h_g$ in (8) geeft na enige herleiding de beide ook direct uit elkaar volgende betrekkingen

$$(10) \quad a_k = a_g^g \log k \\ a_g = a_k^k \log g.$$

Drie in het bijzonder voor de hand liggende waarden voor het grondtal k zijn: e (het grondtal der Neperse logaritmen), 2 en 10. Voor deze drie grondtallen kunnen wij de parameter a_k de namen "kans-nepererings-verhoging", "kanshalveringsverhoging" resp. "kansdecimeringsverhoging" geven, waarbij "verhoging" in de zin van "hoogteverschil" bedoeld is.

De parameter h_k kan de "neperinehoogte", "mediaanhoogte" resp. "decilehoogte" genoemd worden. Deze verschillende parameters zullen wij, al naar het uitkomt, naast elkaar gebruiken en opgeven. De door (7) aangegeven lijnen zullen wij voortaan H.W.-overschrijdingslijnen noemen, ook als de onderstelling van lineariteit op halflogaritmisch papier niet gemaakt wordt.

2. Groepering der maanden van een jaar.

Het is bekend, dat hoge H.W.'s in de zomer minder voorkomen dan in de winter. Het is daarom wenselijk, het jaar te splitsen in verschillende groepen van maanden, binnen ieder waarvan de verdeling der H.W.'s homogener is dan over het gehele jaar. Bij bepaalde statistische onderzoeken kan men dan volstaan met bewerking van de waarnemingen van de gevaarlijkste periode. Dit heeft twee voordelen: ten eerste geeft het een aanzienlijke werkbesparing en ten tweede worden de uitkomsten betrouwbaarder, omdat men minder inhomogeniteit in het waarnemingsmateriaal mag verwachten.

De splitsing in groepen van maanden kan wellicht het beste geschieden door voor een bepaalde periode na te gaan hoeveel keren de hoogste stand van een jaar in de verschillende maanden viel. Immers de hoogste stand in een jaar zal bepalen of er in dat jaar gevaar voor overstroming bestaat of niet. In tabel 1 vindt men deze aantallen vermeld voor de periode 1901 tot en met 1951. Wij hebben ons daarbij gebaseerd op de gegevens van 3 plaatsen: Hoek van Holland, Willemstad en Vlissingen. Wegens de afhankelijkheid van de waarnemingen voor verschillende plaatsen is het overbodig nog meer plaatsen in deze beschouwingen te betrekken.

Tabel 1
Aantallen hoogste standen per jaar in verschillende maanden
(1901-1951)

	Jan.	Feb.	Mrt	Apr.	Mei	Juni	Juli	Aug.	Sep.	Oct.	Nov.	Dec.
H.v.H.	10	4	3	1					2	4	14	13
Willemstad	10 $\frac{1}{2}$ ³⁾	4	4	1			1		2	4 $\frac{1}{2}$ ³⁾	13	11
Vlissingen	11	2	6	3			1		4	3	15	6

Volgens de cijfers van deze tabel kan men drie duidelijk verschillende en ieder op zich zelf min of meer homogene periodes in het jaar onderscheiden:

- 1e De winterperiode: November, December en Januari,
- 2e Lente en herfst: September, October, Februari en Maart,
- 3e Zomer: April tot en met Augustus ⁴⁾.

Een strenge scheiding tussen de perioden valt natuurlijk niet te trekken, maar niettemin mag men van een scheiding in deze drie perioden een goed resultaat verwachten. In de volgende paragraaf wordt nagegaan, welk verschil er bestaat tussen de H.W.-overschrijdingslijn te H.v.H. voor de winter- en de lente- en herfstperiode. Men ziet daar duidelijk, dat bij dergelijke onderzoekingen het gebruik van de aan de winterperiode ontleende gegevens

- 3) In 1935 werd een hoogste stand in Januari en in October waargenomen.
- 4) De stormvloedcommissie kwam in 1943 eveneens tot de conclusie, dat November, December en Januari de drie gevaarlijkste maanden zijn (zie Nota "Beschouwingen en berekeningen over de benedenrivieren", Hoofdstuk V, par. 25).

vens tot andere resultaten voert, dan wanneer men de lente- en herfstperiode erbij neemt. Daar het uiteraard van belang is, juist de gevaarlijkste perioden van het jaar in het onderzoek te betrekken, zullen vele onderzoeken tot deze winterperiode beperkt kunnen blijven. Dat daarbij verschillende stormen (zoals die van 1953!) buiten het waarnemingsmateriaal blijven, heeft weinig invloed. Indien men nl. één zeer hoog punt aan een H.W.-overschrijdingslijn toevoegt, verandert deze vrijwel niet van ligging. Het enige gevolg is, dat de punten "in de staart" slechts bij de lijn aansluiten dan zonder dat hoge punt het geval zou zijn. Bij onderzoeken, waarin niet alle H.W.'s, maar speciaal de hoogste standen per jaar betrokken worden (of enkele zeer hoge standen van ieder jaar), is beperking tot de 3 wintermaanden uiteraard niet gewenst.

3. Het verschil tussen de winterperiode en de lente- en herfstperiode.

De H.W.-overschrijdingslijn voor Hoek van Holland voor de jaren 1888-1948 is in figuur 1 geschetst op halflogarithmisch papier voor:

- I. de wintermaanden: November, December en Januari,
- II. de lente- en herfstperiode: Februari, Maart, September en October,
- III. de beide perioden tezamen.

De waarnemingen (die in de figuur zijn weggelaten) liggen voor niet te kleine h bij benadering op rechte lijnen⁵⁾, die op het oog door deze waarnemingspunten getrokken zijn. Deze waarnemingspunten strekten zich uit over de getrokken gedeelten van de lijnen, terwijl de gestippelde gedeelten daarvan de verlengden zijn. De constanten van vergelijking (7) voor $k = 2$ resp. 10 vindt men in tabel 2.

Tabel 2

Mediaan- en decilehoogte en kanshalverings- resp. decimeringsverhoging voor de winterperiode, de lente-herfstperiode en beide perioden tezamen (1888-1948)

	h_2	a_2	h_{10}	a_{10}
I winter	9	18,5	135	6
II lente-herfst	8	15,5	124,5	50,5
III tezamen		17	129,5	5,5

5) Vgl. ook par. 4.

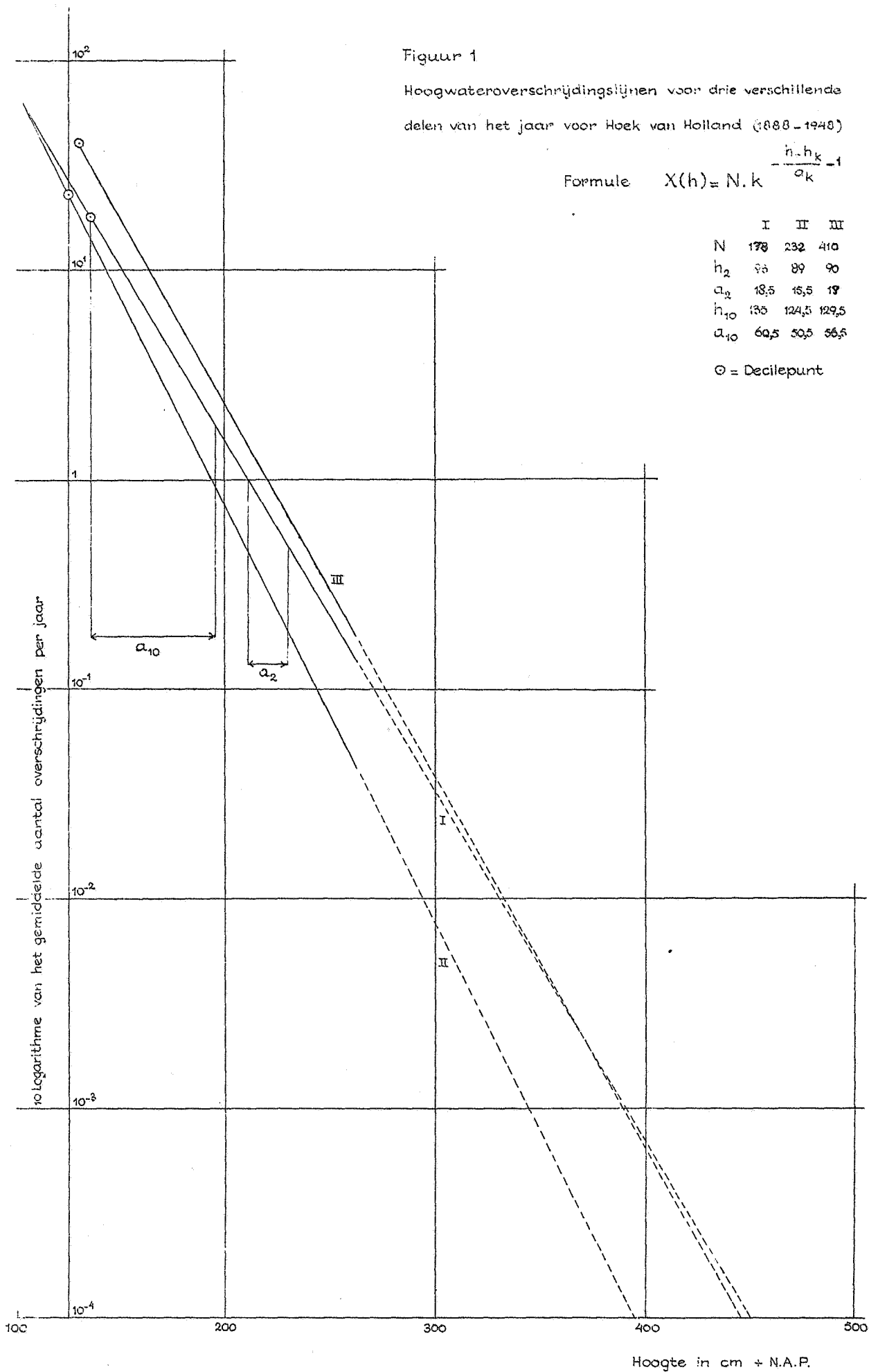
Figuur 1

Hoogwateroverschrijdingslijnen voor drie verschillende delen van het jaar voor Hoek van Holland (1888-1948)

$$\text{Formule } X(h) = N \cdot k \frac{h - h_k - 1}{a_k}$$

	I	II	III
N	178	232	410
h_2	95	89	90
a_2	18,5	16,5	19
h_{10}	135	124,5	129,5
a_{10}	60,5	50,5	56,5

○ = Decilepunt



De lijnen buigen niet ver boven het decilepunt om naar een vrijwel horizontale stand en het mediaanpunt ligt niet meer in het rechtlijnige gebied. Het is daarom in figuur 1 niet aangegeven.

Wij zien uit deze figuur, dat er een duidelijk verschil is tussen de perioden I en II, maar dat I en III weinig verschillen.

Dit wijst erop, dat het er niet veel toe doet hoe lang men de winterperiode precies neemt. Voor voorbereidende onderzoeken kan men zich derhalve bedienen van de drie maanden November, December en Januari. Na voltooiing van het voorbereidende onderzoek zal men, ter berekening van een uiteindelijke dijkhoogte, deze kwestie nader moeten beschouwen. Hieraan worden in de volgende paragraaf enkele voorbereidende opmerkingen gewijd.

4. Over rechtlijnigheid en extrapolatie

Aan figuur 1 zijn twee opmerkelijke eigenschappen waar te nemen. In de eerste plaats zou men niet verwachten, dat alle drie daarin voorkomende lijnen recht zouden kunnen zijn en strict genomen is dat ook niet zo. Immers de som van twee exponentiële functies is niet weer een exponentiële functie.

Nemen wij aan, dat de perioden I en II ieder op halflogarithmisch papier een rechte lijn geven, dan zijn de vergelijkingen daarvan blijkens (7) en de tabel bij figuur 1, met $k = 10$:

$$(8) \quad X_I(h) = 18,1 \cdot 10^{-\frac{h-135}{62}}$$

en

$$(9) \quad X_{II}(h) = 23,6 \cdot 10^{-\frac{h-124,5}{50,5}},$$

zodat wij voor periode III krijgen:

$$(10) \quad X_{III}(h) = X_I(h) + X_{II}(h) = 18,1 \cdot 10^{-\frac{h-135}{62}} + 23,6 \cdot 10^{-\frac{h-124,5}{50,5}}$$

De vergelijkingen (8) en (9) geven op logarithmisch papier een rechte lijn, maar (10) doet dit niet. Berekent men echter de kromtestraal van (10) op logarithmisch papier, dus de kromtestraal van de functie

$$(11) \quad y = {}^{10}\log \left\{ 18,1 \cdot 10^{-\frac{h-135}{62}} + 23,6 \cdot 10^{-\frac{h-124,5}{50,5}} \right\},$$

dan blijkt deze voor iedere h op de schaal van figuur 1 groter dan 87,5 m te zijn, zodat het geen verbazing wekt, dat de geringe afwijking van rechtlijnigheid niet uit de waarnemingen te zien is. Men kan er zelfs geenszins zeker van zijn, dat de rechtlijnigheid

niet beter is voor periode III dan voor I en II apart, daar men uit de waarnemingen hierover geen uitsluitend verkrijgt, terwijl men de bovenstaande redenering natuurlijk ook in andere volgorde kan geven, b.v. uitgaande van een rechtlijnig verband in periode III en niet in I en II.

Een tweede merkwaardigheid staat in verband met het probleem der extrapolatie. Het ligt voor de hand de gevonden lijnen rechtlijnig te extrapoleren, om de kans op overschrijding van grote waarden van h te schatten, een kans waar men geen directe experimentele gegevens over heeft. Afgezien van het speculatieve karakter van een dergelijke extrapolatie, waarop elders⁶⁾ uitvoerig is ingegaan, kunnen wij opmerken, dat deze extrapolaties, toegepast op periode I en III met elkaar in strijd zijn. Immers, doordat de kansdecimeringsverhoging van I groter is dan van III snijden de verlengden van deze twee lijnen elkaar, zodat bij b.v. $h = 4 \text{ m} + \text{N.A.P.}$ het verlengde van III lager ligt dan dat van I. Dit zou dus betekenen, dat er in de drie maanden November, December en Januari gemiddeld meer overschrijdingen per jaar zouden plaatsvinden dan als men daar nog de 4 maanden September, October, Februari en Maart aan toe zou voegen, en dit is uiteraard onmogelijk. Men kan dus niet beide verlengden als aanvaardbare extrapolaties zien en hiermee zal bij het zoeken naar een zo verantwoord mogelijke wijze van extrapoleren rekening gehouden moeten worden. Wij hopen hierop in een later rapport terug te komen.

5. Samenvatting

Uit het onderhavige onderzoek blijkt duidelijk, dat er een aanzienlijk verschil tussen de verschillende perioden van het jaar bestaat. Voor voorbereidende onderzoekingen over de eigenschappen der hoogwateroverschrijdingslijnen kan men volstaan met de maanden November, December en Januari daaropvolgend, die de gevaarlijkste periode van het jaar vormen. Uitbreiding van deze periode naar voren of naar achteren heeft weinig invloed, indien men de periode niet veel groter maakt. Bij het zoeken naar een extrapolatiemethode en bij een uiteindelijke bepaling van een dijkhoogte houde men rekening met het verschil tussen de hoogwateroverschrijdingslijnen van verschillende perioden van het jaar.

6) Zie b.v. "Over de mogelijkheid van statistische voorspelling van extreem hoge waterstanden en haar grenzen", door Prof. Dr D.van Dantzig en Prof. Dr J.Hemelrijk, Rapport S 114, Mathematisch Centrum.