

K O N I N K L I J K   N E D E R L A N D S  
M E T E O R O L O G I S C H   I N S T I T U U T

Wetenschappelijk Rapport W.R. 55-011 (R III-161)

P.J.Rijkoort

De Frequentieverdeling van de tijdsintervallen tussen  
de opeenvolgende NW stormen en de onderlinge  
afhankelijkheid in het optreden van deze stormen.

De Bilt, 1955.

All Rights Reserved.

Nadruk zonder toestemming van het K.N.M.I. is verboden.

P.J. Rijkoort

De frequentieverdeling van de tijdsintervallen tussen  
de opeenvolgende NW stormen en de onderlinge  
afhankelijkheid in het optreden van deze stormen.

Summary

This study is a part of the Dutch research work on storm-surges and their meteorological causes, which has been performed after the heavy storm-surge of February 1, 1953. It will also be included in the series of meteorological reports which are prepared for the governmental commission (Deltacommission), which was installed after the disaster in order to investigate all its technical consequences.

The storms on December 22 and 24, 1954 gave rise to the present investigation. From a statistical study of the length of intervals between successive storms it is concluded that intervals of two (and perhaps three) days are more frequent than would be expected if the occurrence of storms is mutually independent.

INHOUD

	<u>blz.</u>
0. Inleiding .....	19
1. Homogenisering van het materiaal .....	20
2. De jaarlijkse gang van de stormaantallen .....	20
3. De verdeling van de lengten van de tijdsintervallen tussen de stormen .....	22
4. Toetsing van de afhankelijkheid in het optreden van stormen ..	23
5. Afleiding der formules .....	29
6. Conclusies en resultaten .....	33

0. INLEIDING

0.0 Vlak voor Kerstmis 1954 zijn kort na elkaar twee vrij zware NW stormen opgetreden. Hierdoor is de aandacht gevallen op wat we zouden kunnen noemen: "het verschijnsel van de tweelingstorm". Een synoptische studie van deze december-stormen door C.J. van der Ham: Meteorologisch verslag betreffende de stormen van 21 t/m 24 december 1954 (K.N.M.I., W.R. 55-003 (II-003)) is aan dit rapport voorafgegaan. Daarin wordt op blz. 3 een globale schatting gegeven van de kans op het optreden van stormen kort na elkaar. Deze schatting was gemaakt met behulp van de stormstatistiek van Den Helder (zie te vervolg Meteorologisch rapport stormvloed 1 februari 1953 van mei 1954; voortaan aangeduid met M.R.St.2.) en berustte op de onderstelling, dat het optreden van een storm geheel onafhankelijk is van het optreden van voorafgaande stormen.

Om een indruk te krijgen van de zeldzaamheid van dergelijke tweelingstormen hebben we een onderzoek ingesteld met het materiaal van Den Helder, waarbij de frequentieverdeling van de tijdsintervallen tussen opeenvolgende stormmaxima werd bepaald en waarbij de onderstelling van onafhankelijkheid in het optreden van stormen werd getoetst.

0.1 Het onderzoek is in de volgende fasen uitgevoerd:

a Homogenisering van het materiaal door

1. Verwijdering van z.g. triviale gevallen van afhankelijkheid. Dit zijn gevallen die een gevolg zijn van het feit, dat de windsnelheid tijdelijk beneden de drempelwaarde volgens de stormdefinitie valt door een onregelmatige vorm van het depressie gebied.

2. Splitsing in twee perioden n.l. vóór en ná 1 juli 1922. Er is reden om aan te nemen dat de dagelijkse kans op een storm vóór en ná deze datum verschillend is.

b Aanpassing van een sinusvormige jaarlijkse gang.

c Bepaling van de verdeling (histogram) van de tijdsintervallen tussen opeenvolgende stormmaxima en van de theoretische verdeling die volgt uit de jaarlijkse gang en uit de onderstelling van onafhankelijkheid.

d Toetsing van het grote aantal korte intervallen met behulp van een combinatie van dubbele dichotomiën.

## 0.2 Enkele begrippen

0.2.1 Aangezien bij de jaarlijkse gang het maximum in de winter valt, zullen we het "stormjaar" rekenen te beginnen op 1 juli en te eindigen op 30 juni.

0.2.2 Onder een stormdag zullen we in dit onderzoek verstaan een dag waarop een maximum van de windsnelheid valt, en wel een maximum in de zin van de stormdefinitie op blz. 22 van M.R.St.2.

## 1. HOMOGENISERING VAN HET MATERIAAL

1.1 Met behulp van de weerkaarten is door v.d. Ham nagegaan welke stormen aan één en dezelfde depressie zijn toe te schrijven. Als gevolg van het feit dat bij de definitie van een storm een zeker grensniveau is aangenomen (waarvan het passeren het begin en einde van de storm aangeeft: zie M.R.St.2.), kan het voorkomen dat, bij een onregelmatig gevormde depressie of depressiebaan, de windsnelheid tijdelijk beneden dit niveau daalt, zodat er volgens onze definitie twee stormen zijn geweest, die echter synoptisch gezien als één moeten worden beschouwd. Van deze gevallen is dus alleen de dag met het hoogste maximum als stormdag genomen.

1.2 In fig. 1 hebben we de jaarlijkse aantallen stormen aangegeven. We krijgen hieruit de indruk, dat in het begin de aantallen groter zijn dan later. In M.R.St.2. op blz. 29, is reeds aangetoond dat er significant verschil bestaat in de aantallen stormen vóór en ná omstreeks 1920. Aangezien dit verschil ook bij het materiaal van De Bilt en Vlissingen wordt opgemerkt, is het stellig een reëel meteorologisch effect. Echter is het, wat Den Helder betreft, niet onmogelijk dat de vernieuwde opstelling van de anemometer in 1922 ook enige invloed heeft gehad. We hebben daarom als splitsingsdatum 1 juli 1922 genomen en dus verder met twee afzonderlijke perioden gewerkt:

I	1903/04	.....	1921/22		=	19 jr.
II	1922/23	.....	1943/44, 1945/46	.....	1952/53	= 30 jr.

## 2. DE JAARLIJKSE GANG VAN DE STORMAANTALLEN

Het doel van ons onderzoek is in de eerste plaats de bepaling van de frequentieverdeling van de tijdsintervallen tussen de stormen. Als de kans op een storm het gehele jaar door even groot is, is de

theoretische verdeling van de intervallen zeer eenvoudig te bepalen. Het is echter duidelijk, dat het aanwezig zijn van een jaarlijkse gang in de stormkans de vorm van de frequentieverdeling zal beïnvloeden. Enerzijds zal de grotere stormkans die in de winter bestaat ten gunste van de kortere intervallen werken; anderzijds zal de kleinere stormkans in de zomer ten gunste van langere intervallen werken. Het is dus nodig de jaarlijkse gang in een formule uit te drukken. Aangezien er, zoals we in fig. 2 kunnen zien, geen reden is om in de verdeling over het jaar enige scheefheid te vermoeden, zoals b.v. wel bij het stormvloedmateriaal in De Bilt het geval is (zie M.R.St.2., fig. 2 t.o. blz. 39), hebben we de volgende sinusvormige jaarlijkse gang aangepast:

$$q_i = \alpha + \beta \sin \left( \frac{2\pi i}{J} + \varphi \right) \quad (2.1)$$

hierin geeft

$i = 1, \dots, J$  de dag van het jaar aan ( $1 = 1$  juli enz.)

$J =$  het aantal dagen van het jaar

$q_i =$  kans dat dag  $i$  een stormdag is.

#### Opmerkingen:

1. In fig. 2 zijn de frequenties per dekade uitgezet; hierbij is de juiste lengte van de derde dekade (10, 11 of 8 dagen) in aanmerking genomen.

2. De schrikkel dag is buiten beschouwing gelaten.

Het resultaat van de aanpassing is als volgt:

$$\text{Periode I: } q_i = 0.0292 + 0.01915 \sin \left( \frac{2\pi i}{J} - 73^\circ \right) \quad (2.2)$$

$$\text{Periode II: } q_i = 0.0195 + 0.0143 \sin \left( \frac{2\pi i}{J} - 76^\circ 36' \right) \quad (2.3)$$

De bijzonderheden van de berekening zullen we in 5.1 behandelen. In fig. 2 is door de vloeiende lijn het sinusvormig verloop aangegeven.

Om de  $\chi^2$ -toets toe te passen, ten einde te controleren of de aanpassing bevredigend is, hebben we de theoretische frequenties van de grafiek afgelezen. (Om het rekenwerk niet te langdurig te maken zijn de theoretische frequenties slechts voor een 20-tal  $i$  waarden berekend).

*Uitwerking van de toets voor de stormkans 420*

We hebben daarbij het materiaal in cellen verdeeld, zodat de theoretische frequentie minstens 5 per cel is. Voor periode I vinden we  $\chi^2 = 27.13$  met  $\nu = 23$  dus  $P \sim 0,25$  en voor periode II:  $\chi^2 = 22.07$  met  $\nu = 24$  dus  $P \sim 0,60$ . De overschrijdingskansen zijn dermate groot dat er geen reden is om de onderstelling van een sinusvormige jaarlijkse gang in de stormkans te verwerpen.

3. DE VERDELING VAN DE LENGTEN VAN DE TIJDSINTERVALLEN TUSSEN DE STORMEN

We zullen thans de verdeling van de intervallen tussen de stormen nader onderzoeken.

We definiëren als een tijdsinterval van de lengte k of kortweg "een k-interval": Een periode waarbij dag i en dag i + k stormdag zijn, terwijl de dagen i + 1, ..... i + k - 1 storm-maximum-vrij zijn.

In fig. 3 is met twee histogrammen de verdeling van de k-intervallen aangegeven, zoals die uit het materiaal voor de twee perioden volgt. In beide gevallen zien we een bijzonder sterk maximum bij k = 2.

Hoe zou nu de verdeling van de k-intervallen zijn als we onderstellen dat het optreden van stormen geheel onafhankelijk van elkaar plaats vindt?

Uitgaande van deze onderstelling en met behulp van de jaarlijkse gang, zoals die door formule (2.1) wordt aangegeven, laat zich voor de verwachtingswaarde van het aantal k-intervallen afleiden:

$$A_k = NJ(1-\alpha)^{k-1} \left\{ \left( \alpha^2 + \frac{1}{2}\beta^2 \cos \frac{2\pi k}{J} \right) J_0(iB) + 2\alpha\beta \cos \frac{2\pi k}{J} \cdot iJ_1(iB) - \frac{1}{2}\beta^2 J_2(iB) \right\} \quad (3.1)$$

waarin:  $B = \frac{J\gamma}{\pi} \sin \frac{(k-1)\pi}{J}$

J = lengte van het jaar in dagen

N = aantal jaren van de beschouwde periode

$\alpha$  en  $\beta$  : constanten uit (2.1)

$$\gamma = \frac{\beta}{1-\alpha}$$

$J_0, J_1, J_2$ : Besselse functies van de orde 0, 1, 2.

$$i = \sqrt{-1}$$

De afleiding van deze formule geven we in 5.2.



Aangezien de berekening met deze formule vrij veel tijd kost, hebben we  $A_k$  niet voor alle  $k$ -waarden uitgerekend, maar ons beperkt tot  $k = 1, 5, 10 \dots, 100$ . De aldus verkregen waarden  $A_k$  zijn gebruikt om in fig. 3 de vloeiende kromme te verkrijgen, die dus de theoretische dichtheidsfunctie van de  $k$ -intervallengten is, als we onafhankelijkheid in het optreden van stormen onderstellen.

De vraag die nu onmiddellijk gesteld wordt is: Zijn histogram en theoretische verdeling met elkaar in overeenstemming, of m.a.w.: Is de onderstelling van onafhankelijkheid in het optreden van stormen juist?

In eerste instantie is men geneigd de aanpassing te toetsen met behulp van de  $\chi^2$  methode. Echter is hier een groot bezwaar tegen: Het aantal  $k$ -intervallen voor een bepaalde  $k$ -waarde heeft betrekking op stormen, die niet allen dezelfde kans van optreden bezitten ten gevolge van het aanwezig zijn van de jaarlijkse gang en dit is in strijd met de voorwaarden die nodig zijn om de  $\chi^2$ -toets te kunnen toepassen.

Een tweede bezwaar is nog het volgende: Als het verschil tussen histogram en theoretische verdeling significant zou blijken te zijn, dan zou dit i.p.v. aan een onafhankelijkheid in het optreden van de stormen, ook toegeschreven kunnen worden aan het feit dat in bepaalde jaren meer stormen voorkomen dan in andere jaren.

We zouden natuurlijk kunnen trachten te onderzoeken of dit laatste al dan niet het geval is, echter is tijdens een bespreking op het Mathematisch Centrum door Prof. Hemelrijk een methode bedacht en aanbevolen die de mogelijkheid biedt beide effecten te scheiden en de al of niet afhankelijkheid der stormen te toetsen zonder daarbij hinder te ondervinden van het verschillende aantal stormen in verschillende jaren.

We zullen dus in de volgende paragraaf overgaan tot de bespreking van deze methode en de hiermede bereikte resultaten.

#### 4. TOETSING VAN DE AFHANKELIJKHEID IN HET OPTREDEN VAN STORMEN

##### 4.1 De toetsingsmethode

We kunnen onze nulhypothese als volgt formuleren:

De kans dat dag  $i + k$  een stormdag is, wanneer dag  $i$  een stormdag is, is even groot als de kans dat dag  $i + k$  een stormdag is, wanneer dag  $i$  geen stormdag is.

Om deze hypothese te toetsen voor een bepaalde  $i$  en  $k$  waarde kunnen we gebruik maken van de methode van de dubbele dichotomie. Een uitvoerige bespreking van deze toetsingsmethode is b.v. te vinden in Statistica 7 no 3. 1953 p. 141: "Methoden voor het vergelijken van onbekende kansen" door C. van Eeden.

Om de methode handig uiteen te kunnen zetten, voeren we de volgende symboliek in:

$S_i$  betekent: dag  $i$  is een stormdag

$\bar{S}_i$  betekent: dag  $i$  is geen stormdag.

We bepalen nu het aantal gevallen  $(S_i, S_{i+k}) = a$ ,  $(S_i, \bar{S}_{i+k}) = b$ ,  $(\bar{S}_i, S_{i+k}) = c$  en  $(\bar{S}_i, \bar{S}_{i+k}) = d$ . In een z.g. dubbele dichotomie gezet wordt dit als volgt:

	$S_{i+k}$	$\bar{S}_{i+k}$	
$S_i$	<u>a</u>	<u>c</u>	n
$\bar{S}_i$	<u>b</u>	<u>d</u>	m
	r	s	N

Hierbij is  $n = \underline{a} + \underline{c}$

$m = \underline{b} + \underline{d}$

$r = a + b$

$s = \underline{c} + \underline{d}$

en  $N = r + s = n + m$

De strepen onder  $a$ ,  $b$ ,  $c$ ,  $d$ , willen zeggen dat deze grootheden als stochastische worden beschouwd.

Als we nu de kans op  $(S_i, S_{i+k})$  d.w.z. kans dat dag  $i+k$  stormdag is als dag  $i$  stormdag is, kortweg met  $P(S_{i+k} | S_i)$  aangeven, dan kunnen we het volgende zeggen: als b.v.  $P(S_{i+k} | S_i) > P(S_{i+k} | \bar{S}_i)$  dan is er dus een zekere afhankelijkheid en zal a relatief groot worden. a kan b.v. als toetsingsgrootheid worden gebruikt. Voor nadere bijzonderheden verwijzen wij naar het artikel van Mej. van Eeden.

Om de getalwaarden van  $a$ ,  $b$ ,  $c$  en  $d$  snel te bepalen maken we gebruik van de doorlopende nummering van de dagen van het jaar: 1 ..... J. We plaatsen in een diagram de jaren onder elkaar en laten dus overeenkomende data in één kolom voorkomen. In dit diagram geven we de dagen die "stormdagen" zijn met een kruis aan, b.v. als volgt:

Jaar \ dag .....	i	i+1	i+2	i+3	i+4
j	-	-	-	x	-
j+1	-	x	-	-	-
j+2	x	-	x	-	-
j+3	-	-	-	x	-
j+4	-	x	-	-	-
j+5	x	-	-	-	x
j+6	-	-	-	-	x

In dit voorbeeld zouden we voor i en i+4 b.v. vinden a = 1, b = 1, c = 1 en d = 4.

Aangezien wij de significantie niet voor de i-waarden afzonderlijk willen onderzoeken kunnen we voor één bepaalde k-waarde, i de waarden 1, ..... J-k laten doorlopen en aldus een reeks dubbele dichotomieën verkrijgen. Een simultane toetsingsgrootheid voor een dergelijke combinatie van dubbele dichotomieën vinden we weer in het artikel in Statistica n.1.

$$\underline{x} = \frac{\sum_i (a_i - \frac{n_i r_i}{N_i})}{\sqrt{\sum_i \frac{n_i m_i r_i s_i}{N_i^2 (N_i - 1)}}$$

Onder de nulhypothese is deze grootheid bij benadering normaal verdeeld met gemiddelde 0 en spreiding 1.

Deze grootheid  $\underline{x}$  is nu bepaald voor k-waarden 2 tot en met 15. Bij k = 2 is nog een enigszins gewijzigde methode gevolgd om de waarden van a, b, c en d te bepalen. We moeten namelijk rekening houden met het feit dat iedere storm een zekere uitgestrektheid in de tijd bezit. Om een indruk te geven van deze uitgestrektheid geven we de volgende cijfers van de verdeling van de stormduren op het  $1\frac{1}{2}$  m niveau ( $A_{1\frac{1}{2}}$  uit M.R.St.2):

$$A_{1\frac{1}{2}} > 24 \text{ uur in } 60 \% \text{ van het totale aantal}$$

$$A_{1\frac{1}{2}} > 48 \text{ uur in } 11 \% \text{ van het totale aantal.}$$

Aangezien we kunnen aanemen, dat de duur van maximum tot einde van een storm onafhankelijk is zowel van de duur van begin tot maximum van dezelfde storm als van duur van begin tot maximum

van de volgende storm, geven bovenstaande cijfers tevens een indruk van de duren van maxima tot maxima indien alle stormen op het  $1\frac{1}{2}$  m niveau zouden aaneen sluiten. Hieruit volgt, dat in een groot aantal gevallen de dag vlak voor of na een stormdag geen storm(maximum)dag kan zijn, m.a.w. als in een bepaald jaar dag  $i+1$  een stormdag is dan is al min of meer bepaald dat hierdoor voor de kolommen  $i$  en  $i+2$  een geval  $(\bar{S}_i, \bar{S}_{i+2})$  wordt geleverd. Daarom hebben we bij telling in de kolommen  $i$  en  $i+2$  die gevallen waarbij in kolom  $i+1$  een kruis staat weggelaten.

Bij  $k = 3$  is dit effect verwaarloosd.

4.2

Wat kunnen we nu verwachten van de uitkomsten van de berekening van de  $\underline{x}$ -waarden? Als de nulhypothese juist is dan zullen de  $\underline{x}$ -waarden voor 95% binnen de grenzen  $\pm 196$  blijven en gemiddeld vrijwel nul zijn. Als er echter afhankelijkheid bestaat, dan is te verwachten dat de  $\underline{x}$ -waarden voor kleine  $k$ -waarden sterk positief zijn en daarna afnemen totdat de afhankelijkheid niet meer bestaat.

Echter kunnen de  $\underline{x}$ -waarden ook nog op grond van een ander effect positief uitvallen. Aan het eind van 3 hebben we reeds op de mogelijkheid gewezen, dat er een effect ontstaat doordat sommige jaren meer stormen bevatten dan andere.

Voor de dubbele dichotomie betekent dit, dat de  $n$  gevallen  $S_i$  vooral van de stormrijke jaren afkomstig zijn en de  $m$  gevallen  $\bar{S}_i$  van de stormarme jaren. Dan zal echter de kans op  $S_{i+k}$  bij de  $n$   $S_i$ -gevallen groter zijn dan bij de  $m$   $\bar{S}_i$ -gevallen, zodat ook nu  $\underline{x}$  gemiddeld  $> 0$  zal zijn. Echter zal dit nu onafhankelijk van  $k$  zijn.

Het resultaat van de berekening is als volgt:

Tabel 1  
 Waarden van de toetsingsgrootheid  $\underline{x}$

k	Periode I	Periode II	Periode I en II simultaan
2	+3,75	+4,00	+5,42
3	+2,87	+1,26	+2,93
4	+0,83	-0,18	+0,52
5	+2,78	-1,09	+1,41
6	-0,64	+1,19	+0,31
7	-0,09	+1,05	+0,61
8	+2,08	-0,82	+1,05
9	+0,80	-1,12	-0,17
10	+1,00	+0,40	+1,03
11	+1,69	+0,47	+1,66
12	-0,24	+1,87	+1,11
13	+1,40	+1,85	+1,85
14	-1,00	+1,72	+0,20
15	-0,79	+1,21	+0,23

De laatste kolom van deze tabel geeft de resultaten als we beide perioden simultaan beschouwen, d.w.z. als we voor de bepaling van  $\underline{x}$  de sommatie uitstrekken over alle dubbele dichotomiën zowel van periode I als van periode II.

We zien uit de tabel dat de  $\underline{x}$ -waarde voor  $k = 2$  sterk significant positief is. Aangezien we slechts geïnteresseerd zijn voor de vraag of  $P(S_{i+k} > S_i) > P(S_{i+k} > \bar{S}_i)$  kunnen we hier eenzijdig toetsen, zodat de 5 % grens bij +1,63 ligt. Simultaan is ook de  $\underline{x}$ -waarde voor  $k = 3$  significant. Voor grotere  $k$ -waarden neemt  $\underline{x}$  af maar blijft toch nog overwegend positief. Het is dus mogelijk dat het effect van jaren met veel stormen tegenover jaren met weinig stormen een zekere rol speelt. We hebben daarom nog een nieuwe berekening gemaakt, waarvoor we het materiaal nog verder gesplitst hebben ten einde dit effect enigszins te elimineren.

De splitsing is als volgt:

groep	$I_0$	:	Jaren met meer dan 10 stormen van I		
"	$I_1$	:	" " hoogstens 10	"	" I
"	$II_0$	:	" " meer dan 8	"	" II
"	$II_1$	:	" " 6, 7 of 8	"	" II
"	$II_2$	:	" " hoogstens 5	"	" II.

Het resultaat geeft tabel 2:

Tabel 2

Waarden van de toetsingsgrootheid  $\underline{x}$

k	$I_0$	$I_1$	$I_0+I_1$ sim.	$II_0$	$II_1$	$II_2$	$II_0+II_1+II_2$ sim.	$I_0+\dots+II_2$ sim.
2	+2,59	+2,80	+3,80	+2,31	+1,98	+2,23	+3,52	+5,20
3	+2,57	+0,29	+2,40	+2,16	+0,58	-0,33	+2,00	+3,10
4	+1,86	-0,22	+1,58	-1,66	+1,63	-0,82	-0,70	+0,75
5	+1,66	+3,60	+3,32	+0,75	-1,13	-0,45	-0,18	+2,65
6	-0,41	-1,11	-0,96	+0,38	+0,96	-0,95	+0,46	-0,43
7	-0,57	+0,08	-0,43	+1,07	-0,53	+1,05	+0,93	+0,25
8	+3,32	-0,09	+2,71	-0,13	-0,90	-0,50	-0,60	+1,77
9	-0,49	+1,36	+0,18	-0,84	-1,36	-	-1,48	-0,79
10	+0,43	+0,77	+0,77	+0,30	-1,16	-0,60	-0,45	+0,31
11	+0,38	+0,75	+0,73	+0,65	-0,28	+1,40	+0,73	+0,99
12	-0,50	+0,09	-0,38	+0,87	+0,87	+1,07	+1,56	+0,78
13	+1,42	+0,50	+1,48	+1,34	+1,34	-0,74	+1,52	+2,11
14	-1,18	+0,00	-0,97	+1,95	+1,23	+0,57	+2,33	+0,66
15	-0,78	-0,21	-0,79	-1,09	+2,87	-0,83	+0,40	-0,31

In fig. 4 hebben we de totale simultane resultaten van beide berekeningen in beeld gebracht. Hoewel het verschil tussen de resultaten niet groot is, is er wel de tendens dat de  $\underline{x}$ -waarden in het tweede geval sterker naar nul gaan dan bij de eerste methode; dit blijkt uit het volgende staatje:

	$I + II$	$I_0 + I_1 + II_0 + II_1 + II_2$
$\bar{x}$ (k = 2, ....., 5)	+2,57	+2,92
$\bar{x}$ (k = 6, ....., 15)	+0,80	+0,53

Hoewel ook in het tweede geval de  $\underline{x}$ -waarden nog overwegend aan de positieve kant liggen, is het verschil tussen de  $\underline{x}$ -waarde voor k = 2 en de rest toch wel zo groot, dat aan de realiteit van het afhankelijkheidseffect niet behoeft te worden getwijfeld.

## 5. AFLEIDING DER FORMULES

### 5.1 De jaarlijkse gang van het stormmaximum

Aanpassing aan het histogram van de functie

$$q_i = \alpha + \beta \sin\left(\frac{2\pi i}{J} + \varphi\right)$$

$q_i$  is de kans dat dag  $i$  een stormdag is. Als schatting voor deze kans gebruiken we  $f_i = \frac{S_d}{a \cdot N}$  waarin:

$S_d$  = aantal stormen per dekade

$a$  = aantal dagen van de beschouwde dekade (10, 11 of 8)

$N$  = aantal jaren van de beschouwde periode.

Voor de aanpassing maken we volgens de methode der kleinste kwadraten de volgende grootte minimaal:

$$F = \sum_{i=1}^J \left\{ f_i - \alpha - \beta \sin\left(\frac{2\pi i}{J} + \varphi\right) \right\}^2 \quad (5.1.1)$$

Dus

$$1^\circ \quad \frac{\partial F}{\partial \alpha} = 0 \quad \text{of} \quad \sum_{i=1}^J \left\{ f_i - \alpha - \beta \sin\left(\frac{2\pi i}{J} + \varphi\right) \right\} = \sum_{i=1}^J f_i - J\alpha = 0$$

derhalve

$$\alpha = \frac{\sum_{i=1}^J f_i}{J} = \bar{f}_i \quad (5.1.2)$$

$$2^\circ \quad \frac{\partial F}{\partial \varphi} = 0 \quad \text{of} \quad -\sin \varphi \sum_{i=1}^J f_i \sin \frac{2\pi i}{J} + \cos \varphi \sum_{i=1}^J f_i \cos \frac{2\pi i}{J} = 0$$

waaruit

$$\operatorname{tg} \varphi = \frac{\sum_{i=1}^J f_i \cos \frac{2\pi i}{J}}{\sum_{i=1}^J f_i \sin \frac{2\pi i}{J}} \quad (5.1.3)$$

$$3^\circ \quad \frac{\partial F}{\partial \beta} = 0 \quad \text{of} \quad \cos \varphi \sum_{i=1}^J f_i \sin \frac{2\pi i}{J} + \sin \varphi \sum_{i=1}^J f_i \cos \frac{2\pi i}{J} - \beta \sum_{i=1}^J \sin^2\left(\frac{2\pi i}{J} + \varphi\right) = 0$$

waaruit

$$\beta = \frac{2}{J} \left\{ \cos \varphi \cdot \sum_{i=1}^J f_i \sin \frac{2\pi i}{J} + \sin \varphi \sum_{i=1}^J f_i \cos \frac{2\pi i}{J} \right\} \quad (5.1.4)$$

Met behulp van de formules (5.1.2), (5.1.3) en (5.1.4) werden (2.2) en (2.3) berekend.

5.2 De theoretische verdeling van de lengten van de tijdsintervallen tussen twee opeenvolgende stormen onder de onafhankelijkheids hypothese

De kans op een k-interval dat begint bij dag i noemen we  $p_{k, i}$ . Deze kans kunnen we als volgt schrijven

$$p_{k, i} = (\alpha + \beta s_i') (\alpha + \beta s_{i+k}') \prod_{j=1}^{k-1} (1 - \alpha - \beta s_{i+j}')^2$$

hierin is  $S_i'$  de notatie voor  $\sin(\frac{2\pi i}{J} + \varphi)$

Gemiddeld over het jaar is deze kans:

$$p_k = \frac{1}{J} \sum_{i=1}^J p_{k, i} = \frac{1}{J} \sum_{i=1}^J (\alpha + \beta s_i') (\alpha + \beta s_{i+k}') \prod_{j=1}^{k-1} (1 - \alpha - \beta s_{i+j}')^2$$

met  $i = i' - \frac{J\varphi}{2\pi}$  en  $s_{i' - \frac{J\varphi}{2\pi}} = s_{i'}$  wordt dit

$$\frac{1}{J} \sum_{i=1+\frac{J\varphi}{2\pi}}^{J+\frac{J\varphi}{2\pi}} (\alpha + \beta s_i) (\alpha + \beta s_{i+k}) \prod_{j=1}^{k-1} (1 - \alpha - \beta s_{i+j})^2$$

Als we de index bij i weglaten krijgen we met  $\gamma = \frac{\beta}{1-\alpha}$

$$p_k = \frac{1}{J} (1-\alpha)^{k-1} \sum_{i=1+\frac{J\varphi}{2\pi}}^{J+\frac{J\varphi}{2\pi}} \left\{ \alpha^2 + \alpha\beta(s_i + s_{i+k}) + \beta^2 s_i s_{i+k} \right\} \prod_{j=1}^{k-1} (1 - \gamma s_{i+j})^2$$

Nu is  $\gamma s_{i+j} < 1$  en dus nemen we voor  $1 - \gamma s_{i+j} : e^{-\gamma s_{i+j}}$

Welke fout maken we?

Relatief is de fout

$$\prod_{j=1}^{k-1} \frac{e^{-\gamma s_{i+j}}}{1 - \gamma s_{i+j}} - 1 = \prod_{j=1}^{k-1} \frac{1 - \gamma s_{i+j} + \frac{1}{2} \gamma^2 s_{i+j}^2 + \Delta}{1 - \gamma s_{i+j}} - 1$$

Als  $S_{i+j} > 0$  dan is  $\Delta$  negatief, dus is de fout  $< \prod_{j=1}^{k-1} \frac{1 - \gamma s_{i+j} + \frac{1}{2} \gamma^2 s_{i+j}^2}{1 - \gamma s_{i+j}} - 1 =$

$$\prod_{j=1}^{k-1} \left( 1 + \frac{\gamma^2 s_{i+j}^2}{2(1-\gamma s_{i+j})} \right) - 1 < \exp\left(\sum_{j=1}^{k-1} \frac{\gamma^2 s_{i+j}^2}{2(1-\gamma s_{i+j})}\right) - 1 <$$

$$\exp\left(-\frac{\gamma^2}{2(1-\gamma)} \sum_{j=1}^{k-1} s_{i+j}^2\right) - 1 < \exp\left(\frac{\gamma^2(k-1)}{2(1-\gamma)}\right) - 1 \sim 0,0002 (k-1)$$

voor  $\gamma = 0,02$ , (in beide periode is  $\gamma < 0,02$  zodat de fout in werkelijkheid nog kleiner is)

Als  $S_{i+j} < 0$  dan is  $\Delta$  positief n.l.  $\Delta = \frac{\gamma^3 |s_{i+j}|^3}{3!} + \frac{\gamma^4}{4!} + \dots =$

$$\frac{1}{6} \gamma^3 |s_{i+j}|^3 \left( 1 + \frac{\gamma(s_{i+j})}{4} + \dots \right) < \frac{1}{6} \gamma^3 |s_{i+j}|^3 \exp \gamma (s_{i+j}) < \frac{1}{2} \gamma^3$$

dan is de relatieve fout  $< \exp\left(\frac{\gamma^2 + \gamma^3}{2(1-\gamma)} (k-1)\right) - 1 \sim 0,0002 (k-1)$



Voor  $k \sim 100$  is de fout dus kleiner dan 2 %.

We krijgen nu

$$h_k = \frac{1}{J} (1-\alpha)^{k-1} \sum_{i=1+\frac{y}{2\pi}}^{\frac{y+k}{2\pi}} \left\{ \alpha^2 + \alpha\beta (s_i + s_{i+k}) + \beta^2 s_i s_{i+k} \right\} e^{-\delta \sum_{j=1}^{k-1} s_{i+j}}$$

We voeren nu in  $i = i' - \frac{1}{2}k$ , dan

$$h_k = \frac{1}{J} (1-\alpha)^{k-1} \sum_{i'=1+\frac{y}{2\pi}}^{\frac{y+k}{2\pi}} \left\{ \alpha^2 + \alpha\beta (s_{i'-\frac{1}{2}k} + s_{i'+\frac{1}{2}k}) + \beta^2 s_{i'-\frac{1}{2}k} s_{i'+\frac{1}{2}k} \right\} e^{-\delta \sum_{j=1}^{k-1} s_{i'-\frac{1}{2}k+j}}$$

Nu is 
$$\sum_{i=p}^q \sin ai = \frac{\sin \frac{q+p}{2} a \cdot \sin \frac{1+q-p}{2} a}{\sin a/2}$$

Dus

$$\sum_{j=1}^{k-1} s_{i'-\frac{1}{2}k+j} = \sum_{n=i'-\frac{1}{2}k+1}^{i'+\frac{1}{2}k-1} \sin \frac{2\pi}{J} n = \frac{\sin \frac{2\pi i'}{J} \sin \frac{k-1}{2} \cdot \frac{2\pi}{J}}{\sin \frac{\pi}{J}} \approx \frac{J}{\pi} \sin \frac{(k-1)\pi}{J} \sin \frac{2\pi i'}{J}$$

Derhalve:

$$h_k = \frac{(1-\alpha)^{k-1}}{J} \sum_{i'=1+\frac{y}{2\pi}}^{\frac{y+k}{2\pi}} \left( \alpha^2 + 2\alpha\beta C_{\frac{1}{2}k} s_{i'} + \beta^2 s_{i'}^2 - \beta^2 s_{\frac{1}{2}k}^2 \right) e^{-\frac{\delta J}{\pi} s_{\frac{1}{2}k} s_{i'}}$$

$$= C \sum_{i'=1+\frac{y}{2\pi}}^{\frac{y+k}{2\pi}} \left( A_1 + A_2 s_{i'} + A_3 s_{i'}^2 \right) e^{-B s_{i'}}$$

hierin is

$$C_{\frac{1}{2}k} = \cos \frac{\pi}{J}$$

$$C = \frac{1}{J} (1-\alpha)^{k-1}$$

$$A_1 = \alpha^2 - \beta^2 \sin^2 \frac{\pi k}{J}$$

$$A_2 = 2\alpha\beta \cos \frac{\pi k}{J}$$

$$A_3 = \beta^2$$

$$B = \frac{\delta J}{\pi} \sin \frac{(k-1)\pi}{J}$$

We gaan nu van de sommatie  $\sum$  over op de integraal, we vervangen n.l.  $f(i') = e^{-B s_{i'}}$ ,  $s_i e^{-B s_i}$  resp.  $s_i^2 e^{-B s_i}$  door

door 
$$\int_{i=i'-\frac{1}{2}k}^{i'+\frac{1}{2}k} f(u) du$$

waarmee

$$h_k = C A_1 \int_{i=\frac{y}{2\pi}}^{\frac{y+k}{2\pi}} e^{-B s_i} du + C A_2 \int_{i=\frac{y}{2\pi}}^{\frac{y+k}{2\pi}} s_i e^{-B s_i} du + C A_3 \int_{i=\frac{y}{2\pi}}^{\frac{y+k}{2\pi}} s_i^2 e^{-B s_i} du$$

of

$$h_k = C_1 \int_0^{2\pi} e^{-\beta \sin x} dx + C_2 \int_0^{2\pi} \sin x \cdot e^{-\beta \sin x} dx + C_3 \int_0^{2\pi} \sin^2 x \cdot e^{-\beta \sin x} dx \quad (5.2.1)$$

Hierbij is

$$\begin{aligned} C_1 &= \frac{1}{2\pi} \cdot \frac{1}{\beta} (1-\alpha)^{k-1} A_1 = \frac{(1-\alpha)^{k-1}}{2\pi} A_1 \\ C_2 &= \frac{(1-\alpha)^{k-1}}{2\pi} A_2 \\ C_3 &= \frac{(1-\alpha)^{k-1}}{2\pi} A_3 \end{aligned}$$

We kunnen formule (5.2.1) met behulp van Bessel-functies nog anders schrijven. De algemene integraalformule voor Bessel-functies luidt n.l.

$$J_n(z) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{i(z \sin x - nx)} dx = \frac{1}{2\pi} \int_0^{2\pi} (\cos nx - i \sin nx) e^{iz \sin x} dx$$

Dus:

$$J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{iz \sin x} dx \quad \text{of} \quad \int_0^{2\pi} e^{-\beta \sin x} dx = 2\pi J_0(i\beta) \quad (5.2.2)$$

Verder

$$J_1(z) = \frac{1}{2\pi} \int_0^{2\pi} \cos x \cdot e^{iz \sin x} dx - \frac{i}{2\pi} \int_0^{2\pi} \sin x \cdot e^{iz \sin x} dx$$

Nu is  $\int_0^{\frac{1}{2}\pi} \cos x \cdot e^{iz \sin x} dx = - \int_{\frac{1}{2}\pi}^{\pi} \cos x \cdot e^{iz \sin x} dx$  enz.

dus  $\int_0^{2\pi} \cos x \cdot e^{iz \sin x} dx = 0$

Derhalve  $J_1(z) = \frac{1}{2\pi i} \int_0^{2\pi} \sin x \cdot e^{iz \sin x} dx$

of  $\int_0^{2\pi} \sin x \cdot e^{-\beta \sin x} dx = 2\pi i J_1(i\beta) \quad (5.2.3)$

Tenslotte

$$J_2(z) = \frac{1}{2\pi} \int_0^{2\pi} \cos 2x \cdot e^{iz \sin x} dx + \frac{1}{2\pi i} \int_0^{2\pi} \sin 2x \cdot e^{iz \sin x} dx$$

only derived with  $J_1 = \frac{1}{2\pi i} \int_0^{2\pi} \sin x \cdot e^{iz \sin x} dx$

Nu is

$$\int_0^{2\pi} \sin 2x e^{i2 \sin x} dx = 2 \int_0^{2\pi} \sin x \cdot \cos x \cdot e^{i2 \sin x} dx = 0$$

en dus

$$f_2(x) = \frac{1}{2\pi} \int_0^{2\pi} (1 - 2 \sin^2 x) e^{i2 \sin x} dx = \frac{1}{2\pi} \int_0^{2\pi} e^{i2 \sin x} dx - \frac{1}{\pi} \int_0^{2\pi} \sin^2 x e^{i2 \sin x} dx$$

Derhalve

$$\int_0^{2\pi} \sin^2 x e^{-iB \sin x} dx = \pi f_0(iB) - \pi f_2(iB) \quad (5.2.4)$$

We krijgen met (5.2.2)°, (5.2.3) en (5.2.4) in (5.2.1):

$$\begin{aligned} p_k &= 2\pi C_1 f_0(iB) + 2\pi i C_2 f_1(iB) + \pi C_3 f_0(iB) - \pi C_3 f_2(iB) = \\ &= (1-\alpha)^{k-1} \left\{ (\alpha^2 - \beta^2 \sin^2 \frac{2\pi k}{J} + \frac{1}{2} \beta^2) f_0(iB) + 2\alpha\beta \cos \frac{\pi k}{J} \cdot i f_1(iB) - \frac{1}{2} \beta^2 f_2(iB) \right\} \end{aligned}$$

De eindformule wordt:

$$A_k = N \cdot p_k = N (1-\alpha)^{k-1} \left\{ (\alpha^2 + \frac{1}{2} \beta^2 \cos \frac{2\pi k}{J}) f_0(iB) + 2\alpha\beta \cos \frac{\pi k}{J} \cdot i f_1(iB) - \frac{1}{2} \beta^2 f_2(iB) \right\}$$

met  $B = \frac{J\beta}{\pi} \sin \frac{(k-1)\pi}{J}$

## 6. CONCLUSIES EN RESULTATEN

Uit een bewerking van het stormmateriaal van Den Helder is gebleken:

1° De jaarlijkse gang in de dagelijkse kans op een stormmaximum laat zich beschrijven door een goniometrische functie van de vorm

$$q_i = \alpha + \beta \sin \left( \frac{2\pi i}{J} + \varphi \right)$$

Hierin is  $q_i$  de kans dat dag  $i$  een stormmaximumdag is, terwijl  $i$  de doorlopende nummering (gedurende een jaar) van de dagen volgt.  $J$  is het totaal aantal dagen van het jaar. Verder zijn  $\alpha$ ,  $\beta$  en  $\varphi$  constanten.

2<sup>o</sup>        Uit een statistische bewerking van de tijdsintervallen tussen opeenvolgende stormen blijkt, dat de dagelijkse kans op een storm, zoals die door de onder 1<sup>o</sup> genoemde formule gegeven wordt, slechts betrekkelijke waarde heeft. 2 of 3 dagen na een stormdag is de kans op een nieuw stormmaximum groter dan de kans dat dezelfde dagen stormdagen zijn als er geen stormdag vooraf was gegaan, m.a.w. er is een zekere afhankelijkheid in het optreden van stormen.

De schrijver is Prof. Hemelrijk bijzonder erkentelijk voor zijn belangrijke suggesties voor de oplossing van het probleem, dat in dit artikel behandeld is.

#### Addenda

1        Na het tot stand komen van dit rapport in bovenstaande vorm, leek het ons toch nog gewenst het histogram van de k-waarden voor kleine k-waarden te preciseren. We voegen daarom fig. 5 toe, waarop de frequentie-verdeling van de lengten van de intervallen tussen opeenvolgende stormmaxima in uren ( $k_u$ ) gegeven is voorzover deze lengtenkleiner dan 70 zijn.

Hieruit blijkt dat intervallen kleiner dan 20 uur niet zijn voorgekomen. Verder zien we dat het maximum van het histogram bij 35/39 uur ligt. Dit wil echter niet zeggen dat de modus van het universum van de stormintervallen ook bij 35-39 uur ligt. Hoogstens kunnen we zeggen dat deze vermoedelijk wel niet ver er vandaan zal liggen.

2        Tenslotte zijn ter aanvulling van tabel 2 nog  $\bar{x}$ -waarden berekend voor  $k = 31 \dots 39$ . Deze zijn als volgt  $-0.44, +0.96, -0.05, -0.86, -0.78, -0.63, +0.96, +0.31$  en  $-0.08$ . Ze wijken in het gemiddelde niet significant van 0 af, d.w.z. het op pag. 23 genoemde effect van storm-rijke resp. -arme jaren is dus blijkbaar toch niet aanwezig.

Dat voor  $k < 15$   $\bar{x}$  toch nog overwegend positief is moet derhalve geheel aan het afhankelijkheidseffect worden toegeschreven.

Fig.1. Aantal stormen per jaar(juli-juni)te Den Helder

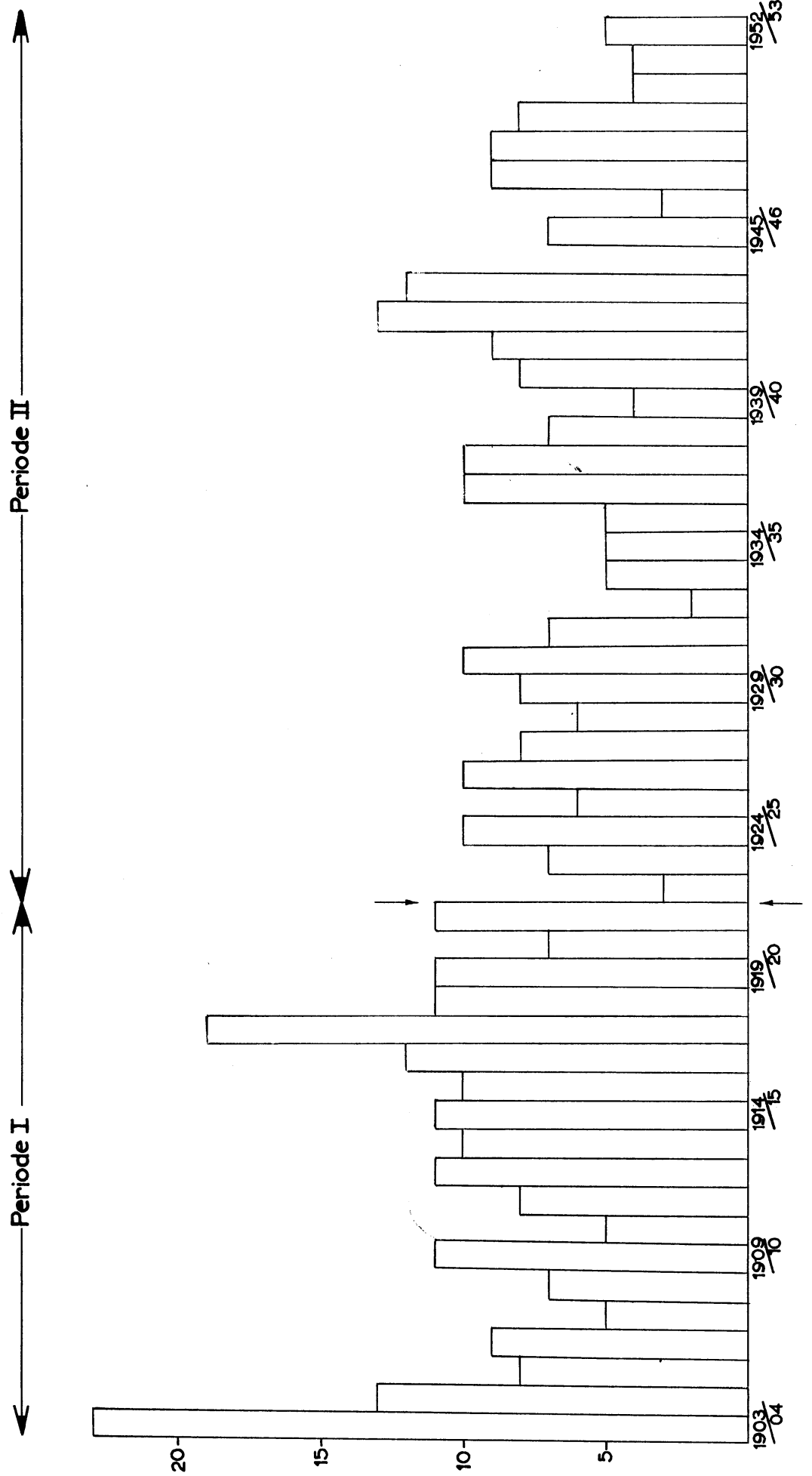


Fig.2 Jaarlijkse gang van de stormfrequentie

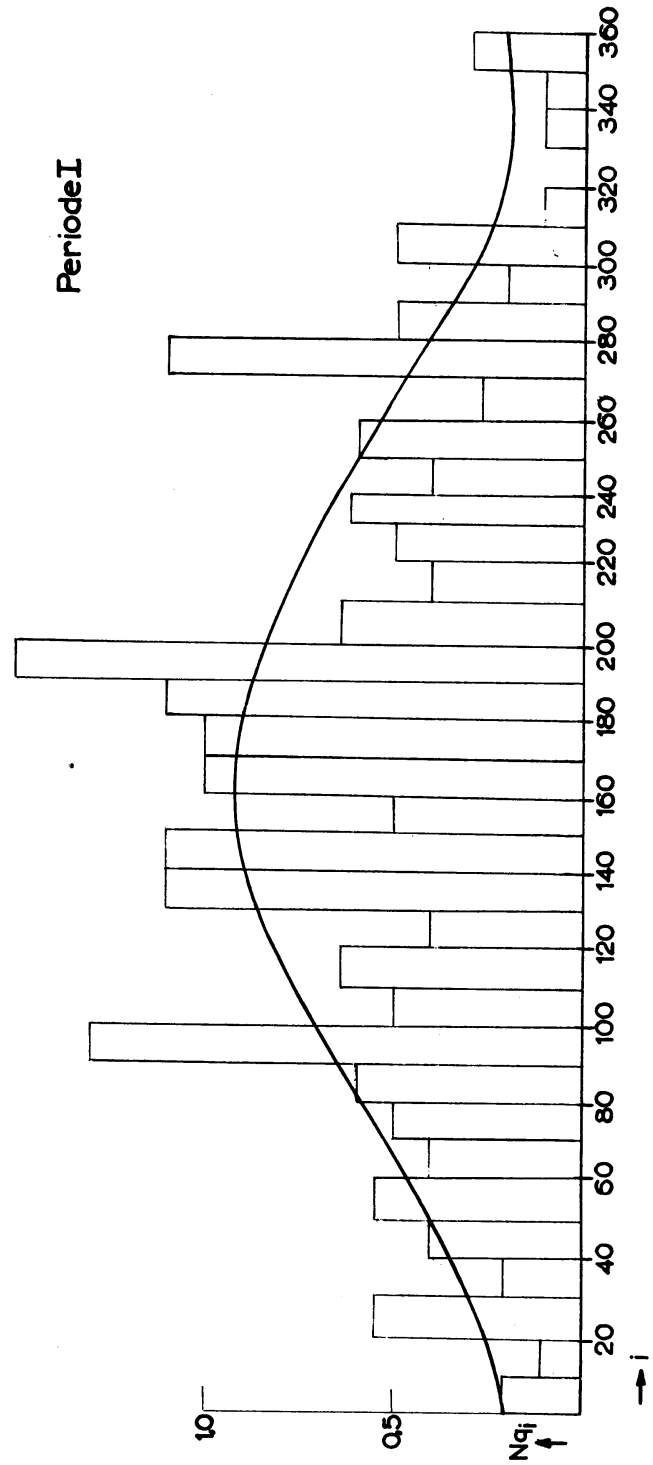
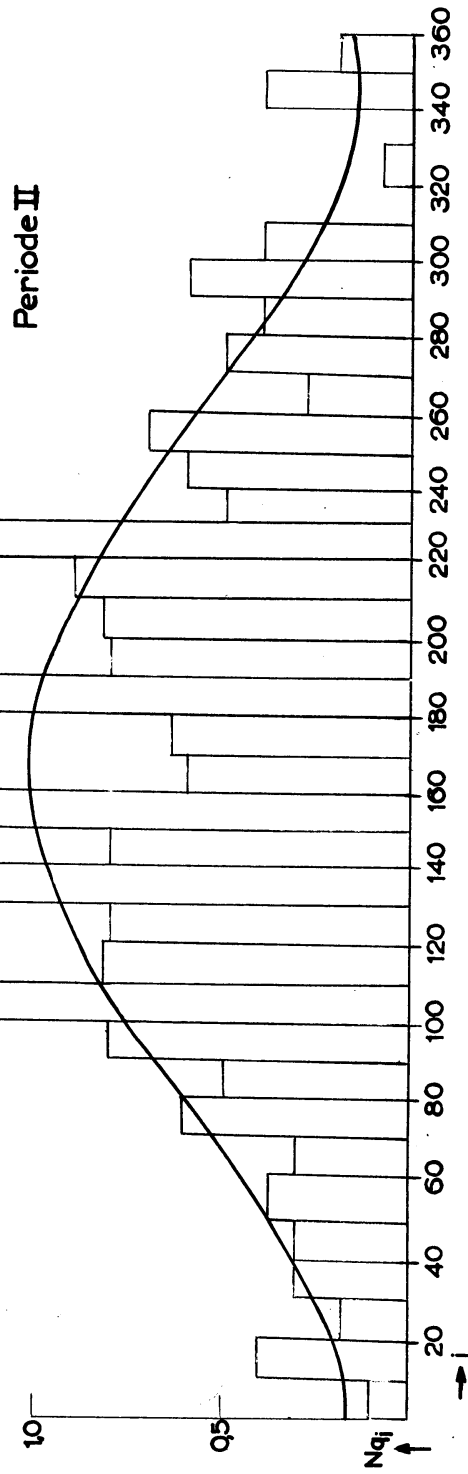


Fig.3 Histogrammen van de intervallengte (in dagen L) tussen opeenvolgende N.W. stormen waargenomen te Den Helder

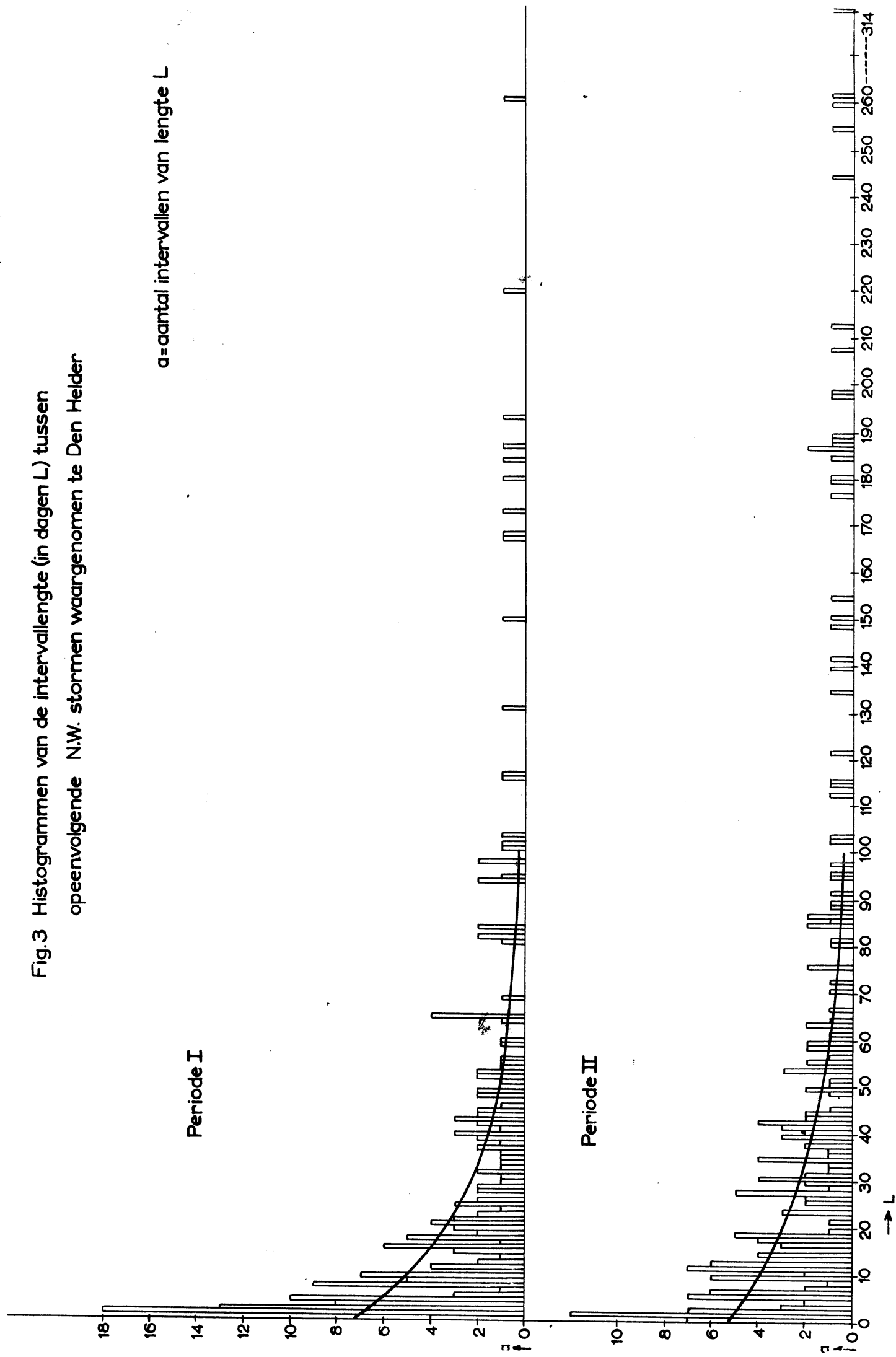


Fig. 4 Combinatie van dubbele dichotomiën  
 Toetsingsgrootheid  $\underline{x}$  als functie van  $k$

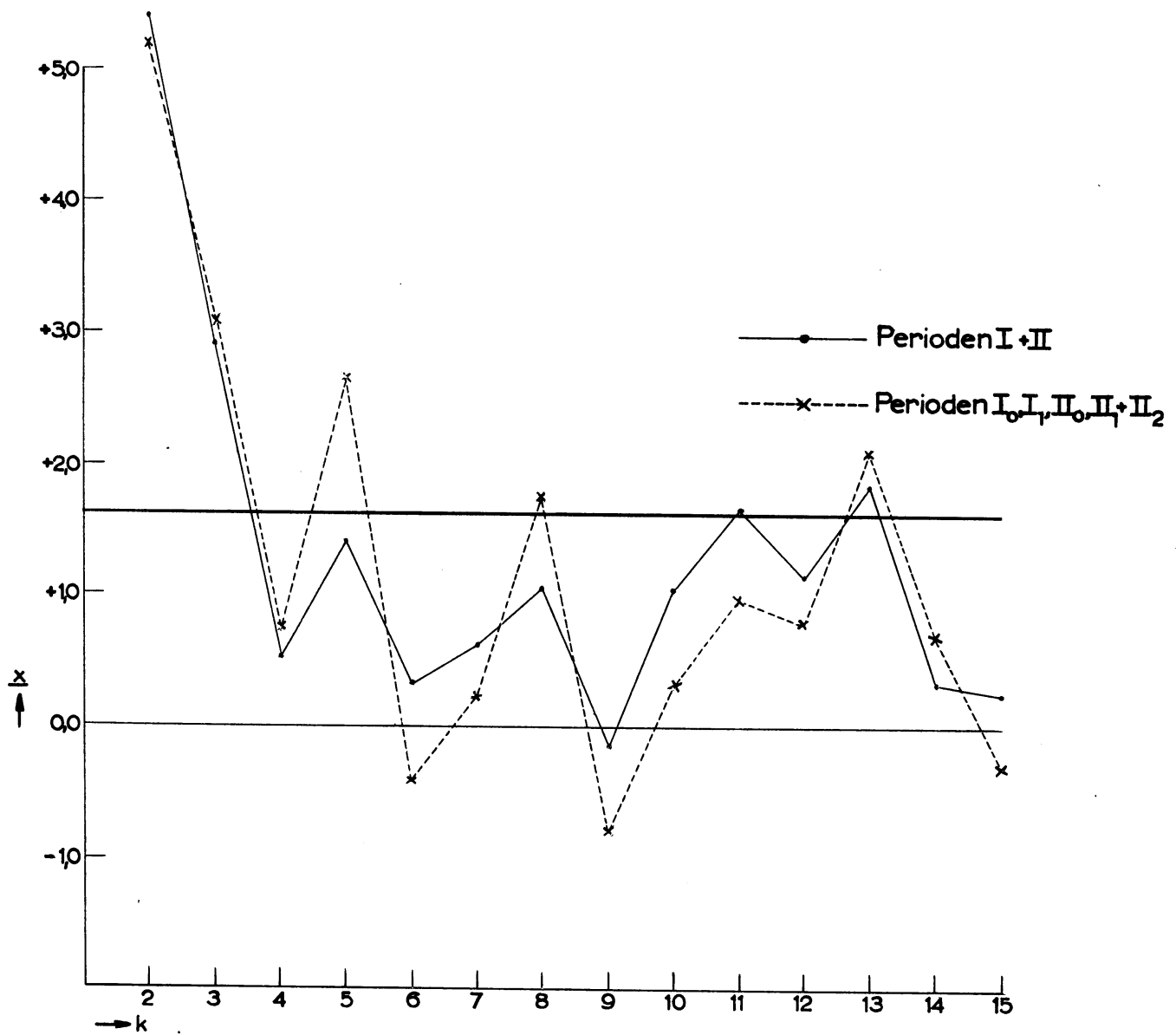




Fig. 5

Frequentie verdeling van storm maximum intervallen in uren voor  $K_y < 70$

