

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

D e B i l t

WETENSCHAPPELIJK RAPPORT

W.R. 72-9

L.J.L. Deij

Over de zeldzaamheid van het optreden
van een droge winter gevolgd door een
droge lente, en van een droge herfst+
winter gevolgd door een droge lente +
zomer.

De Bilt, 1972.

Publikationsnummer: K.N.M.I. W.R. 72-9 (MO)

U.D.C.: 551.509.331 :
551 509.335 :
551.577.35

Over de zeldzaamheid van het optreden van een droge winter
gevolgd door een droge zomer en van een droge herfst + winter
gevolgd door een droge lente + zomer

door

Dr. L.J.L.Deij

1. Inleiding

In het kader van de werkzaamheden van de Werkgroep Hydrologisch Onderzoek Overijssel werd, zeer algemeen, de vraag gesteld hoe vaak na een droog tijdvak in de herfst en winter een droog tijdvak in de lente en zomer volgt. Deze vraag, die op zichzelf nog weinig gepreciseerd is, werd gesteld tegen de achtergrond van een hydrologische vraag nl. naar een zo doelmatig mogelijk beheer van eventueel te stichten waterreservoirs ten behoeve van de toekomstige watervoorziening. Teneinde de gestelde vraag verder te preciseren en deze te beantwoorden met behulp van direct beschikbare gegevens van de neerslag, en eventueel ook van de verdamping, leek het voor een eerste oriëntatie het eenvoudigst als tijdvakken te kiezen de seizoenen en als waarnemingsreeks die van de neerslag van Zwanenburg-Hoofddorp, gepubliceerd door Dr. A.Labrijn in Mededelingen en Verhandelingen 49: "Het klimaat van Nederland gedurende de laatste twee en een halve eeuw", van 1735 af en uitgebreid met de overeenkomstige gegevens van Hoofddorp tot en met het jaar 1971.

Hoewel de bovengenoemde vraag en de hydrologische achtergrond daarvan ingegeven zijn door het hydrologisch onderzoek dat thans in Salland plaatsvindt, zodat het voor de hand gelegen zou hebben in dit geval van gegevens uit het gebied zelf gebruik te maken, werd niettemin de voorkeur gegeven aan de beschouwing van een zo lang mogelijke reeks van waarnemingen, nl. die van Zwanenburg-Hoofddorp. Het neerslagregime in het westelijk deel van ons land zal stellig niet geheel overeenstemmen met dat in het oostelijk deel. Waar wij echter voornamelijk het oog richten op langdurige droge tijdvakken, hier van 3 of 6 maanden lengte, kan men veilig aannemen dat deze in voorkomende gevallen doorgaans vrijwel het gehele land zullen betreffen.

Het zou zeker de voorkeur hebben verdiend onze beschouwing niet uitsluitend te baseren op de neerslag, maar liever op het neerslagoverschot, d.i. het verschil neerslag minus verdamping, omdat dit verschil aan het bodemwater en de waterafvoer ten goede komt. Voor Zwanenburg-Hoofddorp staan de gegevens van verdamping echter niet ter beschikking. Om toch enige indruk van het gedrag van het neerslagoverschot voor de seizoenen te leren kennen werden de aanwezige gegevens van De Bilt over het tijdvak 1911-1960 gebruikt.

Men kan zich afvragen tot welke grenswaarde van de hoeveelheid neerslag, resp. neerslagoverschot, over een tijdvak van één of twee seizoenen dit tijdvak als "droog" moet worden aangemerkt. Tot op zekere hoogte is deze grenswaarde willekeurig. Men kan er een bepaald vast bedrag voor kiezen, men kan haar ook koppelen aan de standaardafwijking van de frequentieverdeling der hoeveelheden, men kan haar ook laten afhangen van de meer specifieke hydrologische doelstelling. Voorlopig willen we de keuze nog open laten.

2. De neerslagwaarnemingen van Zwanenburg-Hoofddorp (1735-1971)
- 2.1. De neerslagsommen van de wintertijdvakken en van de daaropvolgende lentetijdvakken
- 2.1.1. Het stippendiagram

De neerslagsommen (in mm) van de wintertijdvakken (december + januari + februari) en van de daaropvolgende lentetijdvakken (maart + april + mei) werden als abscis x_w en bijbehorende ordinaat y_l voor de jaren 1735 tot en met 1971 in een stippendiagram afgebeeld (fig. 1). In totaal bevat de stippenwolk 237 punten. Voor het gegeven tijdvak 1735-1971 gelden als uiterste waarden van x_w en y_l resp.: $41 \leq x_w \leq 331$ en $44 \leq y_l \leq 235$. De punten van de stippenwolk liggen onregelmatig verspreid en op het gezicht beoordeeld lijkt het er niet op dat tussen de waarden van y_l en x_w een correlatief verband bestaat. Teneinde de afwezigheid van een correlatie te toetsen werden de frequenties van de combinaties (x_w, y_l) binnen gegeven klassegrenzen van x_w en y_l geteld en in een contingentie-tabel samengevat (tabel 1). Als klassegrenzen van x_w werden gekozen de waarden 100, 150 en 200 mm en van y_l de waarden 100 en 150 mm.

Tabel 1

w \ 1	≤ 100	100-150	≥ 150	som
≤ 100	10 (9,2)	19 (17,6)	8 (10,3)	37
100-150	$19\frac{1}{2}$ (18,6)	$35\frac{1}{2}$ (35,4)	$19\frac{1}{2}$ (20,5)	$74\frac{1}{2}$
150-200	$14\frac{1}{2}$ (17,3)	42 (36,8)	21 (21,4)	$77\frac{1}{2}$
≥ 200	15 (12,0)	16 (22,8)	17 (13,3)	48
som	59	$112\frac{1}{2}$	$65\frac{1}{2}$	237

Noemen we de elementen van de frequentiematrix n_{ij} ($i = 1..4$, $j = 1..3$), de som van de elementen in de j^{de} kolom $n_{.j}$ en die van de elementen in de i^{de} rij $n_{i.}$. We toetsen nu de hypothese dat de verdeling der n_{ij} in een willekeurige kolom dezelfde is, onafhankelijk van de verdeling der n_{ij} in een willekeurige rij. We berekenen de frequenties $n_{ij}^{\#}$, welke bij onafhankelijkheid der elementen zou bestaan, uit de randtotalen aldus:

$$n_{ij}^{\#} = \frac{n_{i.} \cdot n_{.j}}{n} \quad (1)$$

De uitkomsten zijn in de tabel tussen () geplaatst.

Vervolgens bepalen we de toetsingsgrootte u volgens:

$$u = \sum_{i=1}^3 \sum_{j=1}^4 \frac{(n_{ij} - n_{ij}^{\#})^2}{n_{ij}^{\#}} \quad (2)$$

Deze grootte volgt een χ^2 -verdeling met $(3-1)(4-1) = 6$ graden van vrijheid. De berekening van u volgens (2) levert $u = 6.52$. Deze waarde blijft beneden de grenswaarde 12.59 van het 95% aanvaardings-traject van de χ^2 -verdeling met 6 graden van vrijheid. De hypothese, dat de neerslagsommen van de lentetijdvakken niet gecorreleerd zijn met die van de voorafgaande wintertijdvakken, behoeft derhalve niet te worden verworpen.

2.1.2. De distributieve en cumulatieve verdeling der neerslagsommen van de wintertijdvakken (x_w) en van de lentetijdvakken (y_l).
Gemiddelden en standaarddeviaties.

Met gebruikmaking van het stipendiagram (fig. 1) werden de frequenties f_w en f_l van de grootheden x_w en y_l in klassen van 10 mm breedte bepaald. De uitkomsten zijn in tabel 2 samengevat. De eerste kolom bevat steeds het klassemidden $x_w^{\#}$ en $y_l^{\#}$ in mm. Uiteraard geldt $\sum f_w = \sum f_l = 237$.

Tabel 2

$x_w^{\#}, y_l^{\#}$	f_w	f_l	$x_w^{\#}, y_l^{\#}$	f_w	f_l	$x_w^{\#}, y_l^{\#}$	f_w	f_l
45	2	2	145	$23\frac{1}{2}$	22	245	3	1
55	1	3	155	$13\frac{1}{2}$	$12\frac{1}{2}$	255	$3\frac{1}{2}$	2
65	$4\frac{1}{2}$	$5\frac{1}{2}$	165	$16\frac{1}{2}$	15	265	3	0
75	10	14	175	24	10	275	1	0
85	$11\frac{1}{2}$	$15\frac{1}{2}$	185	$15\frac{1}{2}$	$10\frac{1}{2}$	285	$2\frac{1}{2}$	1
95	8	19	195	8	$7\frac{1}{2}$	295	$2\frac{1}{2}$	
105	$6\frac{1}{2}$	$29\frac{1}{2}$	205	$14\frac{1}{2}$	$1\frac{1}{2}$	305	0	
115	$11\frac{1}{2}$	15	215	7	$2\frac{1}{2}$	315	0	
125	16	23	225	$7\frac{1}{2}$	0	325	0	
135	17	23	235	$2\frac{1}{2}$	2	335	1	

Op grond van de verkregen uitkomsten werden de grootheden x_w en y_l resp. de gemiddelde waarden $\hat{\mu}_w$ en $\hat{\mu}_l$ berekend, alsmede de standaarddeviaties $\hat{\sigma}_w$ en $\hat{\sigma}_l$. De resultaten zijn in tabel 3 samengevat:

Tabel 3

$\hat{\mu}_w = 157,4 \text{ mm}$	$\hat{\mu}_l = 129,5 \text{ mm}$
$\hat{\sigma}_w = 53,0 \text{ mm}$	$\hat{\sigma}_l = 41,5 \text{ mm}$

Zoals uit het vervolg zal blijken zijn dit tevens de maximale aan-nemelijkheidsschatters van de gemiddelden (μ) en de standaarddeviaties (σ) van de (normaal verdeelde) populaties van x_w en y_l .

Vervolgens zijn de cumulatieve frequenties F_w en F_1 bepaald voor de gekozen klassengrenzen. Deling van F_w en F_1 door $(237 + 1)$ geeft tevens de geschatte onderschrijdingskans voor de gekozen klassengrenshoeveelheid x_w en y_1 . De uitkomsten zijn in tabel 4 (zie bladzijde 6) samengevat.

Met behulp van de aldus verkregen uitkomsten is op lineair waarschijnlijkheidspapier een grafiek samengesteld van $\frac{F_w}{238}$ als functie van x_w en van $\frac{F_1}{238}$ als functie van y_1 (fig.2). Tevens zijn in deze figuur de rechten getrokken welke de cumulatieve normale verdelingen weergeven van x_w en y_1 , resp. gebaseerd op de gevonden waarden $(\hat{\mu}_w, \hat{\sigma}_w)$ en $(\hat{\mu}_1, \hat{\sigma}_1)$ van tabel 3. De punten passen redelijk goed bij deze rechten aan, met andere woorden de verdelingsfuncties van x_w en y_1 kunnen bij benadering als normaal worden beschouwd. De cumulatieve kansverdelingen van x_w en y_1 kunnen dan worden voorgesteld door:

$$\int_{-\infty}^{x_w} \frac{1}{\sqrt{2\pi} \hat{\sigma}_w} e^{-\frac{1}{2} \left(\frac{\xi - \hat{\mu}_w}{\hat{\sigma}_w} \right)^2} d\xi = \int_{-\infty}^{\frac{x_w - \hat{\mu}_w}{\hat{\sigma}_w}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} t^2} dt = \Phi \left(\frac{x_w - \hat{\mu}_w}{\hat{\sigma}_w} \right) \quad (3)$$

$$\int_{-\infty}^{y_1} \frac{1}{\sqrt{2\pi} \hat{\sigma}_1} e^{-\frac{1}{2} \left(\frac{\eta - \hat{\mu}_1}{\hat{\sigma}_1} \right)^2} d\eta = \int_{-\infty}^{\frac{y_1 - \hat{\mu}_1}{\hat{\sigma}_1}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} t^2} dt = \Phi \left(\frac{y_1 - \hat{\mu}_1}{\hat{\sigma}_1} \right) \quad (4)$$

waarin $\Phi(u)$ de cumulatieve standaard normale verdeling voorstelt.

2.1.3. Betrouwbaarheidsintervallen

Voor een normale verdeling, gekarakteriseerd door haar gemiddelde μ en standaarddeviatie σ , wordt de cumulatieve kansfunctie gegeven door:

Tabel 4

x_w, y_l	F_w	$\frac{F_w}{238}$	F_l	$\frac{F_l}{238}$	x_w, y_l	F_w	$\frac{F_w}{238}$	F_l	$\frac{F_l}{238}$	x_w, y_l	F_w	$\frac{F_w}{238}$	F_l	$\frac{F_l}{238}$
50	2	0,008	2	0,008	150	111½	0,469	171½	0,720	250	223½	0,939	234	0,983
60	3	0,013	5	0,013	160	125	0,526	184	0,772	260	227	0,954	236	0,992
70	7½	0,032	10½	0,044	170	141½	0,595	199	0,836	270	230	0,967	236	0,992
80	17½	0,074	24½	0,103	180	165½	0,695	209	0,878	280	231	0,971	236	0,992
90	29	0,122	40	0,168	190	181	0,761	219½	0,922	290	233½	0,981	237	0,996
100	37	0,155	59	0,248	200	189	0,795	227	0,954	300	236	0,992		
110	43½	0,183	88½	0,372	210	203½	0,855	228½	0,960	310	236	0,992		
120	55	0,231	103½	0,435	220	210½	0,884	231	0,971	320	236	0,992		
130	71	0,298	126½	0,531	230	218	0,916	231	0,971	330	236	0,992		
140	88	0,369	149½	0,628	240	220½	0,927	233	0,979	340	237	0,996		

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left(\frac{\xi - \mu}{\sigma} \right)^2} d\xi = \int_{-\infty}^{\frac{x - \mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} t^2} dt = \Phi \left(\frac{x - \mu}{\sigma} \right) \quad (5)$$

We beschouwen nu de waarde van x waarvoor $F(x)$ een voorgeschreven waarde γ_x aanneemt ($0 \leq \gamma_x \leq 1$). Zij $\frac{x - \mu}{\sigma} = k$, zodat $x = \mu + k\sigma$. Is nu $\hat{\mu}$ de maximale aannemelijkheidsschatter van μ en $\hat{\sigma}$ die van σ van een steekproef ter grootte n uit de normale verdeling (μ, σ) , dan is ook $\hat{x} \equiv \hat{\mu} + k\hat{\sigma}$ de maximale aannemelijkheidsschatter van x in (5) voor de voorgeschreven waarde γ_x .

Trekt men bij herhaling steekproeven ter grootte n uit de normale verdeling (μ, σ) , dan kan men aantonen, zo n voldoende groot is, dat de steekproefwaarden $\hat{\mu}$ en $\hat{\sigma}$ een simultane binormale kansverdeling $g(\hat{\mu}, \hat{\sigma})$ volgen, welke geschreven kan worden als een product:

$$g(\hat{\mu}, \hat{\sigma}) = g_1(\hat{\mu}) \cdot g_2(\hat{\sigma}) \quad (6)$$

waarin

$$g_1(\hat{\mu}) = \frac{1}{\sqrt{2\pi} \frac{\sigma}{\sqrt{n}}} e^{-\frac{1}{2} \left(\frac{\hat{\mu} - \mu}{\frac{\sigma}{\sqrt{n}}} \right)^2} \quad (7)$$

en

$$g_2(\hat{\sigma}) = \frac{1}{\sqrt{2\pi} \frac{\sigma}{\sqrt{2n}}} e^{-\frac{1}{2} \left(\frac{\hat{\sigma} - \sigma}{\frac{\sigma}{\sqrt{2n}}} \right)^2} \quad (8)$$

De grote steekproefverdeling van $\hat{\mu}$ heeft dus als gemiddelde μ en als standaarddeviatie $\frac{\sigma}{\sqrt{n}}$, en die van $\hat{\sigma}$ heeft als gemiddelde σ en als standaarddeviatie $\frac{\sigma}{\sqrt{2n}}$.

Beschouwt men nu de grote steekproefverdeling van $\hat{x} = \hat{\mu} + k\hat{\sigma}$, dan volgt deze lineaire combinatie van $\hat{\mu}$ en $\hat{\sigma}$ eveneens een normale verdeling, waarvan het gemiddelde wordt gegeven door $\mu + k\sigma$ en de standaarddeviatie door

$$\sqrt{\frac{\sigma^2}{n} + k^2 \frac{\sigma^2}{2n}} = \sigma \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2} \right)}. \text{ De verdeling van } \hat{x} \text{ wordt dus}$$

voorgesteld door:

$$h(\hat{x}) = \frac{1}{\sqrt{2\pi} \cdot \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)} \sigma} e^{-\frac{1}{2} \frac{\{\hat{x} - (\mu + k\sigma)\}^2}{\frac{1}{n} \left(1 + \frac{k^2}{2}\right) \sigma^2}} \quad (9)$$

De maximale aannemelijkheidsschatting van de standaardafwijking van \hat{x} is dan blijkbaar bepaald door:

$$\hat{\sigma}_{\hat{x}} = \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)} \cdot \hat{\sigma} \quad (10)$$

Men kan nu een betrouwbaarheidsinterval voor x aangeven volgens:

$$P\left(\hat{x} - d_{\gamma} \cdot \hat{\sigma}_{\hat{x}} < x < \hat{x} + d_{\gamma} \cdot \hat{\sigma}_{\hat{x}}\right) = \gamma \quad (11)$$

waarbij:

$$\gamma = \frac{1}{\sqrt{2\pi}} \int_{-d_{\gamma}}^{+d_{\gamma}} e^{-\frac{1}{2}t^2} dt \quad (12)$$

of:

$$P\left(\hat{\mu} + k\hat{\sigma} - d_{\gamma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)} \hat{\sigma} < \mu + k\sigma < \hat{\mu} + k\hat{\sigma} + d_{\gamma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)} \hat{\sigma}\right) = \gamma \quad (13)$$

Deze betrekking geldt voor iedere waarde van k tussen $-\infty$ en $+\infty$. Voor $\gamma = 0.95$ is $d_{\gamma} = 1.96$. In fig. 2 kan men nu voor ieder punt van de kansrechte van de neerslagsommen van de wintertijdvakken, welke is gebaseerd op de waarden van $\hat{\mu}_w$ en $\hat{\sigma}_w$ en de bij dat punt behorende waarde van k het 95% betrouwbaarheidsinterval bepalen, dat wordt gegeven door:

$$\pm 1.96 \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)} \hat{\sigma}_w \quad (14)$$

Blijkbaar blijft het merendeel van de steekproefpunten binnen de door stippellijnen aangegeven 95% betrouwbaarheidsband. Alleen voor zeer lage, resp. zeer hoge waarden van x_w komen enkele afwijkingen voor. Hetzelfde kan worden gezegd in het geval van de neerslagsommen van de lentetijdvakken. Hiermede wordt nog eens bevestigd dat zowel x_w als y_1 bij redelijk goede benadering als normaal verdeeld kunnen worden beschouwd.

2.1.4. Kansen op neerslaghoeveelheden in opvolgende winters en lentes

Uit de beschouwingen in par. 2.1.1. en par. 2.1.2. blijkt dat de neerslaghoeveelheden van de winters en die van de daaropvolgende lentes een bivariabele normale kansverdeling volgen die kan worden voorgesteld door:

$$f(x_w, y_1) dx_w dy_1 = f_1(x_w) dx_w \cdot f_2(y_1) dy_1 = \frac{1}{\sqrt{2\pi} \hat{\sigma}_w} e^{-\frac{1}{2} \left(\frac{x_w - \hat{\mu}_w}{\hat{\sigma}_w} \right)^2} dx_w \cdot \frac{1}{\sqrt{2\pi} \hat{\sigma}_1} e^{-\frac{1}{2} \left(\frac{y_1 - \hat{\mu}_1}{\hat{\sigma}_1} \right)^2} dy_1 \quad (15)$$

Hieruit volgt dat de kans op een winter met ten hoogste x_w mm neerslag, gevolgd door een lente met ten hoogste y_1 mm neerslag wordt gegeven door:

$$F(x_w, y_1) = \int_{-\infty}^{x_w} \frac{1}{\sqrt{2\pi} \hat{\sigma}_w} e^{-\frac{1}{2} \left(\frac{\xi - \hat{\mu}_w}{\hat{\sigma}_w} \right)^2} d\xi \cdot \int_{-\infty}^{y_1} \frac{1}{\sqrt{2\pi} \hat{\sigma}_1} e^{-\frac{1}{2} \left(\frac{\eta - \hat{\mu}_1}{\hat{\sigma}_1} \right)^2} d\eta = \Phi \left(\frac{x_w - \hat{\mu}_w}{\hat{\sigma}_w} \right) \cdot \Phi \left(\frac{y_1 - \hat{\mu}_1}{\hat{\sigma}_1} \right) \quad (16)$$

Geven we aan $F(x_w, y_1)$ een voorgeschreven waarde P dan wordt in het (x_w, y_1) -vlak een relatie $F(x_w, y_1) = P$ tussen x_w en y_1 vastgelegd, welke door een kromme kan worden afgebeeld. Zulks is in fig. 1 geschied voor $P = 0.005, 0.01, 0.02, 0.05, 0.10, 0.20$ en 0.30 . Deze kanswaarden corresponderen achtereenvolgens met de gemiddelde herhalingsstijden $T = 200, 100, 50, 20, 10$ en 3.3 jaar.

Geheel algemeen kan de kans op een winter met ten minste x'_w en ten hoogste x''_w mm neerslag, gevolgd door een lente met tenminste y'_1 en ten hoogste y''_1 mm neerslag gemakkelijk worden berekend volgens de formule:

$$\Phi(x'_w, x''_w; y'_1, y''_1) = \int_{x'_w}^{x''_w} \frac{1}{\sqrt{2\pi} \hat{\sigma}_w} e^{-\frac{1}{2} \left(\frac{\xi - \hat{\mu}_w}{\hat{\sigma}_w} \right)^2} d\xi \cdot \int_{y'_1}^{y''_1} \frac{1}{\sqrt{2\pi} \hat{\sigma}_1} e^{-\frac{1}{2} \left(\frac{\eta - \hat{\mu}_1}{\hat{\sigma}_1} \right)^2} d\eta =$$

$$\left\{ \Phi\left(\frac{x_w'' - \hat{\mu}_w}{\hat{\sigma}_w}\right) - \Phi\left(\frac{x_w' - \hat{\mu}_w}{\hat{\sigma}_w}\right) \right\} \left\{ \Phi\left(\frac{y_1'' - \hat{\mu}_1}{\hat{\sigma}_1}\right) - \Phi\left(\frac{y_1' - \hat{\mu}_1}{\hat{\sigma}_1}\right) \right\} \quad (17)$$

Vatten we bijv. een droge winter op als een winter waarin ten hoogste 100 mm valt en evenzo een droge lente als een lente waarin ten hoogste 100 mm valt, dan vinden we met de waarden van $\hat{\mu}_w$, $\hat{\sigma}_w$, $\hat{\mu}_1$, $\hat{\sigma}_1$ uit tabel 3 en $x_w = y_1 = 100$ mm met toepassing van (16):

$$F(x_w, x_1) = \Phi\left(\frac{100-157,4}{53,0}\right) \cdot \Phi\left(\frac{100-129,5}{41,5}\right) =$$

$$= 0,1394 \times 0,2387 = 0,0333$$

De simultane kans op deze winter-lente combinatie is dus 3,33%, of ander uitgedrukt: deze situatie doet zich 1 maal in gemiddeld 30 jaren voor. Uit fig. 1 zien we dat zich 10 gevallen op de 237 hebben voorgedaan waarvoor $x_w \leq 100$ en $y_1 \leq 100$ mm waren, d.i. de combinatie (x_w, y_1) die aan deze voorwaarden voldoet trad met een frequentie van $\frac{10}{237} = 0,0423$ op of met een gemiddelde herhalings-tijd van ongeveer $\frac{237}{10}$ jaren. Met een kans van 3,33% op 237 gevallen zouden zich ca. 8 gevallen hebben mogen voordoen. In werkelijkheid hebben zich 10 gevallen voorgedaan, wat nog in redelijke overeenstemming is met de berekening wanneer de betrouwbaarheid der uitkomsten in aanmerking wordt genomen.

2.2. De neerslagsommen van het (herfst + winter)-tijdvak en van het daarop volgende (lente + zomer)-tijdvak

2.2.1. Het stippendiagram

De neerslagsommen (in mm) van de (herfst + winter)-tijdvakken (september, oktober, november, december, januari, februari) en van de daarop volgende (lente + zomer)-tijdvakken (maart, april, mei, juni, juli, augustus) werden als abscis x_{HW} en bijbehorende ordinaat y_{LZ} voor de jaren 1735 tot en met 1971 in een stippendiagram afgebeeld (fig. 3). In totaal bevat de stippenwolk 236 punten. Voor het gegeven tijdvak gelden als uiterste waarden van x_{HW} en y_{LZ} resp. $201 \leq x_{HW} \leq 763$ en $181 \leq y_{LZ} \leq 523$. Ook nu liggen de punten van de stippenwolk onregelmatig verspreid. Teneinde de al of niet aanwezigheid van een correlatie tussen de waarden van y_{LZ} en x_{HW} te

onderzoeken werden op overeenkomstige wijze als in par.2.1.1. beschreven de frequenties van de combinaties (x_{HW}, y_{LZ}) binnen gegeven klassegrenzen van x_{HW} en y_{LZ} geteld en in een contingentietabel samengevat (tabel 5). Als klassegrenzen van x_{HW} werden gekozen de waarden 300, 400 en 500 mm en van y_{LZ} de waarden 300 en 400 mm.

Tabel 5

H+W \ L+Z	≤ 300	300-400	≥ 400	som
≤ 300	$11\frac{1}{2}$ (9,6)	11 (11,7)	8 (9,2)	$30\frac{1}{2}$
300-400	$28\frac{1}{2}$ (28,2)	37 (34,5)	$24\frac{1}{2}$ (27,3)	90
400-500	30 (27,4)	$28\frac{1}{2}$ (33,6)	29 (26,5)	$87\frac{1}{2}$
≥ 500	4 (8,8)	14 (10,7)	10 (8,5)	28
som	74	$90\frac{1}{2}$	$71\frac{1}{2}$	236

De getallen tussen haakjes in de frequentie matrix stellen de frequenties n_{ij}^* voor berekend volgens (1), welke zouden gelden indien de verdeling der n_{ij} in een willekeurige kolom dezelfde is, onafhankelijk van de verdeling der n_{ij} in een willekeurige rij. De toetsingsgrootte u , berekend volgens (2), levert nu de waarde $u = 6,21$. Deze waarde is kleiner dan de grenswaarde 12,59 van het met 95% kans aanvaardingstraject van de door u gevolgde χ^2 -verdeling met 6 graden van vrijheid. De hypothese dat de neerslagsommen der (lente + zomer)-tijdvakken niet gecorreleerd zijn met die van de voorafgaande (herfst + winter)-tijdvakken behoeft derhalve niet te worden verworpen.

2.2.2. De distributieve en de cumulatieve verdeling der neerslagsommen van de (herfst + winter)-tijdvakken (x_{HW}) en van de (lente + zomer)-tijdvakken (y_{LZ}).

Gemiddelden en standaarddeviaties

Met gebruikmaking van het stippendiagram (fig. 3) werden de frequenties f_{HW} en f_{LZ} van de grootheden x_{HW} en y_{LZ} in klassen van 10 mm breedte bepaald. De uitkomsten zijn in tabel 6 samengevat.

De eerste kolom bevat steeds het klassemidden x_{HW}^* resp. y_{LZ}^* in mm. Uiteraard geldt $\sum f_{HW} = \sum f_{LZ} = 236$.

Tabel 6

x_{HW}^*, y_{LZ}^*	f_{HW}	f_{LZ}	x_{HW}^*, y_{LZ}^*	f_{HW}	f_{LZ}	x_{HW}^*, y_{LZ}^*	f_{HW}	f_{LZ}
185		3	355	7	8	525	3	2
195	0	1	365	$13\frac{1}{2}$	$7\frac{1}{2}$	535	2	0
205	4	2	375	$6\frac{1}{2}$	9	545	$4\frac{1}{2}$	
215	1	7	385	$10\frac{1}{2}$	$8\frac{1}{2}$	555	$\frac{1}{2}$	
225	$3\frac{1}{2}$	$1\frac{1}{2}$	395	10	$3\frac{1}{2}$	565	5	
235	$2\frac{1}{2}$	4	405	7	$12\frac{1}{2}$	575	3	
245	2	9	415	$11\frac{1}{2}$	7	585	0	
255	4	7	425	8	$9\frac{1}{2}$	595	1	
265	3	$8\frac{1}{2}$	435	9	$7\frac{1}{2}$	605	0	
275	1	$7\frac{1}{2}$	445	9	7	615	1	
285	3	$9\frac{1}{2}$	455	$7\frac{1}{2}$	$8\frac{1}{2}$	625	$2\frac{1}{2}$	
295	$6\frac{1}{2}$	14	465	$11\frac{1}{2}$	$4\frac{1}{2}$	635	$\frac{1}{2}$	
305	4	$11\frac{1}{2}$	475	$7\frac{1}{2}$	$4\frac{1}{2}$	645	0	
315	11	8	485	$7\frac{1}{2}$	$\frac{1}{2}$	---		
325	$11\frac{1}{2}$	$9\frac{1}{2}$	495	9	2	---		
335	8	13	505	3	2	---		
345	3	12	515	1	4	765	1	

Op grond van de verkregen uitkomsten werden de grootheden x_{HW} en y_{LZ} resp. de gemiddelde waarden $\hat{\mu}_{HW}$ en $\hat{\mu}_{LZ}$ berekend, alsmede de standaarddeviaties $\hat{\sigma}_{HW}$ en $\hat{\sigma}_{LZ}$. De resultaten zijn in tabel 7 samengevat.

Tabel 7

$\hat{\mu}_{HW} = 400,9 \text{ mm}$	$\hat{\mu}_{LZ} = 348,0 \text{ mm}$
$\hat{\sigma}_{HW} = 93,9 \text{ mm}$	$\hat{\sigma}_{LZ} = 79,8 \text{ mm}$

Vervolgens zijn de cumulatieve frequenties F_{HW} en F_{LZ} bepaald voor de gekozen klassegrenzen. Deling van F_{HW} en F_{LZ} door $(236+1)$ geeft tevens de geschatte onderschrijdingskans voor de gekozen klassegrenshoeveelheid x_{HW} en y_{LZ} . De uitkomsten zijn in tabel 8 (zie bladzijde 14) samengevat.

Met behulp van de aldus verkregen uitkomsten is wederom op lineair waarschijnlijkheidspapier een grafiek samengesteld van $\frac{F_{HW}}{237}$ als functie van x_{HW} en van $\frac{F_{LZ}}{237}$ als functie van y_{LZ} (fig. 4). Tevens zijn in deze figuur de rechten getrokken welke de cumulatieve normale verdelingen weergeven van x_{HW} en y_{LZ} , resp. gebaseerd op de gevonden waarden $(\hat{\mu}_{HW}, \hat{\sigma}_{HW})$ en $(\hat{\mu}_{LZ}, \hat{\sigma}_{LZ})$ van tabel 7. De punten passen redelijk goed bij deze rechten aan, met andere woorden de verdelingsfuncties van x_{HW} en y_{LZ} kunnen bij benadering als normaal worden beschouwd. De cumulatieve kansverdelingen van x_{HW} en y_{LZ} kunnen dan worden voorgesteld door:

$$\int_{-\infty}^{x_{WL}} \frac{1}{\sqrt{2\pi} \hat{\sigma}_{HW}} e^{-\frac{1}{2} \left(\frac{\xi - \hat{\mu}_{HW}}{\hat{\sigma}_{HW}} \right)^2} d\xi = \Phi \left(\frac{x_{HW} - \hat{\mu}_{HW}}{\hat{\sigma}_{HW}} \right) \quad (18)$$

$$\int_{-\infty}^{y_{LZ}} \frac{1}{\sqrt{2\pi} \hat{\sigma}_{LZ}} e^{-\frac{1}{2} \left(\frac{\eta - \hat{\mu}_{LZ}}{\hat{\sigma}_{LZ}} \right)^2} d\eta = \Phi \left(\frac{y_{LZ} - \hat{\mu}_{LZ}}{\hat{\sigma}_{LZ}} \right) \quad (19)$$

waarin $\Phi(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$ de cumulatieve standaard normale verdeling voorstelt.

2.2.3. Betrouwbaarheidsintervallen

Geheel overeenkomstig de methode welke in par. 2.1.3. werd uiteengezet is ook in fig. 4, met toepassing van (14), voor ieder punt van de kansrechte van de neerslagsommen van de (herfst + winter)-tijdvakken het 95% betrouwbaarheidsinterval bepaald. Blijkbaar blijft het merendeel van de steekproefpunten binnen de door de stippellijnen aangegeven 95% betrouwbaarheidsband.

Tabel 8

x_{HW}, y_{LZ}	F_{HW}	$\frac{F_{HW}}{237}$	F_{LZ}	$\frac{F_{LZ}}{237}$	x_{HW}, y_{LZ}	F_{HW}	$\frac{F_{HW}}{237}$	F_{LZ}	$\frac{F_{LZ}}{237}$	x_{HW}, y_{LZ}	F_{HW}	$\frac{F_{HW}}{237}$	F_{LZ}	$\frac{F_{LZ}}{237}$
210	4	0,017	3	0,013	370	93½	0,395	128	0,540	530	215	0,907	230	0,970
220	5	0,021	4	0,017	380	100	0,422	136	0,574	540	217	0,916	234	0,987
230	8½	0,036	6	0,025	390	110½	0,466	143½	0,605	550	221½	0,935	236	0,996
240	11	0,046	13	0,055	400	120½	0,508	152½	0,643	560	222	0,937		
250	13	0,055	14½	0,061	410	127½	0,538	161	0,679	570	227	0,958		
260	17	0,072	18½	0,078	420	139	0,586	164½	0,694	580	230	0,970		
270	20	0,084	27½	0,116	430	147	0,620	177	0,747	590	230	0,970		
280	21	0,089	34½	0,146	440	156	0,658	184	0,776	600	231	0,975		
290	24	0,101	43	0,181	450	165	0,696	193½	0,816	610	231	0,975		
300	30½	0,129	50½	0,213	460	172½	0,728	201	0,848	620	232	0,979		
310	34½	0,146	60	0,253	470	184	0,776	208	0,878	630	234½	0,989		
320	45½	0,192	74	0,312	480	191½	0,808	216½	0,914	640	235	0,992		
330	57	0,241	85½	0,361	490	199	0,840	221	0,932	650	235	0,992		
340	65	0,274	93½	0,395	500	208	0,878	225½	0,951	660	235	0,992		
350	73	0,308	103	0,435	510	211	0,890	226	0,954	670	235	0,992		
360	80	0,338	116	0,489	520	212	0,895	228	0,962	770	236	0,996		

Hetzelfde kan worden gezegd in het geval van de neerslagsommen in de (lente + zomer)-tijdvakken.

Hiermede wordt nog eens bevestigd, dat zowel x_{HW} en y_{LZ} bij redelijk goede benadering als normaal verdeeld kunnen worden beschouwd.

2.2.4. Kansen op neerslaghoeveelheden in opvolgende (herfst + winter)- en (lente + zomer)-tijdvakken

Uit de beschouwingen in par. 2.2.1. en par. 2.2.2. blijkt geheel analoog aan par. 2.1.4. dat de neerslaghoeveelheden x_{HW} en y_{LZ} een bivariabele normale verdeling volgen welke kan worden voorgesteld door:

$$f(x_{HW}, y_{LZ}) dx_{HW} dy_{LZ} = f_1(x_{HW}) dx_{HW} \cdot f_2(y_{LZ}) dy_{LZ} = \frac{1}{\sqrt{2\pi} \hat{\sigma}_{HW}} e^{-\frac{1}{2} \left(\frac{x_{HW} - \hat{\mu}_{HW}}{\hat{\sigma}_{HW}} \right)^2} dx_{HW} \cdot \frac{1}{\sqrt{2\pi} \hat{\sigma}_{LZ}} e^{-\frac{1}{2} \left(\frac{y_{LZ} - \hat{\mu}_{LZ}}{\hat{\sigma}_{LZ}} \right)^2} dy_{LZ} \quad (20)$$

Hieruit volgt wederom dat de kans op een (herfst + winter)-tijdvak met ten hoogste x_{HW} mm neerslag, gevolgd door een (lente + zomer)-tijdvak met ten hoogste y_{LZ} mm neerslag gegeven wordt door:

$$F(x_{HW}, y_{LZ}) = \int_{-\infty}^{x_{HW}} \frac{1}{\sqrt{2\pi} \hat{\sigma}_{HW}} e^{-\frac{1}{2} \left(\frac{\xi - \hat{\mu}_{HW}}{\hat{\sigma}_{HW}} \right)^2} d\xi \cdot \int_{-\infty}^{y_{LZ}} \frac{1}{\sqrt{2\pi} \hat{\sigma}_{LZ}} e^{-\frac{1}{2} \left(\frac{\eta - \hat{\mu}_{LZ}}{\hat{\sigma}_{LZ}} \right)^2} d\eta = \Phi \left(\frac{x_{HW} - \hat{\mu}_{HW}}{\hat{\sigma}_{HW}} \right) \cdot \Phi \left(\frac{y_{LZ} - \hat{\mu}_{LZ}}{\hat{\sigma}_{LZ}} \right) \quad (21)$$

Geven we aan $F(x_{HW}, y_{LZ})$ voorgeschreven waarden $P = 0,005, 0,01, 0,02, 0,05, 0,10, 0,20, 0,30$ dan worden met deze kanswaarden als parameter krommen $F(x_{HW}, y_{LZ}) = P$ vastgelegd welke in fig. 3 grafisch zijn voorgesteld.

Vatten we bijv. een droge (herfst + winter) op als een tijdvak waarin ten hoogste 300 mm valt en evenzo een droge (lente + zomer) als een tijdvak waarin ten hoogste 300 mm valt, dan vinden we met de waarden van $\hat{\mu}_{HW}, \hat{\sigma}_{HW}, \hat{\mu}_{LZ}, \hat{\sigma}_{LZ}$ uit tabel 7 en $x_{HW} = y_{LZ} = 300$ mm met toepassing van (21):

$$F(x_{HW}, y_{LZ}) = \prod \left(\frac{300-400,9}{93,9} \right) \cdot \prod \left(\frac{300-343,0}{79,8} \right)$$

$$= 0,1234 \times 0,2709 = 0,0335$$

De simultane kans op deze (x_{HW}, y_{LZ}) -combinatie is dus 3,35%, hetgeen overeenstemt met een herhalings-tijd van gemiddeld 30 jaren. Theoretisch zouden zich op de 236 gevallen 8 gevallen kunnen voordoen die aan de gestelde voorwaarden voldoen. Uit fig. 3 blijkt dat zich in werkelijkheid 12 van dergelijke gevallen hebben voorgedaan, wat nog in redelijke overeenstemming is met de berekening, wanneer de betrouwbaarheid van de uitkomsten in aanmerking wordt genomen.

2.2.5. De tien droogste winter/lente- en de twaalf droogste (herfst+winter)/(lente+zomer)-tijdvakken

In tabel 9 zijn de tien droogste winter/lente-tijdvakken, waarvoor $x_W \leq 100$ en $y_L \leq 100$ mm geldt, bijeengebracht en tevens de twaalf droogste (herfst+winter)/(lente+zomer)-tijdvakken waarvoor $x_{HW} \leq 300$ en $y_{LZ} \leq 300$ mm geldt, gerangschikt naar opklimmende waarden van de som $(x+y)$.

Tabel 9

	x_W	y_L		x_{HW}	y_{LZ}
1928 - 1929	77	57	1920 - 1921	230	181
1741 - 1742	75	61	1948 - 1949	227	218
1857 - 1858	68	87	1743 - 1744	200	247
1812 - 1813	76	79	1931 - 1932	201	266
1853 - 1854	70	90	1933 - 1934	236	231
1879 - 1880	82	78	1802 - 1803	202	295
1778 - 1779	91	70	1805 - 1806	291	250
1804 - 1805	87	79	1741 - 1742	269	281
1846 - 1847	88	78	1793 - 1794	269	292
1831 - 1832	83	95	1846 - 1847	300	277
			1834 - 1835	294	295
			1742 - 1743	296	299

Wanneer we deze gevallen naar opklimmende jaartallen zouden rangschikken en de opvolgende tijdsverschillen (herhalings-tijden) bepalen dan blijkt voor het geval (x_W, y_L) een gemiddelde herhalings-tijd

van 21 jaren te bestaan met een standaarddeviatie van 15 jaren en voor het geval $(x_{HW}y_{LZ})$ een gemiddelde herhalings-tijd van 19 jaren met een standaarddeviatie van 24 jaren. Daaruit blijkt wel dat de gevallen zeer onregelmatig over de tijd verstrooid liggen en dat hun gemiddelde herhalings-tijd weinig of geen prognostische waarde bezit.

- 3. De neerslag- en verdampingswaarnemingen van De Bilt (1911-1960)
- 3.1. De neerslagsommen van de wintertijdvakken en de daarop volgende lentetijdvakken

Er werd een stippendiagram samengesteld van de neerslagsommen (in mm) van de wintertijdvakken (x_W) en van de daarop volgende lentedtijdvakken (y_L), (fig. 5). In totaal bevat de stippenwolk 49 punten. Voor het gegeven tijdvak 1911-1960 gelden als uiterste waarden van x_W en y_L resp. $69 \leq x_W \leq 309$ en $77 \leq y_L \leq 232$.

Er werden voorts van de grootheden x_W en y_L de gemiddelde waarde en de standaardafwijking berekend. De uitkomsten zijn in tabel 10 verenigd.

Tabel 10

$\hat{\mu}_W = 185,3 \text{ mm}$	$\hat{\mu}_L = 146,3 \text{ mm}$
$\hat{\sigma}_W = 55,9 \text{ mm}$	$\hat{\sigma}_L = 43,5 \text{ mm}$

Tevens werd de correlatiecoëfficiënt tussen de bij elkaar behorende waarden van x_W en y_L bepaald. Deze bedroeg:

$$\hat{\rho}_{WL} = 0,165$$

Deze waarde, afgeleid uit een steekproef van 49 waardeparen (x_W, y_L) , bleek na toetsing, bij een onbetrouwbaarheidsdrempel $\alpha = 0,05$, niet significant van 0 te verschillen. Hieruit volgt dat de verdelingen van x_W en y_L als ongecorrleerd kunnen worden beschouwd, althans niet lineair gecorreleerd.

Door toetsing is nagegaan of de verdelingen van x_W en y_L als normaal kunnen worden beschouwd. Dit bleek inderdaad het geval te zijn.

De cumulatieve verdelingen van x_W en van y_L zijn op lineair waarschijnlijkheidspapier in beeld gebracht (fig. 6). In deze figuur zijn tevens de rechten getekend van de cumulatieve normale verdelingen, gebaseerd op de steekproefwaarden $\hat{\mu}_W$, $\hat{\sigma}_W$ en $\hat{\mu}_L$, $\hat{\sigma}_L$ uit tabel 10. De steekproefpunten slingeren om deze rechten maar blijven binnen de overeenkomstig (14) berekende 95% betrouwbaarheidsmarges.

De neerslagsommen van de wintertijdvakken x_W en de daarop volgende lentetijdvakken y_L kunnen dus bij benadering door een bivariabele normale verdeling worden voorgesteld overeenkomstig (15) en de kans op een winter met ten hoogste x_W mm neerslag, gevolgd door een lente met ten hoogste y_L mm neerslag wordt weergegeven door (16).

Geven we aan $F(x_W, y_L)$ voorgeschreven waarden $P = 0,01, 0,02, 0,05, 0,10, 0,20, 0,30$ dan worden met deze kanswaarden als parameter krommen $F(x_W, y_L) = P$ vastgelegd welke in fig. 5 grafisch zijn voorgesteld.

3.2. De neerslagsommen van de (herfst+winter)-tijdvakken en de daarop volgende (lente+zomer)-tijdvakken

Er werd een stippendiagram samengesteld van de neerslagsommen (in mm) van de (herfst+winter)-tijdvakken (x_{HW}) en van de daarop volgende (lente+zomer)-tijdvakken (y_{LZ}). (Fig. 7). In totaal bevat de stippenwolk 49 punten. Voor het gegeven tijdvak 1911-1960 gelden als uiterste waarden van x_{HW} en y_{LZ} resp.: $192 \leq x_{HW} \leq 569$ en $166 \leq y_{LZ} \leq 568$. Er werden voorts van de grootheden x_{HW} en y_{LZ} de gemiddelde waarde en de standaardafwijking berekend. De uitkomsten zijn in tabel 11 samengevat.

Tabel 11

$\hat{\mu}_{HW} = 395,9$ mm	$\hat{\mu}_{LZ} = 374,3$ mm
$\hat{\sigma}_{HW} = 96,5$ mm	$\hat{\sigma}_{LZ} = 74,0$ mm

Tevens werd de correlatiecoëfficiënt tussen de bij elkaar behorende waarden van x_{HW} en y_{LZ} bepaald. Deze bedroeg:

$$\hat{\rho}_{HZ} = + 0,346$$

Deze waarde, afgeleid uit een steekproef van 49 waardeparen (x_{HW}, y_{LZ}) bleek na toetsing bij een onbetrouwbaarheidsdrempel $\alpha = 0,05$, significant van 0 te verschillen. De verdelingen van x_{HW} en y_{LZ} kunnen dus niet als ongecorrleerd worden beschouwd. Toch willen we deze conclusie met enige reserve aanvaarden en er in par. 5 op terugkomen.

Door toetsing is verder nagegaan of de verdelingen van x_{HW} en y_{LZ} als normaal kunnen worden beschouwd. Dit bleek inderdaad het geval te zijn.

De cumulatieve verdelingen van x_{HW} en van y_{LZ} zijn op lineair waarschijnlijkheidspapier in beeld gebracht (fig. 8 en fig. 9). In deze figuren zijn tevens de rechten getekend van de cumulatieve normale verdelingen, gebaseerd op de steekproefwaarden $\hat{\mu}_{HW}$, $\hat{\sigma}_{HW}$ en $\hat{\mu}_{LZ}$, $\hat{\sigma}_{LZ}$ uit tabel 11. De steekproefpunten slingeren om deze rechten maar blijven binnen de overeenkomstig (14) berekende 95% betrouwbaarheidsmarges. De neerslagsommen van de (herfst+winter)-tijdvakken x_{HW} en de daarop volgende (lente+zomer)-tijdvakken y_{LZ} kunnen dus bij benadering door een bivariabele normale verdeling met een correlatiecoëfficiënt worden voorgesteld van de vorm (met weglating van de indices HW, LZ, HZ):

$$f(x,y)dxdy = \frac{1}{2\pi\hat{\sigma}_x\hat{\sigma}_y\sqrt{1-\hat{\rho}^2}} e^{-\frac{1}{2(1-\hat{\rho}^2)}\left\{\left(\frac{x-\hat{\mu}_x}{\hat{\sigma}_x}\right)^2 - 2\hat{\rho}\left(\frac{x-\hat{\mu}_x}{\hat{\sigma}_x}\right)\left(\frac{y-\hat{\mu}_y}{\hat{\sigma}_y}\right) + \left(\frac{y-\hat{\mu}_y}{\hat{\sigma}_y}\right)^2\right\}} dxdy \quad (22)$$

en de kans op een (herfst+winter)-tijdvak met ten hoogste x_{HW} mm neerslag, gevolgd door een (lente+zomer)-tijdvak met ten hoogste y_{LZ} mm neerslag wordt (met weglating van de indices HW, LZ, HZ) gegeven door:

$$F(x,y) = \int_{-\infty}^y \int_{-\infty}^x f(\xi, \eta) d\xi d\eta = \frac{1}{2\pi\hat{\sigma}_x\hat{\sigma}_y\sqrt{1-\hat{\rho}^2}} \int_{-\infty}^y \int_{-\infty}^x e^{-\frac{1}{2(1-\hat{\rho}^2)}\left\{\left(\frac{\xi-\hat{\mu}_x}{\hat{\sigma}_x}\right)^2 - 2\hat{\rho}\left(\frac{\xi-\hat{\mu}_x}{\hat{\sigma}_x}\right)\left(\frac{\eta-\hat{\mu}_y}{\hat{\sigma}_y}\right) + \left(\frac{\eta-\hat{\mu}_y}{\hat{\sigma}_y}\right)^2\right\}} d\xi d\eta \quad (23)$$

Deze uitdrukking kan niet in gesloten vorm worden geïntegreerd, doch wel herleid tot de volgende uitdrukking:

voor $\hat{\rho} > 0$:

$$F(x,y) = \frac{1}{\sqrt{2\pi}s_0} \int_{x_0 - \hat{\rho}y_0}^{\infty} e^{-\frac{1}{2} \left(\frac{t-x_0}{s_0} \right)^2} \cdot \bar{\Phi}(t) \cdot dt \quad (24)$$

waarin $\bar{\Phi}(t)$ de cumulatieve standaard normale verdeling voorstelt en

$$x_0 = \frac{x - \hat{\mu}_x}{\hat{\sigma}_x \sqrt{1-\hat{\rho}^2}}, \quad y_0 = \frac{y - \hat{\mu}_y}{\hat{\sigma}_y \sqrt{1-\hat{\rho}^2}}, \quad s_0 = \frac{\hat{\rho}}{\sqrt{1-\hat{\rho}^2}} \quad (25)$$

en voor $\hat{\rho} < 0$:

$$F(x,y) = \frac{1}{\sqrt{2\pi}s_0} \int_{-\infty}^{x_0 - \hat{\rho}y_0} e^{-\frac{1}{2} \left(\frac{t-x_0}{s_0} \right)^2} \cdot \bar{\Phi}(t) \cdot dt \quad (26)$$

met $\bar{\Phi}(t)$, x_0 en y_0 in dezelfde betekenis als hierboven en $s_0 = \frac{-\hat{\rho}}{\sqrt{1-\hat{\rho}^2}}$ (27)

Geven we aan $F(x,y)$ volgens (23) met de gegeven waarden van $\hat{\mu}_{HW}$, $\hat{\sigma}_{HW}$, $\hat{\mu}_{LZ}$, $\hat{\sigma}_{LZ}$ en $\hat{\rho}_{HZ}$ de voorgeschreven waarden $P = 0,01, 0,02, 0,05, 0,10, 0,20, 0,30$, dan worden met deze kanswaarden als parameter krommen $F(x_{HW}, y_{LZ}) = P$ vastgelegd welke in fig. 7 grafisch zijn voorgesteld. De punten van deze krommen zijn niet volgens (23) of (24) berekend, doch bepaald met behulp van de "Tabel van de tweedimensionale kansverdeling" (Mededeling No. 245 (1965) van het Instituut voor de Veredeling van Tuinbouwgewassen, door J.H.A.Ferguson). Deze tabellen leveren bij een gegeven waarde van $r \equiv \hat{\rho}_{HZ}$ en bij gegeven waarden van $P_{xy} \equiv P$ bij elkaar behorende marginale kansen P_x en P_y , zodanig dat $P_{xy} = P(P_x, P_y, r) = P(x \leq h, y \leq k | r)$. Bij elkaar behorende paren waarden h, k zijn als punten in de grafiek gezet en door een vloeiende kromme verbonden.

3.3. De neerslagoverschotsommen van de wintertijdvakken en de daarop volgende lentetijdvakken

Onder het neerslagoverschot wordt verstaan het verschil $N - pE_0$, waarin N de hoeveelheid neerslag voorstelt, E_0 de open water verdamping,

berekend volgens de methode van Penman, en p een factor die voor de wintermaanden de waarde 0,6 heeft, voor de lente- en de herfstmaanden de waarde 0,7 en voor de zomermaanden de waarde 0,8.

Er werd een stippendiagram samengesteld van de neerslagoverschotsommen (in mm) van de wintertijdvakken (x_W) en van de daarop volgende lentetijdvakken (y_L), (fig. 10). In totaal bevat de stippenwolk 49 punten. Voor het gegeven tijdvak 1911-1960 gelden als uiterste waarden van x_W en y_L resp. $58 \leq x_W \leq 290$ en $-110 \leq y_L \leq 63$. Er werden voorts van de grootheden x_W en y_L de gemiddelde waarde en de standaardafwijking berekend. De uitkomsten zijn in tabel 12 samengevat.

Tabel 12

$\hat{\mu}_W = 170,9 \text{ mm}$	$\hat{\mu}_L = -24,9 \text{ mm}$
$\hat{\sigma}_W = 54,8 \text{ mm}$	$\hat{\sigma}_L = 47,8 \text{ mm}$

Tevens werd de correlatiecoëfficiënt tussen de bij elkaar behorende waarden van x_W en y_L bepaald. Deze bedroeg:

$$\hat{\rho}_{WL} = +0,16.$$

Deze waarde, afgeleid uit een steekproef van 49 waardeparen (x_W, y_L), bleek na toetsing, bij een onbetrouwbaarheidsdrempel $\alpha = 0,05$, niet significant van 0 te verschillen. De verdelingen van x_W en y_L kunnen dus als ongecorreleerd worden beschouwd.

Door toetsing is nagegaan of de verdelingen van x_W en y_L als normaal kunnen worden beschouwd. Dit bleek inderdaad het geval te zijn. De cumulatieve verdelingen van x_W en van y_L zijn op lineair waarschijnlijkheidspapier in beeld gebracht (fig. 11 en fig. 12). In deze figuren zijn tevens de rechten getekend van de cumulatieve normale verdelingen, gebaseerd op de steekproefwaarden van $\hat{\mu}_W$, $\hat{\sigma}_W$ en $\hat{\mu}_L$, $\hat{\sigma}_L$ uit tabel 12. De steekproefpunten slingeren om deze rechten maar blijven binnen de overeenkomstig (14) berekende 95% betrouwbaarheidsmarges. De verdeling van de neerslagoverschotsommen van de wintertijdvakken x_W en de daarop volgende lentetijdvakken y_L kan dus bij benadering door een bivariabele normale verdeling worden voorgesteld overeenkomstig (15) en de kans op een winter met ten hoogste x_W mm neerslagoverschot, gevolgd door een

lente met ten hoogste y_L mm neerslagoverschot wordt weergegeven overeenkomstig (16).

Geven we aan $F(x_W, y_L)$ voorgeschreven waarden $P = 0,01, 0,02, 0,05, 0,10, 0,20, 0,30$ dan worden met deze kanswaarden als parameter krommen $F(x_W, y_L) = P$ vastgelegd, welke in fig. 10 grafisch zijn voorgesteld.

5.4. De neerslagoverschotsommen van de (herfst+winter)-tijdvakken en de daarop volgende (lente+zomer)-tijdvakken

Er werd een stippendiagram samengesteld van de neerslagoverschotsommen (in mm), zoals gedefinieerd in 3.3., van de (herfst+winter)-tijdvakken (x_{HW}) en van de daarop volgende (lente+zomer)-tijdvakken (y_{LZ}), (fig. 13). In totaal bevat de stippenwolk 49 punten. Voor het gegeven tijdvak 1911-1960 gelden als uiterste waarden van x_{HW} en y_{LZ} resp.: $108 \leq x_{HW} \leq 492$ en $-311 \leq y_{LZ} \leq 149$. Er werden voorts van de grootheden x_{HW} en y_{LZ} de gemiddelde waarde en de standaardafwijking berekend. De uitkomsten zijn in tabel 13 samengevat.

Tabel 13

$\hat{\mu}_{HW} = 313,7$ mm	$\hat{\mu}_{LZ} = -64,8$ mm
$\hat{\sigma}_{HW} = 98,1$ mm	$\hat{\sigma}_{LZ} = 86,4$ mm

Tevens werd de correlatiecoëfficiënt tussen de bij elkaar behorende waarden van x_{HW} en y_{LZ} bepaald. Deze bedroeg:

$$\hat{\rho}_{HZ} = +0,380$$

Deze waarde, afgeleid uit een steekproef van 49 waardeparen (x_{HW}, y_{LZ}), bleek na toetsing bij een onbetrouwbaarheidsdrempel $\alpha = 0,05$, significant van 0 te verschillen. De verdelingen van x_{HW} en y_{LZ} kunnen dus niet als ongecorrleerd worden beschouwd. Door toetsing is nagegaan of de verdelingen van x_{HW} en y_{LZ} als normaal kunnen worden beschouwd. Dit bleek inderdaad het geval te zijn.

De cumulatieve verdelingen van x_{HW} en van y_{LZ} zijn op lineair waarschijnlijkheidspapier in beeld gebracht (fig. 14 en fig. 15). In deze figuren zijn tevens de rechten getekend van de cumulatieve

normale verdelingen, gebaseerd op de steekproefwaarden van $\hat{\mu}_{HW}$, $\hat{\sigma}_{HW}$, en $\hat{\mu}_{LZ}$, $\hat{\sigma}_{LZ}$ uit tabel 13. De steekproefpunten slingeren om deze rechten maar blijven binnen de overeenkomstig (14) berekende 95% betrouwbaarheidsmarges. De verdeling van de neerslagoverschotsommen van de (herfst+winter)-tijdvakken x_{HW} en de daarop volgende (lente+zomer)-tijdvakken y_{LZ} kan dus bij benadering door een bivariabele normale verdeling worden voorgesteld overeenkomstig (22) en de kans op een (herfst-winter)-tijdvak met ten hoogste x_{HW} mm neerslagoverschot, gevolgd door een (lente+zomer)-tijdvak met ten hoogste y_{LZ} mm neerslagoverschot, wordt weergegeven overeenkomstig (23).

Geven we aan $F(x_{HW}, y_{LZ})$ voorgeschreven waarden $P = 0,01, 0,02, 0,05, 0,10, 0,20, 0,30$ dan worden met deze kanswaarden als parameters krommen $F(x_{HW}, y_{LZ}) = P$ vastgelegd welke in fig. 13 grafisch zijn voorgesteld.

5.5. De vijf droogste neerslagtijdvakken (x_W, y_L) en (x_{HW}, y_{LZ}) en de vijf droogste neerslagoverschottijdvakken (x'_W, y'_L) en (x'_{HW}, y'_{LZ}) in het tijdvak 1911-1960 voor De Bilt

Deze zijn in tabel 14 bijeengebracht, gerangschikt naar opklimmende x-waarden. Tevens zijn in de kolommen R de rangnummers opgenomen volgens opklimmende (x+y)-waarden.

Tabel 14

	R	x_W	y_L		R	x_{HW}	y_{LZ}
1933-1934	3	69	170	1948-1949	2	192	249
1932-1933	2	84	126	1953-1954	5	196	413
1928-1929	1	38	87	1931-1932	3	204	334
1946-1947	5	88	212	1920-1921	1	205	166
1931-1932	4	93	165	1933-1934	4	233	366

	R	x'_W	y'_L		R	x'_{HW}	y'_{LZ}
1933-1934	3	58	0	1953-1954	5	108	-10
1932-1933	2	73	-41	1948-1949	2	111	-218
1928-1929	1	76	-85	1931-1932	3	124	-92
1931-1932	4	78	12	1920-1921	1	127	-311
1946-1947	5	78	34	1933-1934	4	154	-96

Bezien we de $(x_W y_L)$ - en de $(x'_W y'_L)$ -combinaties naar hun volgorde volgens opklimmende $(x_W + y_L)$ - resp. $(x'_W + y'_L)$ -waarden, dan is er voor de rangnummers 1 t/m 5 volledige overeenstemming van de jaren:

(1) 1928-1929, (2) 1932-1933, (3) 1933-1934, (4) 1931-1932, (5) 1946-1947.

Bezien we de (x_{HW}, y_{LZ}) - en de (x'_{HW}, y'_{LZ}) -combinaties naar hun volgorde volgens opklimmende $(x_{HW} + y_{LZ})$ - resp. $(x'_{HW} + y'_{LZ})$ -waarden, dan is er voor de rangnummers 1 t/m 5 volledige overeenstemming van de jaren: (1) 1920-1921, (2) 1948-1949, (3) 1931-1932, (4) 1933-1934, (5) 1953-1954. Blijkbaar maakt het weinig of geen verschil of men de droogste tijdvakcombinaties volgens de neerslag of het neerslagoverschot rangschikt.

Vergelijken we deze uitkomsten welke voor De Bilt gelden met die welke voor Zwanenburg-Hoofddorp in tabel 9 zijn gegeven, dan zien we allereerst dat in deze tabel de combinatie $(x_W y_L)$ van het jaar 1928-1929 als eerste staat evenals dat in tabel 14 het geval was en voorts dat de $(x_{HW} y_{LZ})$ -combinaties van de jaren 1920-1921 en 1948-1949 resp. als (1) en (2) met die van tabel 14 overeenstemmen.

4. Voorwaardelijke kansen

Wanneer wij er van uitgaan dat de grootheden x en y een bivariabele normale kansverdeling volgen, gegeven door:

$$f(x,y)dxdy = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\frac{x-\mu_x}{\sigma_x}\cdot\frac{y-\mu_y}{\sigma_y} + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}} dx dy \quad (28)$$

dan zijn de randverdelingen gegeven door:

$$\varphi_1(x)dx = \int_{-\infty}^{+\infty} f(x,\eta) d\eta \cdot dx = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2} dx \quad (29)$$

en

$$\varphi_2(y)dy = \int_{-\infty}^{+\infty} f(\xi,y) d\xi \cdot dy = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_y}\right)^2} dy \quad (30)$$

De voorwaardelijke verdeling van x bij een gegeven waarde $y^{\#}$ wordt gegeven door:

$$f_1(x|y^{\#}) dx = \frac{f(x;y^{\#}) dx}{\varphi_2(y^{\#})} = \frac{1}{\sqrt{2\pi\sigma_x^{\#}}} e^{-\frac{1}{2}\left(\frac{x-\mu_x^{\#}}{\sigma_x^{\#}}\right)^2} \cdot dx \quad (31)$$

$$\text{waarin: } \begin{cases} \mu_x^{\#} = \mu_x + \rho \frac{\sigma_x}{\sigma_y} (y^{\#} - \mu_y) \\ \sigma_x^{\#} = \sigma_x \sqrt{1-\rho^2} \end{cases} \quad (31a)$$

Evenzo wordt de voorwaardelijke verdeling van y bij een gegeven waarde $x^{\#}$ gegeven door:

$$f_2(y|x^{\#}) dy = \frac{f(y;x^{\#}) dy}{\varphi_1(x^{\#})} = \frac{1}{\sqrt{2\pi\sigma_y^{\#}}} e^{-\frac{1}{2}\left(\frac{y-\mu_y^{\#}}{\sigma_y^{\#}}\right)^2} \cdot dy \quad (32)$$

$$\text{waarin: } \begin{cases} \mu_y^{\#} = \mu_y + \rho \frac{\sigma_y}{\sigma_x} (x^{\#} - \mu_x) \\ \sigma_y^{\#} = \sigma_y \sqrt{1-\rho^2} \end{cases} \quad (32a)$$

Met behulp van (32) en 32a) kunnen we nu bijv. de voorwaardelijke kans berekenen op tenminste y mm neerslag in de lente als $x^{\#}$ mm neerslag in de voorafgaande winter is gevallen. Deze voorwaardelijke kans wordt dan gegeven door:

$$\begin{aligned} P(\eta \geq y | x^{\#}) &= \int_y^{\infty} f_2(\eta | x^{\#}) d\eta = \frac{1}{\sqrt{2\pi\sigma_y^{\#}}} \int_y^{\infty} e^{-\frac{1}{2}\left(\frac{\eta-\mu_y^{\#}}{\sigma_y^{\#}}\right)^2} d\eta = \frac{1}{\sqrt{2\pi}} \int_{\frac{y-\mu_y^{\#}}{\sigma_y^{\#}}}^{\infty} e^{-\frac{1}{2}t^2} dt = \\ &= 1 - \Phi\left(\frac{y-\mu_y^{\#}}{\sigma_y^{\#}}\right) \end{aligned} \quad (33)$$

5. Samenvatting en conclusies

In het voorgaande is een oriënterend onderzoek verricht naar de frequentie van het optreden van een droge winter gevolgd door een droge lente, resp. van een droge (herfst+winter) gevolgd door een droge (lente+zomer). Zulks geschiedde aan de hand van de desbetreffende neerslaggegevens ontleend aan de waarnemingsreeks Zwanenburg-Hoofddorp over de jaren 1735-1971. De neerslagtotalen voor winter, lente, (herfst+winter), (lente+zomer) blijken alle bij goede benadering aan normale verdelingen te voldoen. Tussen de neerslagsommen van winter en opvolgende lente resp. (herfst+winter) en opvolgende (lente+zomer) bestaat geen correlatie. De simultane kans op een hoeveelheid neerslag van ten hoogste x mm in de winter, resp. (herfst+winter), en een hoeveelheid neerslag van ten hoogste y mm in de daarop volgende lente, resp. (lente+zomer) kan op eenvoudige wijze worden berekend als product van de marginale kansen.

Het onderzoek werd voortgezet met de overeenkomstige neerslaggegevens van De Bilt over de jaren 1911-1960, uitgebreid met een daaraan parallel lopend onderzoek van de gegevens van het neerslagoverschot. Alle verdelingen van de neerslaghoeveelheden en van de neerslagoverschothoeveelheden in de winter, lente, (herfst+winter) en (lente+zomer) bleken bij goede benadering normaal verdeeld te zijn. Tussen de neerslagsommen, resp. de neerslagoverschotsommen van de winter en de daarop volgende lente bleek geen correlatie te bestaan. Wel bestaat tussen de hoeveelheden neerslag, resp. neerslagoverschot van (herfst+winter) en van de daarop volgende (lente+zomer) een zwakke, doch significant van 0 verschillende correlatie. In dit laatste geval kan in beginsel van tabellen voor de twee-dimensionale normale verdeling voor specifieke waarden van de correlatiecoëfficiënt worden gebruik gemaakt om de simultane kans op hoeveelheden van ten hoogste x mm in de winter, resp. (herfst+winter), gevolgd door hoeveelheden van ten hoogste y mm in de daarop volgende lente, resp. (lente+zomer) te bepalen.

De omstandigheid echter dat voor De Bilt over de jaren 1911-1960 tussen de neerslaghoeveelheden van de winter en de daarop volgende lente geen significant van 0 afwijkende correlatie werd gevonden, maar wel tussen de neerslaghoeveelheden van (herfst+winter) en daarop volgende (lente+zomer), was aanleiding om ook van Zwanenburg-Hoofddorp

over de jaren 1911-1960 de overeenkomstige correlatie te berekenen. Als uitkomsten werden voor dit geval verkregen:

$$\hat{\rho}'_{WL} = +0,104 \qquad \hat{\rho}'_{HZ} = +0,334$$

terwijl de overeenkomstige correlaties berekend over de jaren 1735-1969 als uitkomsten leverden:

$$\hat{\rho}_{WL} = +0,103 \qquad \hat{\rho}_{HZ} = +0,087$$

Al deze uitkomsten, behalve die voor $\hat{\rho}'_{HZ}$, verschillen niet significant van 0. Evenals voor De Bilt vinden we derhalve voor Zwanenburg-Hoofddorp over hetzelfde tijdvak 1911-1960 een significant van 0 verschillende correlatie resp. met de waarden +0,346 en +0,334, terwijl voor het tijdvak 1735-1969 de niet significant van 0 verschillende waarde 0,087 wordt gevonden. De vraag rijst of de significant van 0 afwijkende waarde van de correlatiecoëfficiënt voor de 50-jarige waarnemingsreeksen van zowel De Bilt als Zwanenburg-Hoofddorp toch niet als gevolg van een steekproefeffect moeten worden gezien. Dat hiertoe aanleiding bestaat blijkt uit de volgende beschouwing.

Duiden we de neerslaghoeveelheden voor herfst, winter, lente en zomer achtereenvolgens met de variabelen x_3, x_4, x_1 en x_2 aan. Algemeen geldt dat wanneer

$y = \sum_{i=1}^k a_i x_i$ een lineaire functie van k variabelen voorstelt en $z = \sum_{i=1}^k b_i x_i$ een tweede lineaire functie van deze k variabelen, waarbij $b_i \neq c a_i$, dat de correlatiecoëfficiënt tussen y en z wordt gegeven door:

$$\rho_{yz} = \frac{\sum_{j=1}^k \sum_{i=1}^k a_i b_j \sigma_{ij}}{\sqrt{\sum_{\beta=1}^k \sum_{\alpha=1}^k a_{\alpha} a_{\beta} \sigma_{\alpha\beta} \cdot \sum_{\mu=1}^k \sum_{\lambda=1}^k b_{\lambda} b_{\mu} \sigma_{\lambda\mu}}}$$

Hierin is:

$$\sigma_{ij} = \mathcal{E} \left[(x_i - \xi_i)(x_j - \xi_j) \right] \text{ (covariantie); schatting: } \hat{\sigma}_{ij} = \frac{1}{n} \sum_{\lambda=1}^n (x_{i\lambda} - \hat{\xi}_i)(x_{j\lambda} - \hat{\xi}_j)$$

$$\sigma_{ii} = \mathcal{E} \left[(x_i - \xi_i)^2 \right] \text{ (variantie); schatting: } \hat{\sigma}_{ii} = \frac{1}{n} \sum_{\lambda=1}^n (x_{i\lambda} - \hat{\xi}_i)^2$$

$$\xi_i = \mathcal{E} \left[x_i \right] \text{ (gemiddelde); schatting: } \hat{\xi}_i = \frac{1}{n} \sum_{\lambda=1}^n x_{i\lambda}$$

terwijl voorts:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii} \sigma_{jj}}} \text{ (correlatie); schatting: } \hat{\rho}_{ij} = \frac{\hat{\sigma}_{ij}}{\sqrt{\hat{\sigma}_{ii} \hat{\sigma}_{jj}}}$$

In ons geval is nu $y = x_3 + x_4$ (d.w.z. $a_1 = a_2 = 0, a_3 = a_4 = 1$) en $z = x_1 + x_2$ (d.w.z. $b_1 = b_2 = 1, b_3 = b_4 = 0$) en dus:

$$\begin{aligned} \rho_{HZ} \equiv \rho_{yz} &= \frac{\sigma_{31} + \sigma_{32} + \sigma_{41} + \sigma_{42}}{\sqrt{(\sigma_{33} + 2\sigma_{34} + \sigma_{44})(\sigma_{11} + 2\sigma_{12} + \sigma_{22})}} = \\ &= \frac{\sqrt{\sigma_{33}\sigma_{11}}\rho_{31} + \sqrt{\sigma_{33}\sigma_{22}}\rho_{32} + \sqrt{\sigma_{44}\sigma_{11}}\rho_{41} + \sqrt{\sigma_{44}\sigma_{22}}\rho_{42}}{\sqrt{(\sigma_{33} + 2\sqrt{\sigma_{33}\sigma_{44}}\rho_{34} + \sigma_{44})(\sigma_{11} + 2\sqrt{\sigma_{11}\sigma_{22}}\rho_{12} + \sigma_{22})}} \end{aligned}$$

terwijl voor $y' = x_4$ en $z' = x_1$ geldt:

$$\rho_{WL} \equiv \rho_{y'z'} = \frac{\sigma_{41}}{\sqrt{\sigma_{44} \cdot \sigma_{11}}} = \rho_{41}$$

We toetsen nu of de schattingswaarden van de correlatiecoëfficiënten $\hat{\rho}_{31}$, $\hat{\rho}_{32}$, $\hat{\rho}_{41}$ en $\hat{\rho}_{42}$ die in de teller van $\hat{\rho}_{yz}$ voorkomen significant van nul verschillen. Voor steekproeven ter grootte n (n voldoende groot) volgt de toetsingsgrootte $u = -n \lg(1 - \hat{\rho}^2)$ onder de nulhypothese $\rho = 0$, bij benadering de χ^2 -verdeling met 1 graad van vrijheid. De nulhypothese behoeft niet te worden verworpen indien $u < 3,84$ blijkt te zijn, behoudens een kans van 5% op een foutieve uitspraak.

Aldus moet voor $n=235$ $\hat{\rho} < 0,128$ zijn en voor $n=50$ $\hat{\rho} < 0,258$. Het blijkt dat voor het geval Zwanenburg-Hoofddorp voor beide steekproeven de correlatiecoëfficiënten $\hat{\rho}_{31}$, $\hat{\rho}_{32}$, $\hat{\rho}_{41}$, $\hat{\rho}_{42}$ alle beneden de aangegeven grenzen liggen en dus inderdaad niet significant van nul verschillen. Hetzelfde geldt voor het geval De Bilt. Dit resultaat versterkt de veronderstelling dat de voor $n=50$ gevonden waarden van $\hat{\rho}_{yz}$ van Zwanenburg-Hoofddorp en van De Bilt, hoewel significant van nul verschillend, niettemin aan een steekproefeffect moeten worden toegeschreven, d.i. de uitkomsten als zuiver toevallig moeten worden beschouwd.

De berekende kansen voor de extreme omstandigheden van droogte bezitten geen grote nauwkeurigheid, zelfs niet in het geval van de ongeveer 235 jaren lange reeks van Zwanenburg-Hoofddorp. Binnen de onzekerheidsmarges bestaat er evenwel een redelijke overeenstemming tussen de volgens de formules berekende kanswaarden, resp. te verwachten aantallen gevallen en de aantallen waargenomen gevallen. De gemiddelde herhalings-tijd, berekend voor de 10 (12) droogste combinaties van winter en lente, resp. (herfst+winter) en (lente+zomer) is van de orde van 20 jaren, echter met standaardafwijkingen van 19 à 24 jaren. De gevallen liggen dan ook zeer onregelmatig over de tijd verdeeld en de gemiddelde herhalings-tijd heeft weinig of geen prognostische waarde.

Uit het onderzoek naar extreme omstandigheden van droogte te De Bilt blijkt dat een rangschikking volgens toenemende neerslaghoeveelheden (per seizoen of per half jaar) vrijwel of geheel overeenstemt met een volgens toenemende neerslagoverschothoeveelheden. Een klassificatie op grond van het neerslagoverschot biedt in dit geval geen voordeel boven die op grond van de neerslag alleen.

Uitgaande van de bivariabele normale kansverdeling met correlatie tussen de variabelen kan gemakkelijk de voorwaardelijke kansverdeling worden afgeleid op bijv. tenminste y mm neerslag in de lente als x^* mm in de voorafgaande winter is gevallen.

De vraag kan worden gesteld of op grond van de gegeven beschouwingen waarbij als tijdvaklengte het seizoen of halfjaar werd gekozen, reeds een strategie zou kunnen worden opgebouwd voor de regeling van de waterinhoud van waterreservoirs, danwel of daarvoor kortere tijdvaklengten gewenst zouden zijn, bijv. de maand en veelvoud daarvan tot 6 maanden

in willekeurige combinatie voor de x- en y-waarden. In beginsel kunnen de bijbehorende bivariabele kansverdelingen op dezelfde wijze worden berekend als in het voorgaande werd gedaan, aannemende dat ook de maandhoeveelheden nog aan een normale verdeling voldoen. Dit laatste zou allereerst dan moeten worden nagegaan. Uiteraard zal een dergelijk onderzoek toegepast op redelijk lange reeksen van waarnemingen een veel omvangrijker werk betekenen dan hier werd verricht.

Het behoeft geen betoog dat de cumulatieve bivariabele normale kansverdeling elke combinatie van grenswaarden (x,y) toelaat. Men kan dus niet alleen kansen voor de combinatie droog/droog nagaan maar ook voor nat/droog, droog/nat en nat/nat, al naar men de grenswaarden x en y vastlegt.

6. Summary and conclusions

In the framework of the activities of the Working Group on Hydrological Research in Overijssel, the question has been posed, very generally, how often a dry period in autumn and winter is followed by a dry period in spring and summer. This question, as yet not well defined, araised against the background of a hydrological question, namely the most efficient control of water reservoirs eventually to be established for the watersupply in future. In order to define the question more precisely and to get an answer with the aid of immediately available data on precipitation, and where possible also on evaporation, it appeared for a first orientation as the simplest way to choose as period the seasons (spring = March + April + May, summer, autumn, winter) and as series of observations that of precipitation of Zwanenburg-Hoofddorp, published by A.Labrijn in Mededelingen en Verhandelingen 49: "Het klimaat van Nederland gedurende de laatste twee en een halve eeuw", since 1735 and extended with the similar data of Hoofddorp up to 1971 inclusive.

Though the question just mentioned and its hydrological background are inspired by the hydrological research which is going on in Salland (Overijssel) so that it would have been natural to use the data of the region itself, it has been preferred to consider a series of observations as long as possible, in this case that of Zwanenburg-Hoofddorp. Certainly the precipitation regime in the western part of the country

differs from that in the eastern part. However, concentrating oneself mainly on the long lasting dry periods, here of 3 or 6 months, one can safely accept that these periods in most cases will affect the whole country.

Undoubtedly it would have been preferable to base our consideration not merely on precipitation, but rather on precipitation surplus, i.e. the difference of the amount of precipitation minus the amount of evaporation, as this difference feeds the groundwater and the runoff. However, the data of evaporation for Zwanenburg-Hoofddorp are not available. In order to get at least an impression of the behaviour of the precipitation surplus in seasons the available data of De Bilt were used from the period 1911-1960.

It may be wondered at what limiting values of the amount of precipitation or precipitation surplus a period of one or two seasons should be considered as "dry". To some extent this limiting value is arbitrary. One could choose a fixed amount, one could also relate the limiting value to the standard deviation of the frequency distribution of the amounts, or one could let it depend on the more specific hydrological aim. For the present we wish to leave the choice open.

In par. 2.1.1. the precipitation amount of Zwanenburg-Hoofddorp of the winter periods (December + January + February) and those of the subsequent spring periods (March + April + May) were plotted as abscis x_W with the corresponding ordinate y_L for the years from 1735 upto 1971 inclusive in a scatter diagram (237 points in total; Fig. 1). The numbers of values within fixed limits were summarized in a contingency table (Table 1) from which it was concluded by means of a statistical test that the hypothesis that the precipitation amounts of the spring periods are not correlated with those of the preceding winter periods needs not to be rejected on the 5% level of significance.

From the scatter diagram (Fig. 1) frequencies f_W and f_L of the amounts x_W and y_L within classes of 10 mm width were derived, as well as means and standard deviations (Table 2 and 3). Cumulative relative frequencies were computed (Table 2) and the results plotted on normal probability paper (Fig. 2) in which also the rectilinear normal distribution curves were drawn, corresponding with the computed sample means and standard deviations.(Par. 2.1.2.).

Moreover the 95% confidence intervals were drawn for these sample distributions in the same graph, from which it became obvious that the sample points could be considered to belong to normal distributions. (Par. 2.1.3.).

The conclusion can be drawn that the amounts of precipitation of winter periods (x_W) and those of the subsequent spring periods (y_L) have a bivariate normal distribution, with a density function that equals the product of the marginal normal density functions of x_W and y_L respectively. On the basis of this result the chance for a specified "dry" period in winter followed by a specified "dry" period in the subsequent spring can easily be computed. So it appears for example that the chance for a winter period with 100 mm of precipitation or less, followed by a spring period with 100 mm precipitation or less is 0,0333. This situation occurs once in 30 years on the mean; in a sample of 237 years 8 occasions could be expected. In reality there were found 10 occasions, which is in good agreement with the number that could be expected if the confidence intervals of the distribution functions are taken into account. In Fig. 1 curves are drawn representing specified combinations of winter (x_W) - spring (y_L) amounts, not to be exceeded, for a given simultaneous chance or return period. (Par. 2.1.4.).

Similar procedures have been used for the investigation of "dry" (autumn + winter)-periods followed by "dry" (spring + summer)-periods, with similar results. (Par. 2.2.1., 2.2.2., 2.2.3., 2.2.4.).

In par. 2.2.5. the 10 driest winter/spring periods and the 12 driest (autumn + winter)/(spring + summer)-periods are put together in increasing order of $x+y$. (Table 9). The limits used for x and y were 100 mm for the first case and 300 mm for the second. If the cases were ordered with time and the time differences measured between consecutive occasions than a mean return period can be obtained of 21 years with a standard deviation of 15 years for the first case and a mean return period of 19 years with a standard deviation of 24 years in the second case. This means that the occurrence of the mentioned occasions is very irregularly spread over the time scale; so their mean return period, even based on a very long period of observations, has little or no prognostic value.

The investigation has been continued with the evaluation of the observations of precipitation and precipitation surplus of De Bilt for the period 1911-1960 along similar lines as before. Starting with the scatter diagram of the precipitation totals in winter (x_W) and in the subsequent spring (y_L) it appeared that the correlation coefficient between the 49 pairs of observations did not differ significantly from zero. The distributions of x_W and y_L could be considered to be normal as was illustrated by the graphs of their relative frequencies on normal probability paper together with the corresponding confidence intervals. (Par. 3.1.).

Similar results were obtained for the totals of precipitation in (autumn + winter) and the subsequent (spring + summer), with the exception however that there proved to be a small correlation between their amounts ($\hat{\rho}_{HZ} = +0,346$), differing significantly from zero. For the construction of the lines of equal probability for different simultaneous values of the x and y parameters not to be exceeded this correlation was taken into account. (Par. 3.2.).

In par. 3.3. the precipitation surplus was introduced, defined as the difference $N - pE_0$, in which N represents the amount of precipitation, E_0 the open water evaporation computed according the method of Penman and p a factor which was taken 0,6 for the winter period, 0,7 for the spring and autumn periods and 0,8 for the summer period. The results obtained with the precipitation surplus were quite similar to those with the precipitation. (Par. 3.3. and 3.4.).

In par. 3.5. the 5 driest winter/spring periods, the 5 driest (autumn + winter)/(spring + summer) periods of precipitation and of precipitation surplus are given for De Bilt in the period 1911-1960. When comparing the results there appears to be but little advantage in using the precipitation surplus instead of the precipitation for the ranging of dry periods.

In par. 4 formulae are given for conditional normal distributions and probabilities.

Apart from the concluding remarks already made par. 5 contains an additional investigation on the correlation of the amounts of precipitation of the winter and the subsequent spring period, as well

of the amounts of precipitation of the (autumn + winter) and the subsequent (spring + summer) period for Hoofddorp in the period 1911-1960. The results for this period were quite similar to those for De Bilt. That means that the correlation coefficients significantly differing from zero as found in the case (autumn + winter)/(spring + summer) for the short period 1911-1960, both for De Bilt and for Zwanenburg-Hoofddorp, very probably are to be ascribed to a sampling effect. This hypothesis is confirmed by a further analysis of the correlation coefficient.

One might question whether on the basis of the preceding considerations, where the seasons or the half year are chosen as basic periods, already a strategy could be built up for the controlling of water-reservoirs or whether shorter periods, e.g. a month and multiples of it up to 6 months in arbitrary combinations for the x- and y-values would be preferred. In principle the bivariate probability distribution can be obtained in the same way as before, assuming that monthly values of precipitation are again normally distributed. Such an investigation however, using reasonably long series of observations, would require far more work than has been done here.

It needs no argument that the cumulative bivariate normal probability distribution admits every combination of limiting values of x and y not to be exceeded in subsequent periods. Therefore all combinations of dry/wet, wet/dry, wet/wet can be handled in just the same way and according to the limits fixed.

7. Naschrift

De schrijver dankt de heer A.Denkema voor de uitvoering van de statistische toetsen en andere berekeningen alsmede voor het ontwerp van een aantal figuren.

Voorts wordt dank gebracht aan Drs. S.Kruizinga voor een aantal kritische opmerkingen.

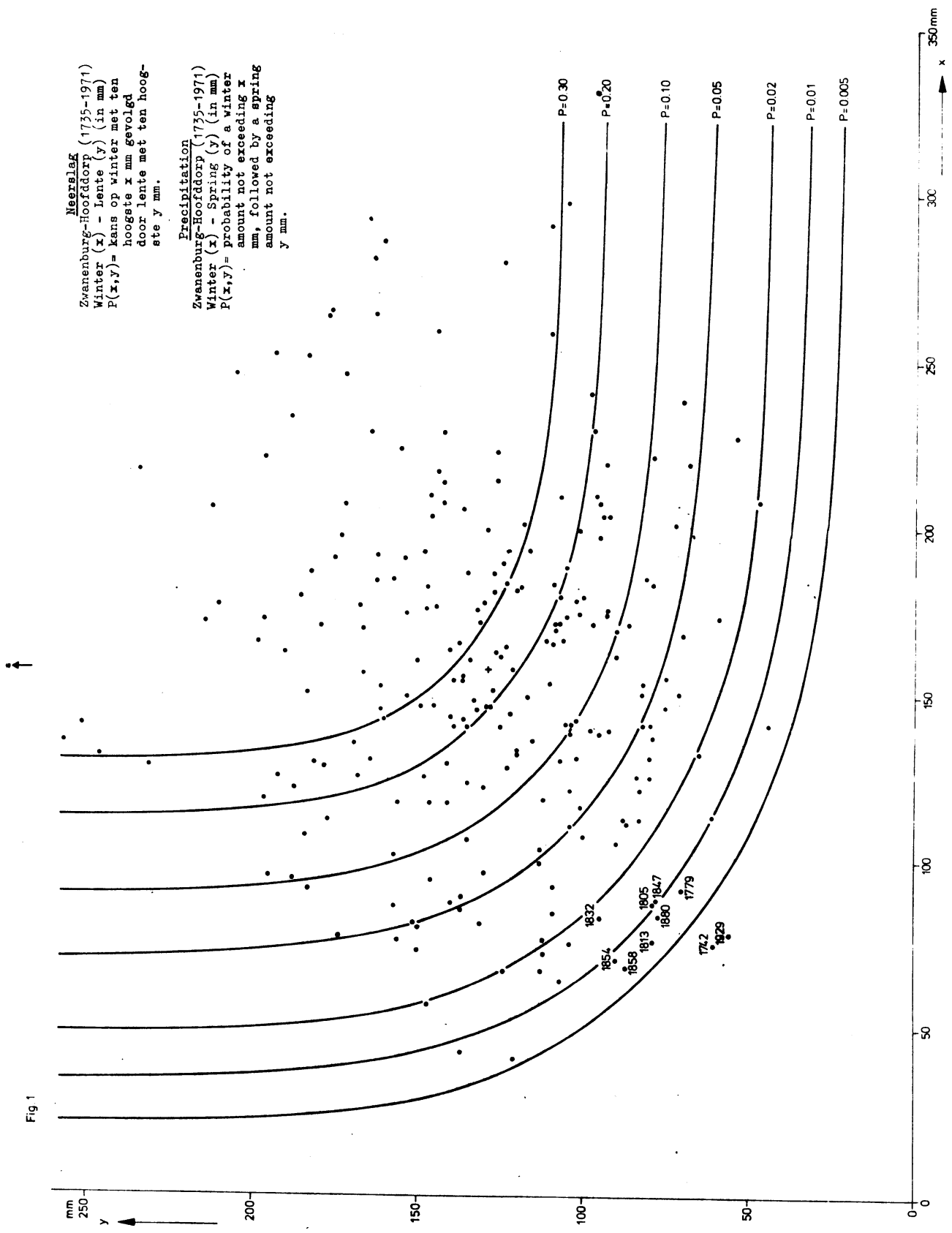
8. Literatuur

- (1) A.Labrijn: "Het klimaat van Nederland gedurende de laatste twee en een halve eeuw". Med. en Verh. 49, KNMI, (1945).

- (2) J.H.A.Ferguson: "Tabel van de twee-dimensionale kansverdeling".
Med. No. 245, Instituut voor de Veredeling van
Tuinbouwgewassen, (1965).
- (3) A.M.Mood & F.A.Graybill: "Introduction to the Theory of Statistics",
McGraw-Hill Book Company, Inc., 2nd Ed. (1963).

De Bilt, 28 februari 1972.

.....



Neerslag
 Zwanenburg-Hoofddorp (1735-1971)
 Winter (x) - Lente (y) (in mm)
 $P(x,y)$ = kans op winter met ten
 hoogste x mm gevolgd
 door lente met ten hoog-
 ste y mm.

Precipitation
 Zwanenburg-Hoofddorp (1735-1971)
 Winter (x) - Spring (y) (in mm)
 $P(x,y)$ = probability of a winter
 amount not exceeding x
 mm, followed by a spring
 amount not exceeding
 y mm.

Fig. 1

Fig. 2

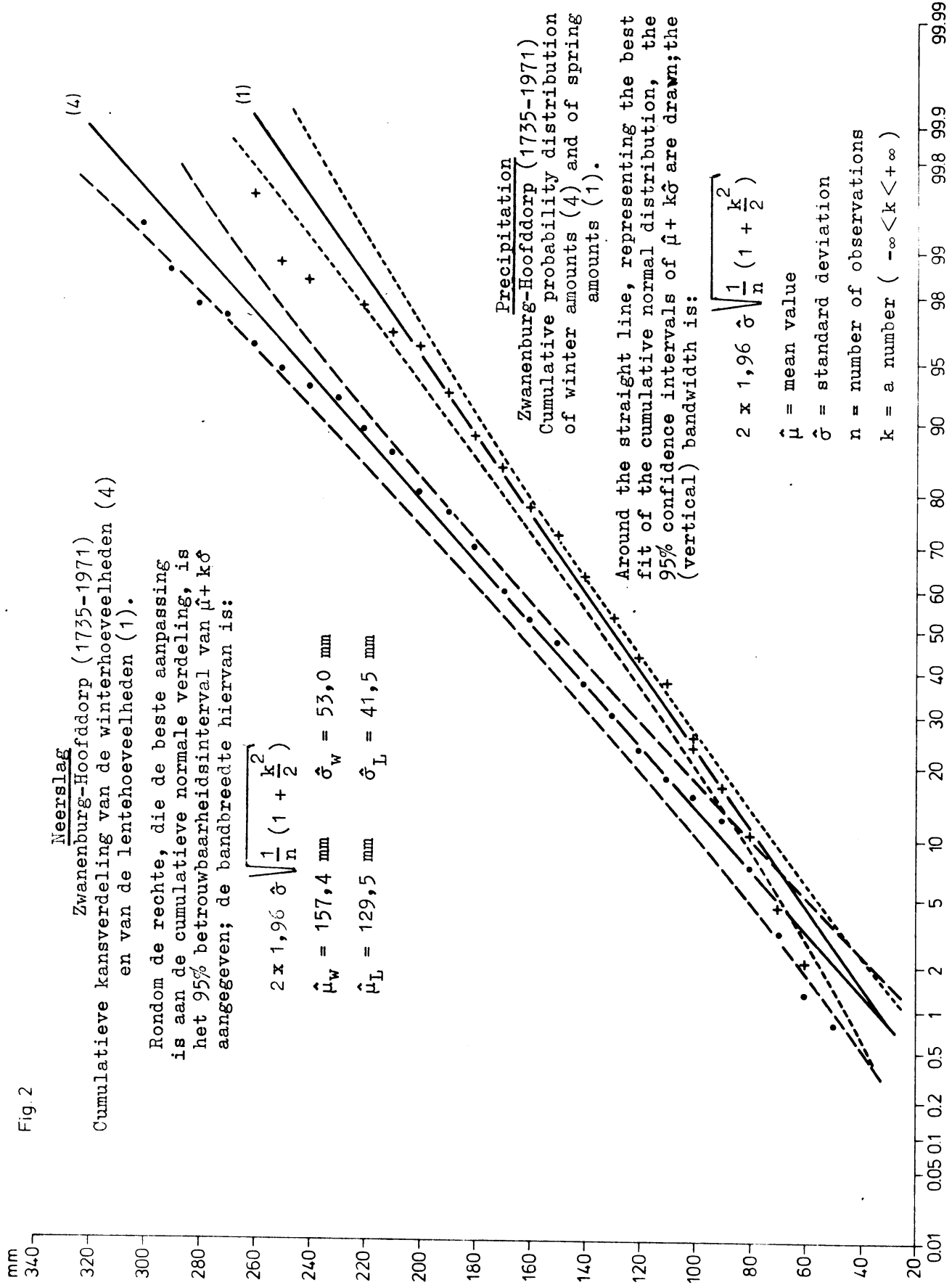
Neerslag
 Zwanenburg-Hoofddorp (1735-1971)
 Cumulatieve kansverdeling van de winterhoeveelheden (4)
 en van de lentehoeveelheden (1).

Rondom de rechte, die de beste aanpassing is aan de cumulatieve normale verdeling, is het 95% betrouwbaarheidsinterval van $\hat{\mu} + k\hat{\sigma}$ aangegeven; de bandbreedte hiervan is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$$\hat{\mu}_w = 157,4 \text{ mm} \quad \hat{\sigma}_w = 53,0 \text{ mm}$$

$$\hat{\mu}_L = 129,5 \text{ mm} \quad \hat{\sigma}_L = 41,5 \text{ mm}$$



Precipitation
 Zwanenburg-Hoofddorp (1735-1971)
 Cumulative probability distribution
 of winter amounts (4) and of spring
 amounts (1).

Around the straight line, representing the best fit of the cumulative normal distribution, the 95% confidence intervals of $\hat{\mu} + k\hat{\sigma}$ are drawn; the (vertical) bandwidth is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

- $\hat{\mu}$ = mean value
- $\hat{\sigma}$ = standard deviation
- n = number of observations
- k = a number ($-\infty < k < +\infty$)

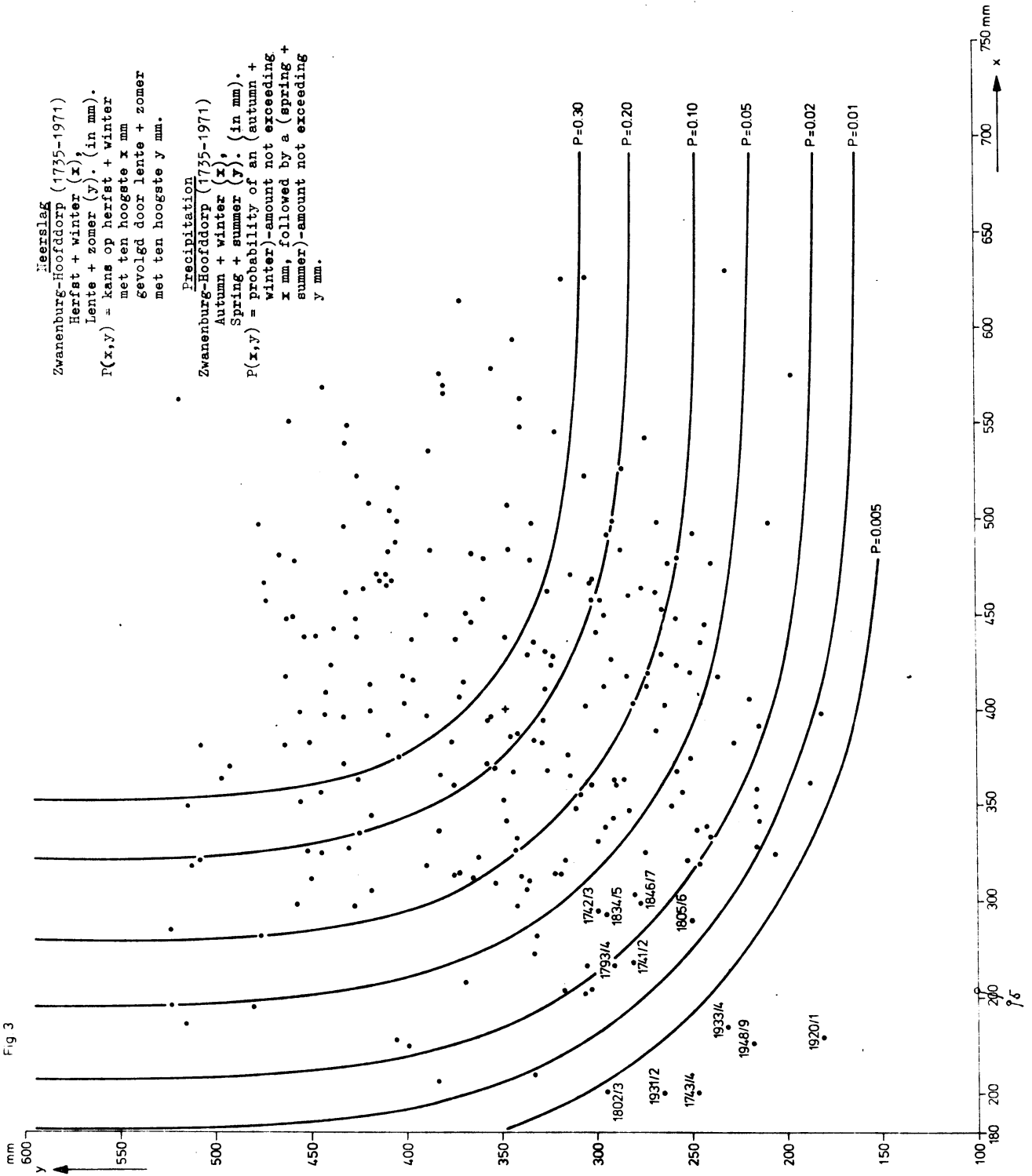


Fig. 3

Neerslag
 Zwanenburg-Hoofddorp (1735-1971)

Herfst + winter (x),
 Lente + zomer (y). (in mm).
 $P(x,y)$ = kans op herfst + winter
 met ten hoogste x mm
 gevolgd door lente + zomer
 met ten hoogste y mm.

Precipitation

Zwanenburg-Hoofddorp (1735-1971)

Autumn + winter (x),
 Spring + summer (y). (in mm).
 $P(x,y)$ = probability of an autumn +
 winter)-amount not exceeding
 x mm, followed by a (spring +
 summer)-amount not exceeding
 y mm.

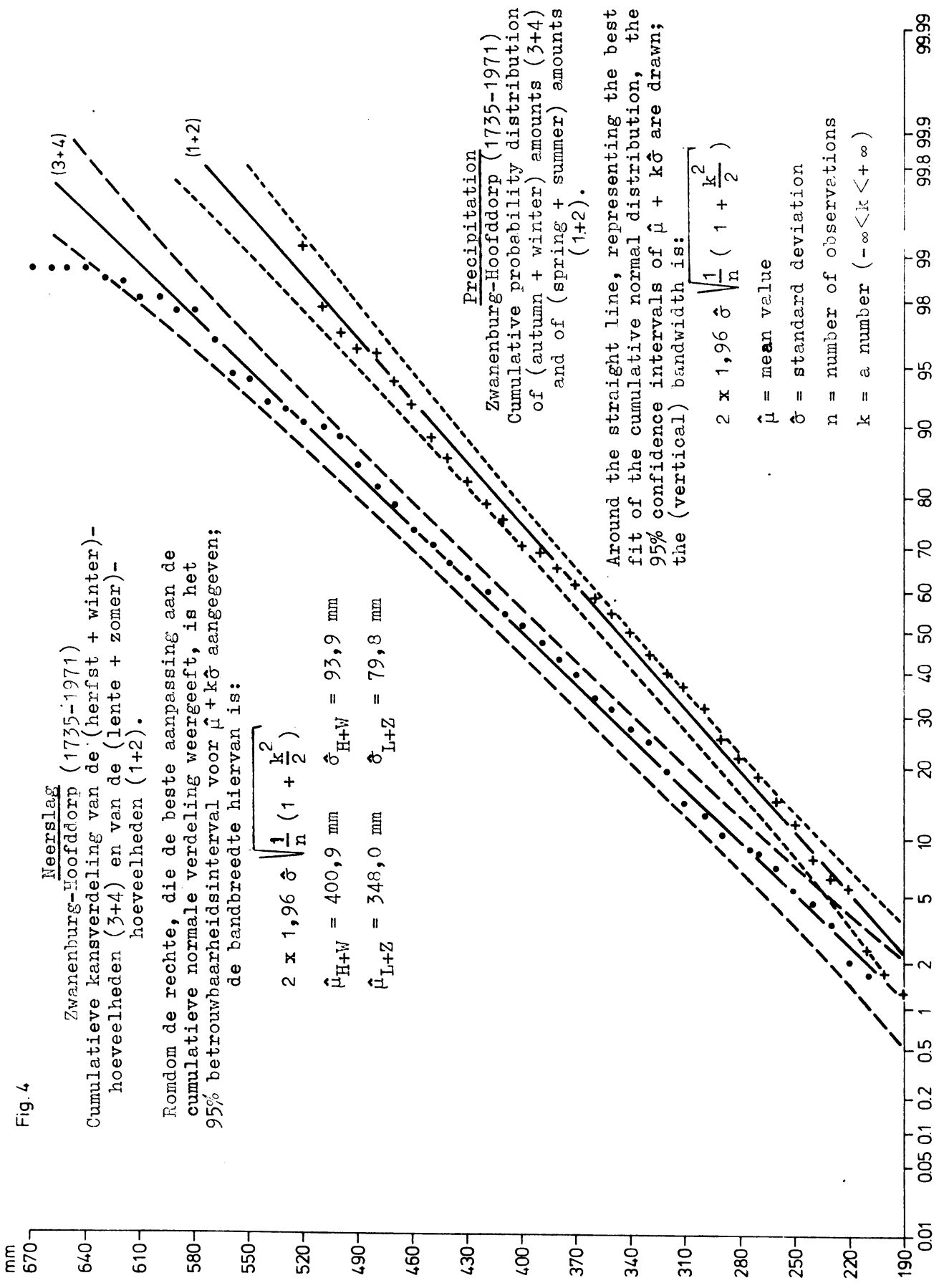


Fig. 4

Neerslag
 Zwanenburg-Hoofddorp (1735-1971)
 Cumulatieve kansverdeling van de (herfst + winter)-
 hoeveelheden (3+4) en van de (lente + zomer)-
 hoeveelheden (1+2).

Romdom de rechte, die de beste aanpassing aan de
 cumulatieve normale verdeling weergeeft, is het
 95% betrouwbaarheidsinterval voor $\hat{\mu} + k\hat{\sigma}$ aangegeven;
 de bandbreedte hiervan is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$$\hat{\mu}_{H+W} = 400,9 \text{ mm} \quad \hat{\sigma}_{H+W} = 93,9 \text{ mm}$$

$$\hat{\mu}_{L+Z} = 348,0 \text{ mm} \quad \hat{\sigma}_{L+Z} = 79,8 \text{ mm}$$

Precipitation
 Zwanenburg-Hoofddorp (1735-1971)
 Cumulative probability distribution
 of (autumn + winter) amounts (3+4)
 and of (spring + summer) amounts
 (1+2).

Around the straight line, representing the best
 fit of the cumulative normal distribution, the
 95% confidence intervals of $\hat{\mu} + k\hat{\sigma}$ are drawn;
 the (vertical) bandwidth is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

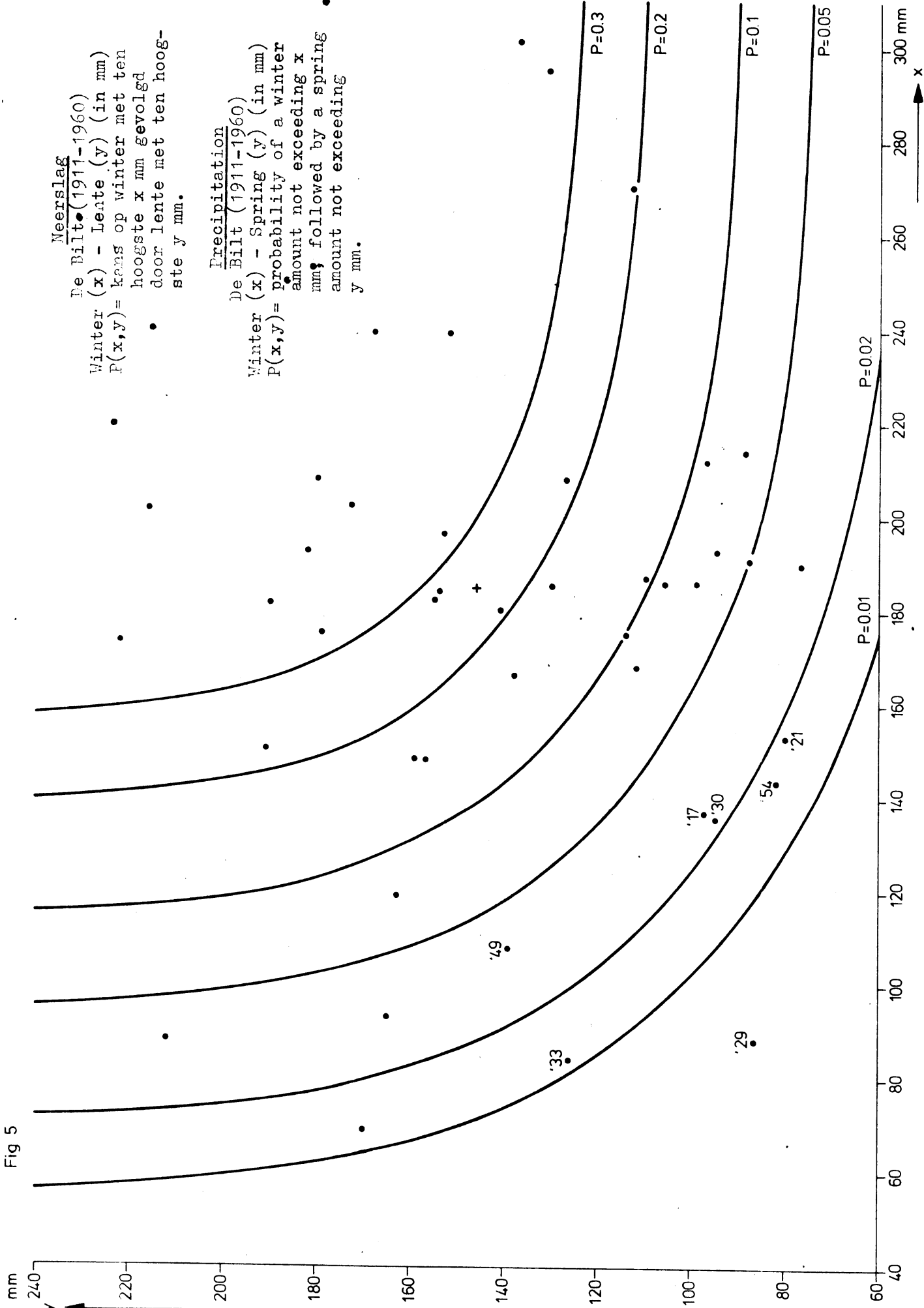
$\hat{\mu}$ = mean value

$\hat{\sigma}$ = standard deviation

n = number of observations

k = a number $(-\infty < k < +\infty)$

Fig 5



Neerslag

De Bilt (1911-1960)

Winter (x) - Lente (y) (in mm)

$P(x,y)$ = kans op winter met ten hoogste x mm gevolgd door lente met ten hoogste y mm.

Precipitation

De Bilt (1911-1960)

Winter (x) - Spring (y) (in mm)

$P(x,y)$ = probability of a winter amount not exceeding x mm, followed by a spring amount not exceeding y mm.

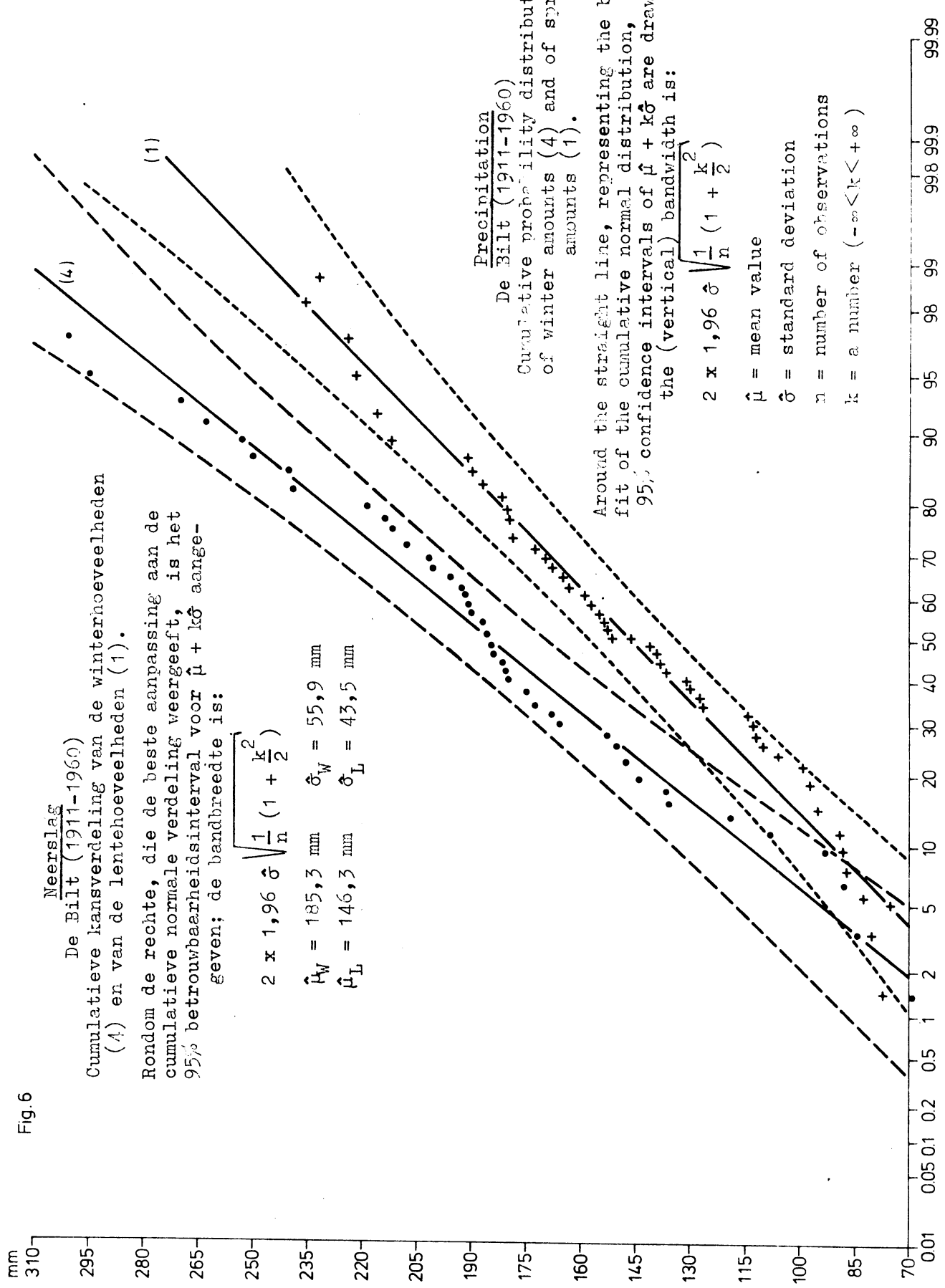


Fig. 6

Neerslag

De Bilt (1911-1960)

Cumulative kansverdeling van de winterhoeveelheden (4) en van de lentehoeveelheden (1).

Rondom de rechte, die de beste aanpassing aan de cumulatieve normale verdeling weergeeft, is het 95% betrouwbaarheidsinterval voor $\hat{\mu} + k\hat{\sigma}$ aangegeven; de bandbreedte is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$$\hat{\mu}_W = 185,3 \text{ mm} \quad \hat{\sigma}_W = 55,9 \text{ mm}$$

$$\hat{\mu}_L = 146,3 \text{ mm} \quad \hat{\sigma}_L = 43,5 \text{ mm}$$

Precipitation

De Bilt (1911-1960)

Cumulative probability distribution of winter amounts (4) and of spring amounts (1).

Around the straight line, representing the best fit of the cumulative normal distribution, the 95% confidence intervals of $\hat{\mu} + k\hat{\sigma}$ are drawn; the (vertical) bandwidth is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

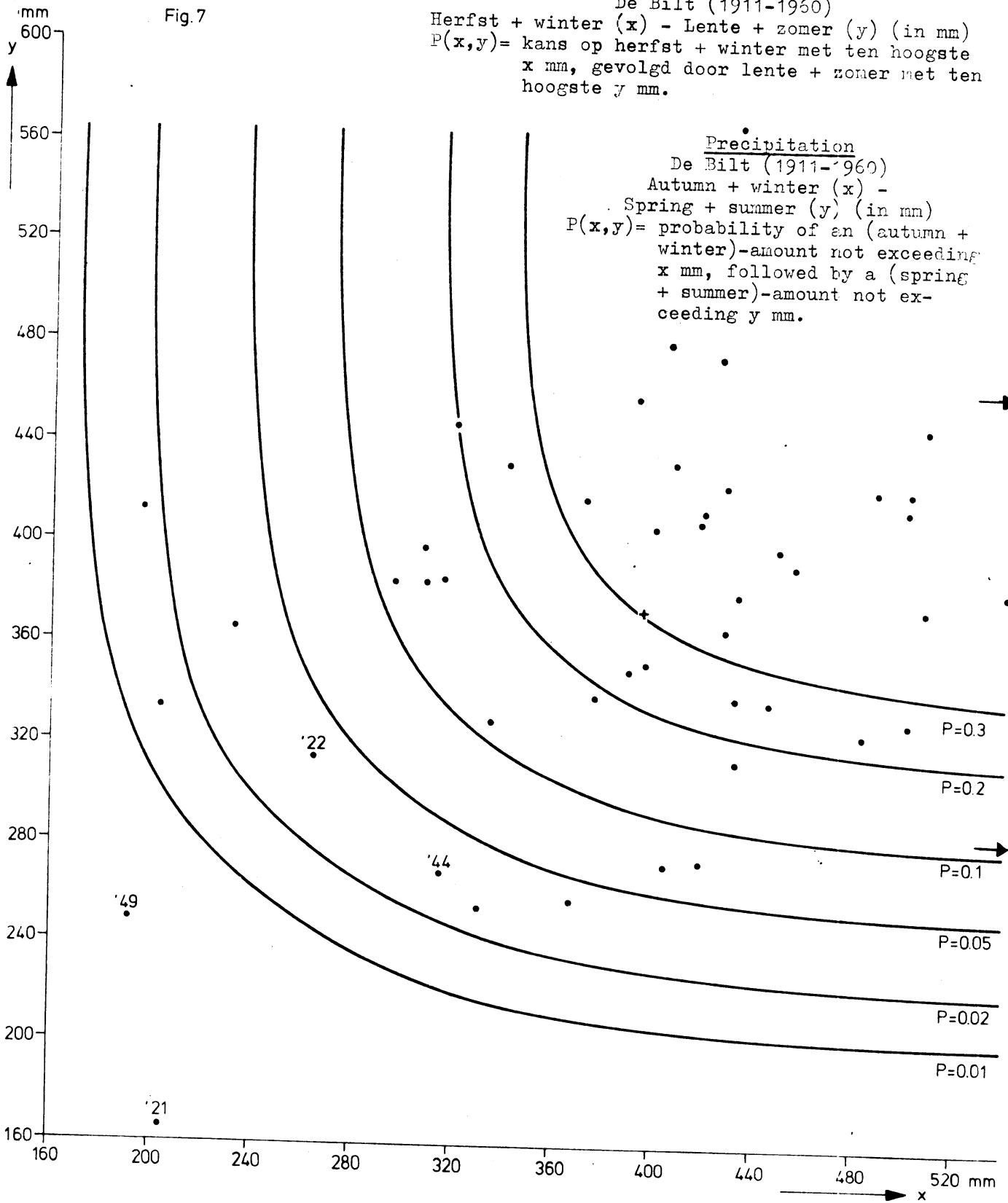
$\hat{\mu}$ = mean value

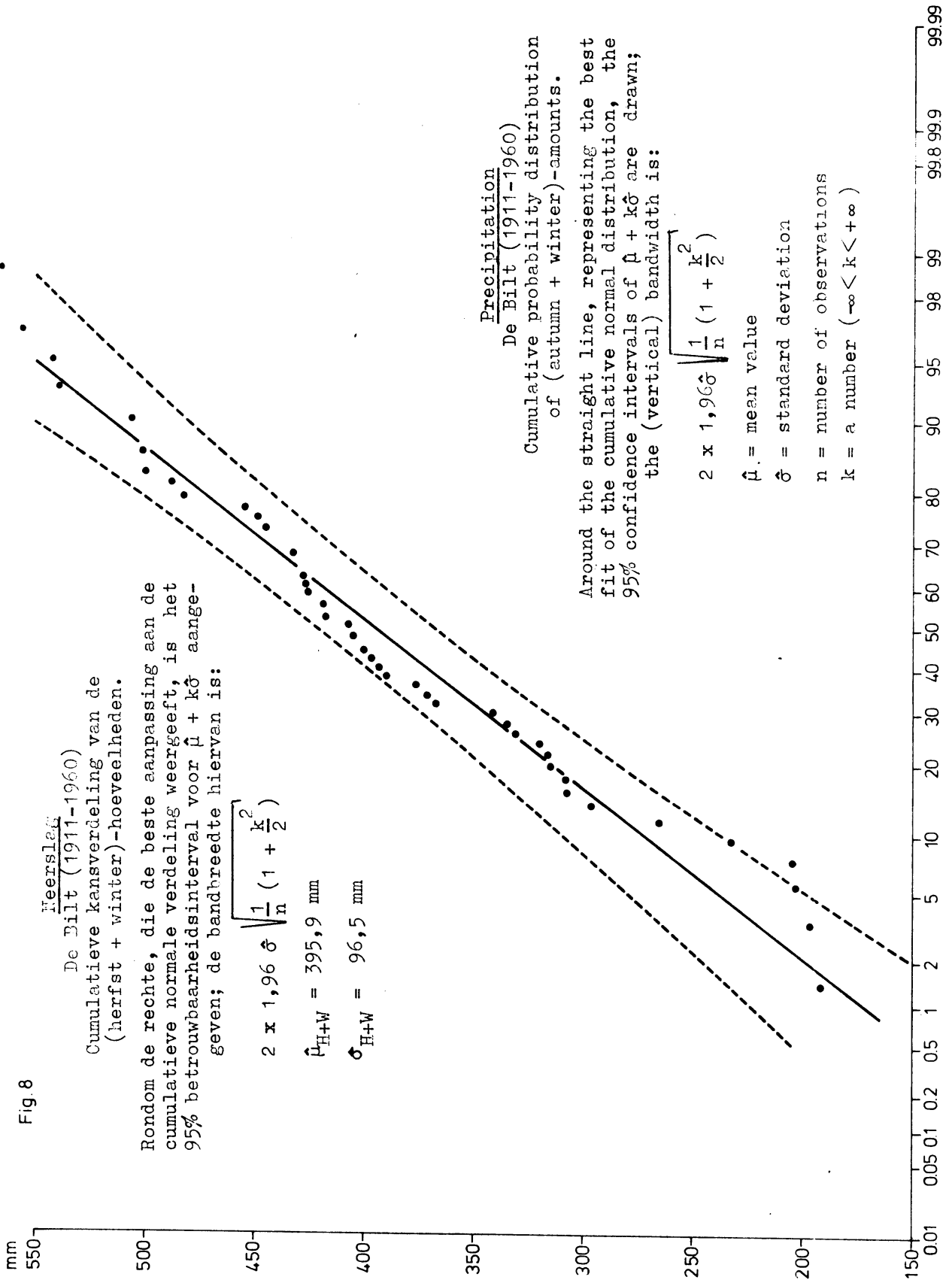
$\hat{\sigma}$ = standard deviation

n = number of observations

k = a number ($-\infty < k < +\infty$)

Fig. 7





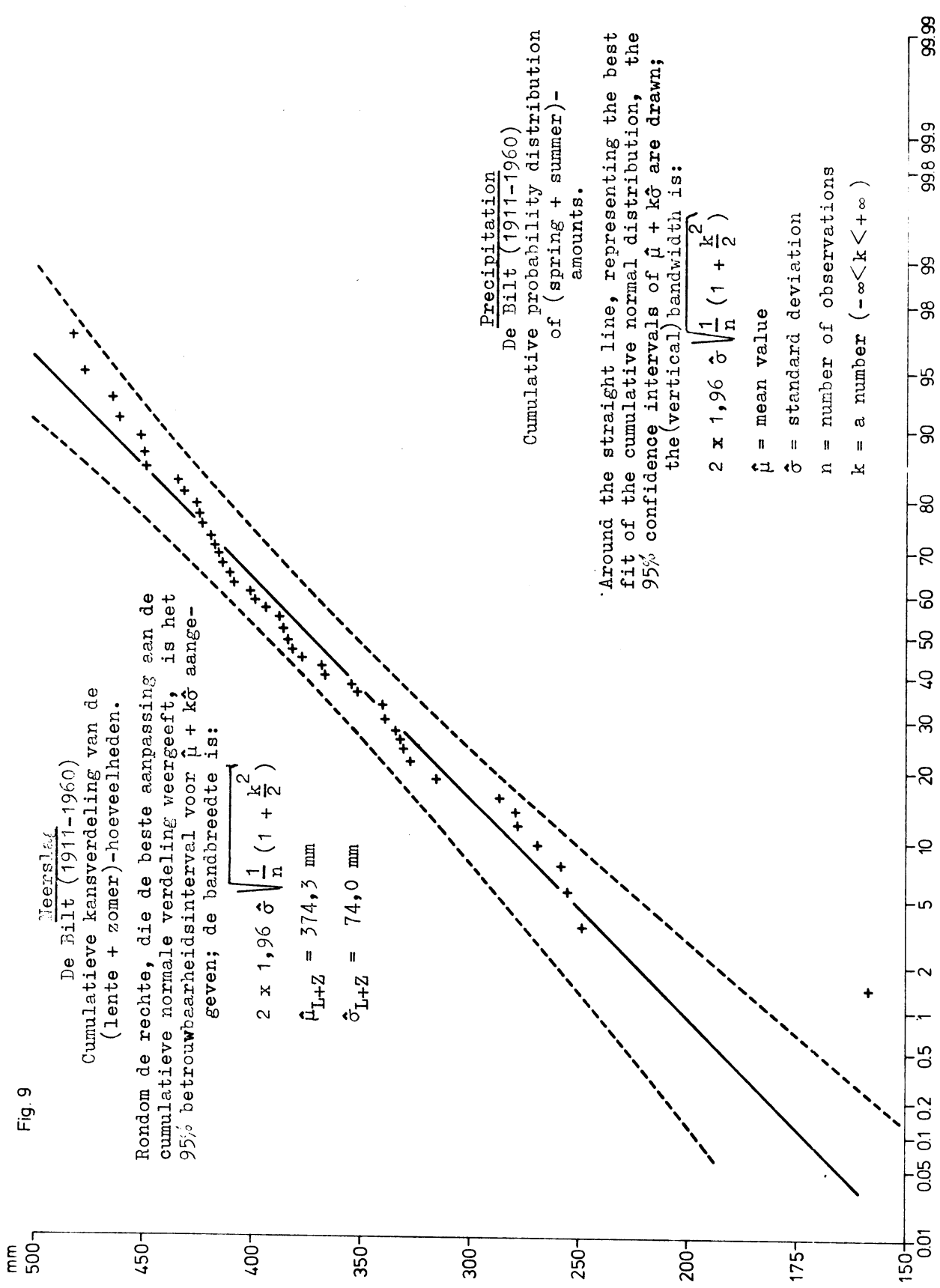


Fig. 9

Neerslag
 De Bilt (1911-1960)
 Cumulatieve kansverdeling van de
 (lente + zomer)-hoeveelheden.

Rondom de rechte, die de beste aanpassing aan de
 cumulatieve normale verdeling weergeeft, is het
 95% betrouwbaarheidsinterval voor $\hat{\mu} + k\hat{\sigma}$ aange-
 geven; de bandbreedte is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$$\hat{\mu}_{L+Z} = 374,3 \text{ mm}$$

$$\hat{\sigma}_{L+Z} = 74,0 \text{ mm}$$

Precipitation
 De Bilt (1911-1960)
 Cumulative probability distribution
 of (spring + summer)-
 amounts.

Around the straight line, representing the best
 fit of the cumulative normal distribution, the
 95% confidence intervals of $\hat{\mu} + k\hat{\sigma}$ are drawn;
 the (vertical) bandwidth is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

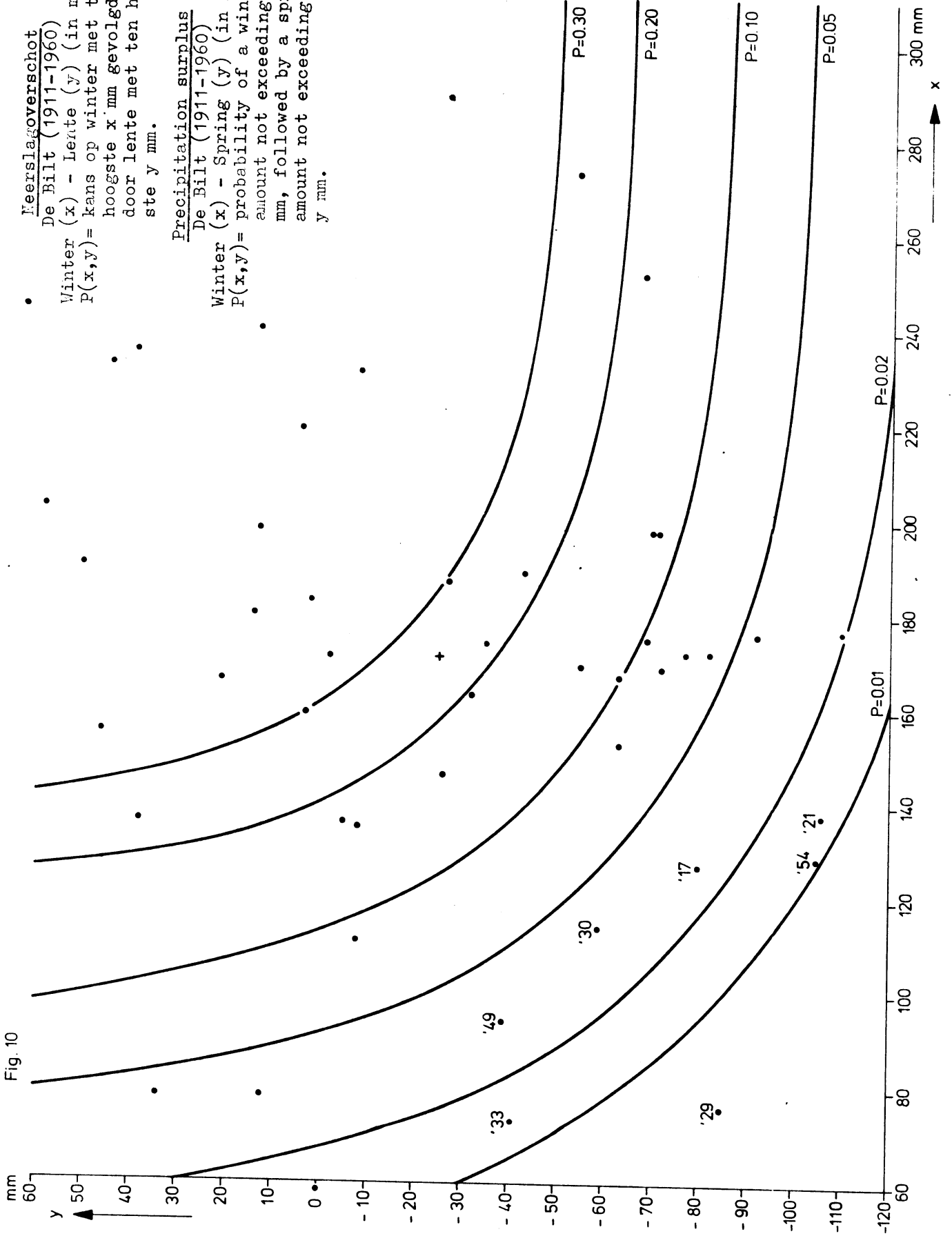
$\hat{\mu}$ = mean value

$\hat{\sigma}$ = standard deviation

n = number of observations

k = a number ($-\infty < k < +\infty$)

Fig. 10



Meerslagoverschot

De Bilt (1911-1960)

Winter (x) - lente (y) (in mm)
 $P(x,y)$ = kans op winter met ten hoogste x mm gevolgd door lente met ten hoogste y mm.

Precipitation surplus

De Bilt (1911-1960)

Winter (x) - Spring (y) (in mm)
 $P(x,y)$ = probability of a winter amount not exceeding x mm, followed by a spring amount not exceeding y mm.

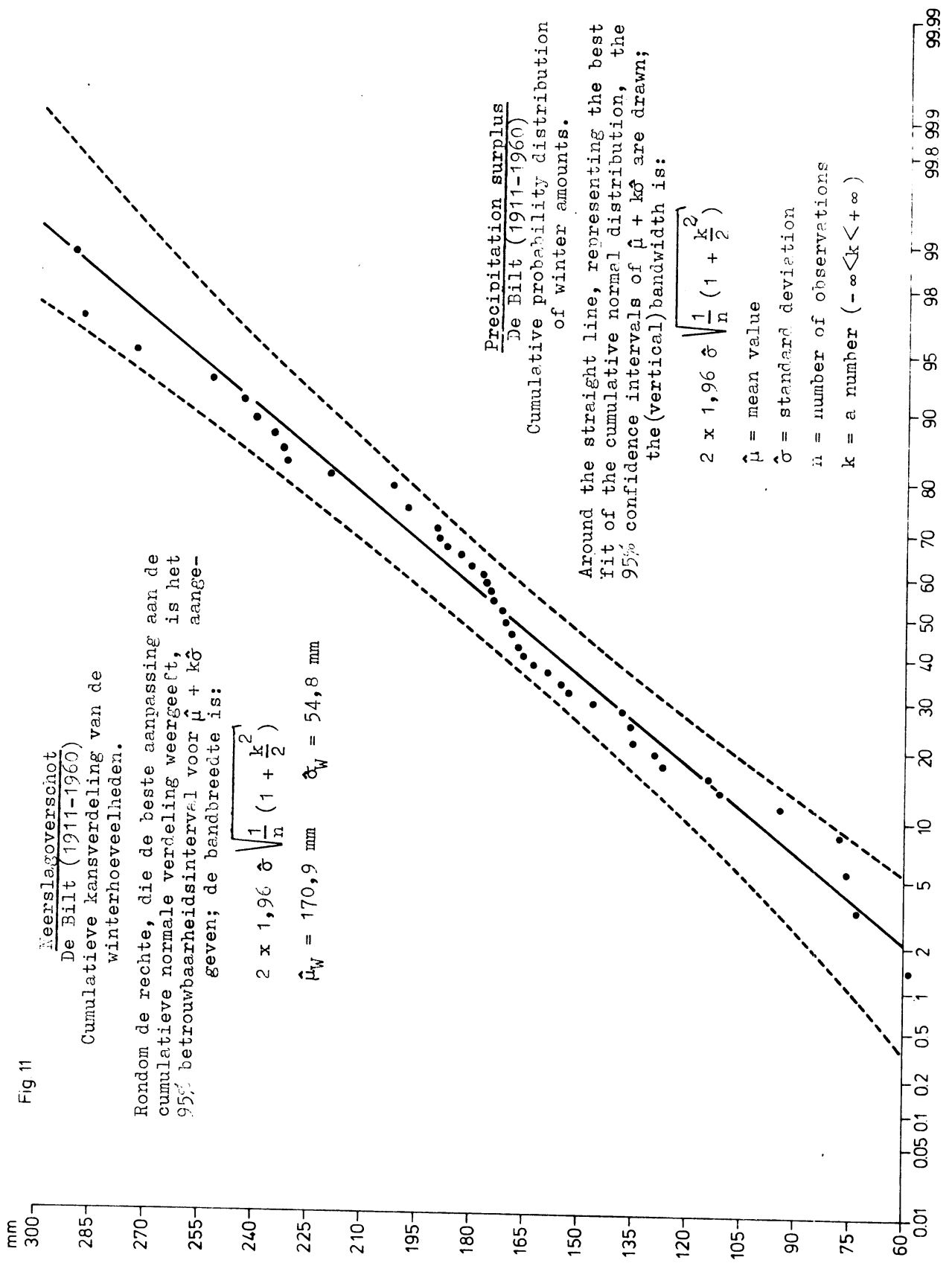
Fig 11

Neerslagoverschot
De Bilt (1911-1960)
 Cumulatieve kansverdeling van de
 winterhoeveelheden.

Rondom de rechte, die de beste aanpassing aan de
 cumulatieve normale verdeling weergeeft, is het
 95% betrouwbaarheidsinterval voor $\hat{\mu} + k\hat{\sigma}$ aange-
 geven; de bandbreedte is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$$\hat{\mu}_W = 170,9 \text{ mm} \quad \hat{\sigma}_W = 54,8 \text{ mm}$$



Around the straight line, representing the best
 fit of the cumulative normal distribution, the
 95% confidence intervals of $\hat{\mu} + k\hat{\sigma}$ are drawn;
 the (vertical) bandwidth is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$\hat{\mu}$ = mean value

$\hat{\sigma}$ = standard deviation

n = number of observations

k = a number $(-\infty < k < +\infty)$

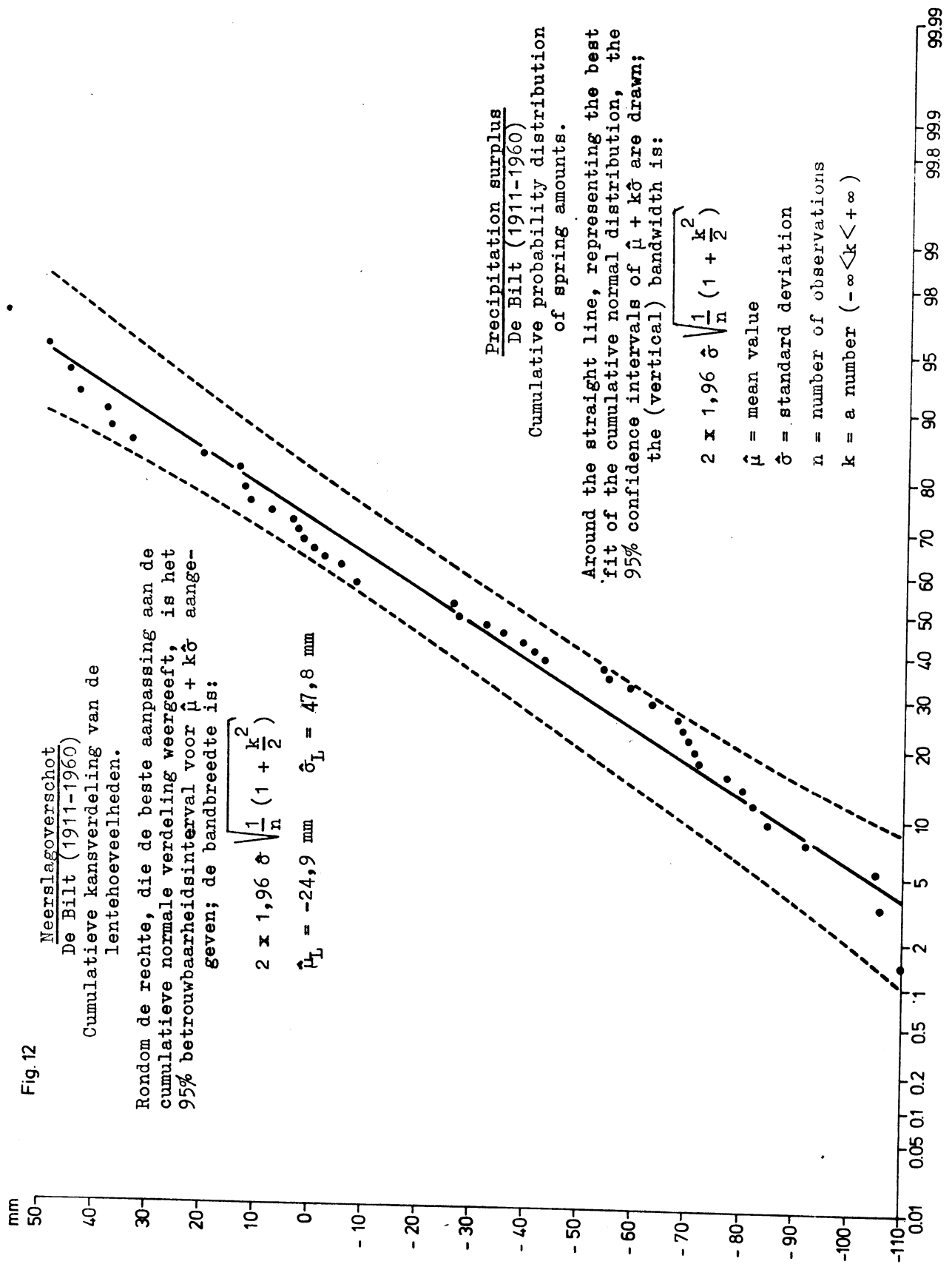


Fig. 12

Meerslagoverschot
 De Bilt (1911-1960)
 Cumulatieve kansverdeling van de
 lentehoeveelheden.

Rondom de rechte, die de beste aanpassing aan de
 cumulatieve normale verdeling weergeeft, is het
 95% betrouwbaarheidsinterval voor $\hat{\mu} + k\hat{\sigma}$ aange-
 geven; de bandbreedte is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$$\hat{\mu}_L = -24,9 \text{ mm} \quad \hat{\sigma}_L = 47,8 \text{ mm}$$

Precipitation surplus
 De Bilt (1911-1960)
 Cumulative probability distribution
 of spring amounts.

Around the straight line, representing the best
 fit of the cumulative normal distribution, the
 95% confidence intervals of $\hat{\mu} + k\hat{\sigma}$ are drawn;
 the (vertical) bandwidth is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$\hat{\mu}$ = mean value

$\hat{\sigma}$ = standard deviation

n = number of observations

k = a number $(-\infty < k < +\infty)$

Neerslagoverschot
 De Bilt (1911-1960)
 Herfst + winter (x) -
 Lente + zomer (y) (in mm).

$P(x,y)$ = kans op herfst + winter
 met ten hoogste x mm,
 gevolgd door lente + zomer
 met ten hoogste y mm.

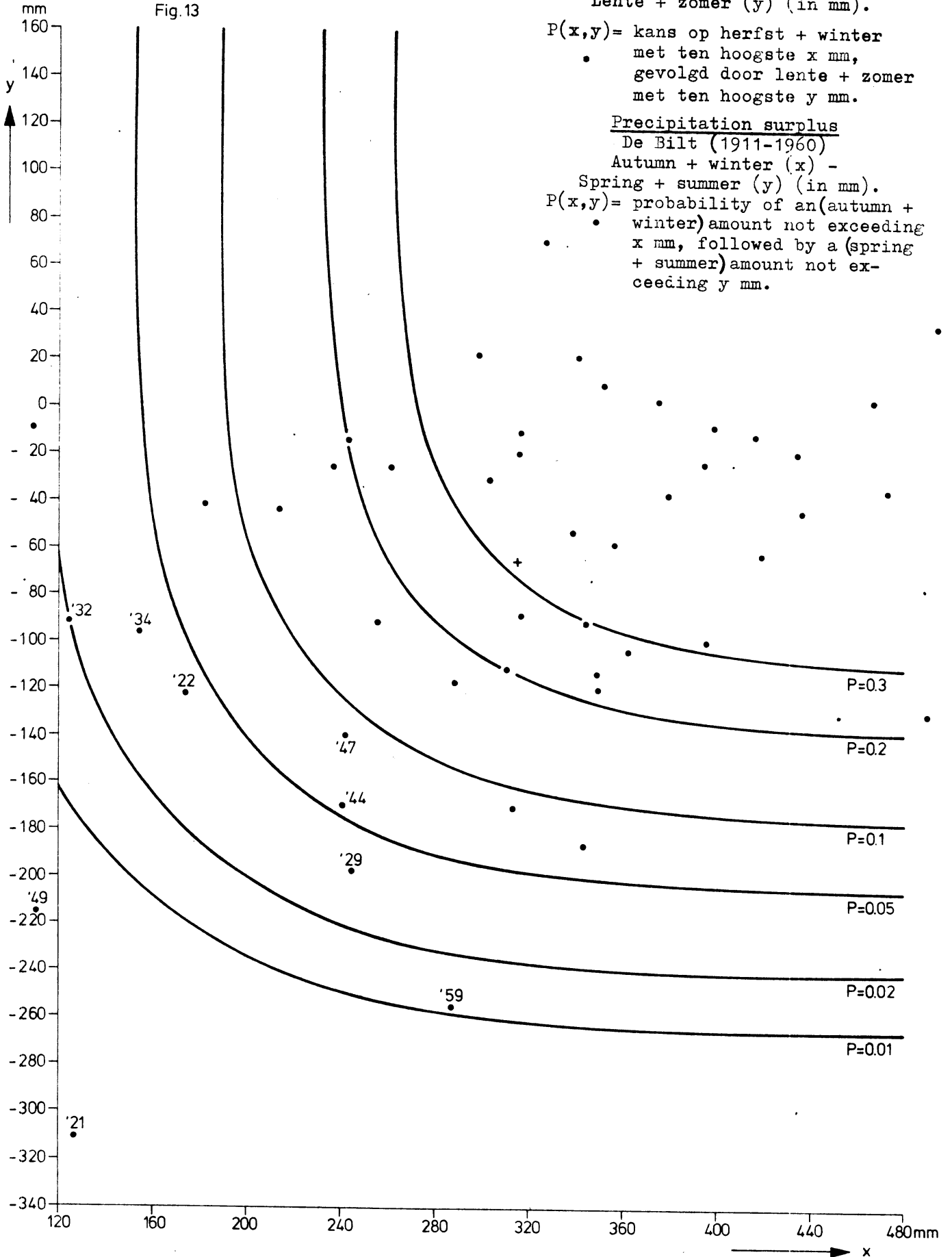
Precipitation surplus

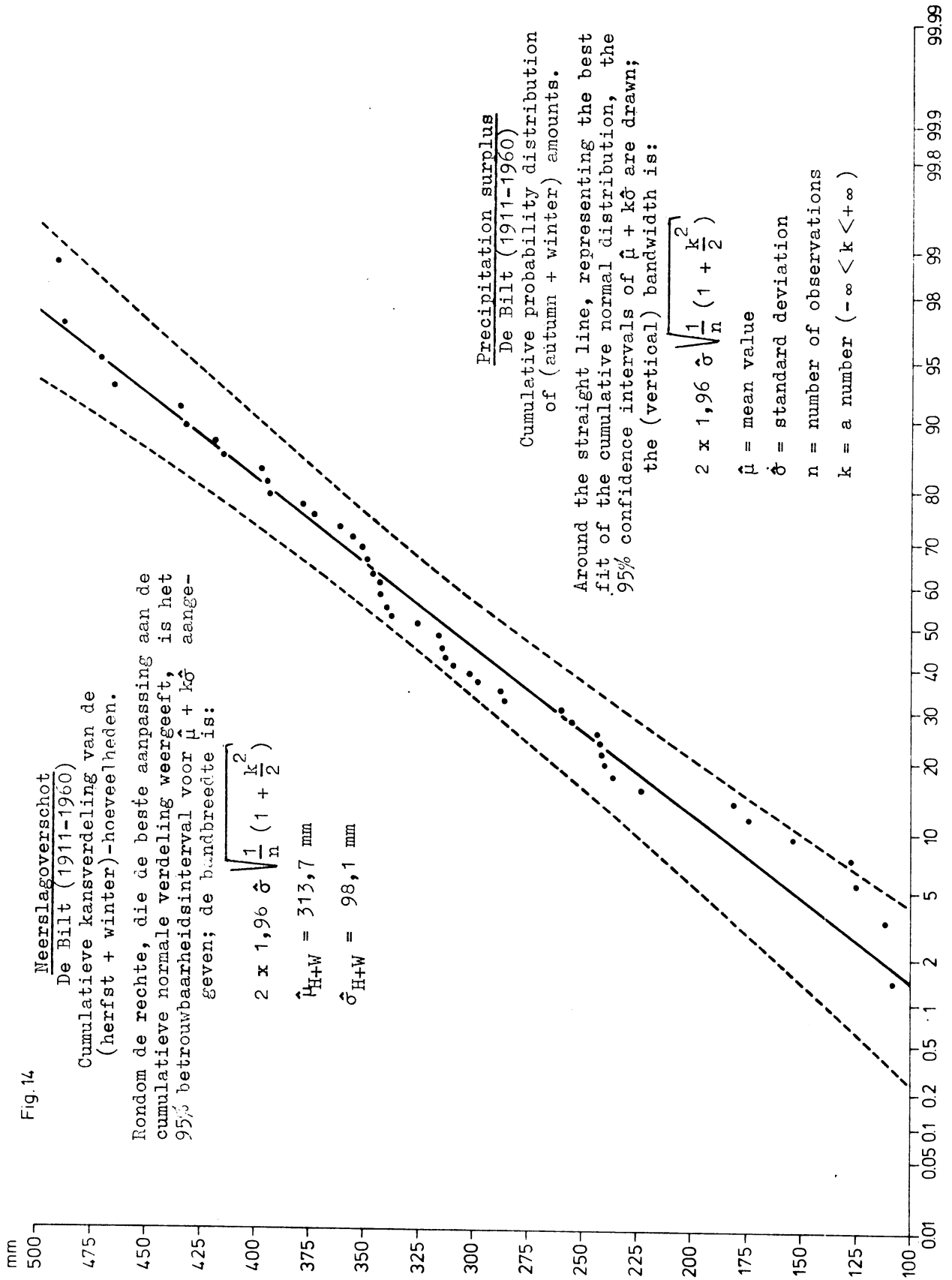
De Bilt (1911-1960)

Autumn + winter (x) -

Spring + summer (y) (in mm).

$P(x,y)$ = probability of an (autumn +
 winter) amount not exceeding
 x mm, followed by a (spring
 + summer) amount not ex-
 ceeding y mm.





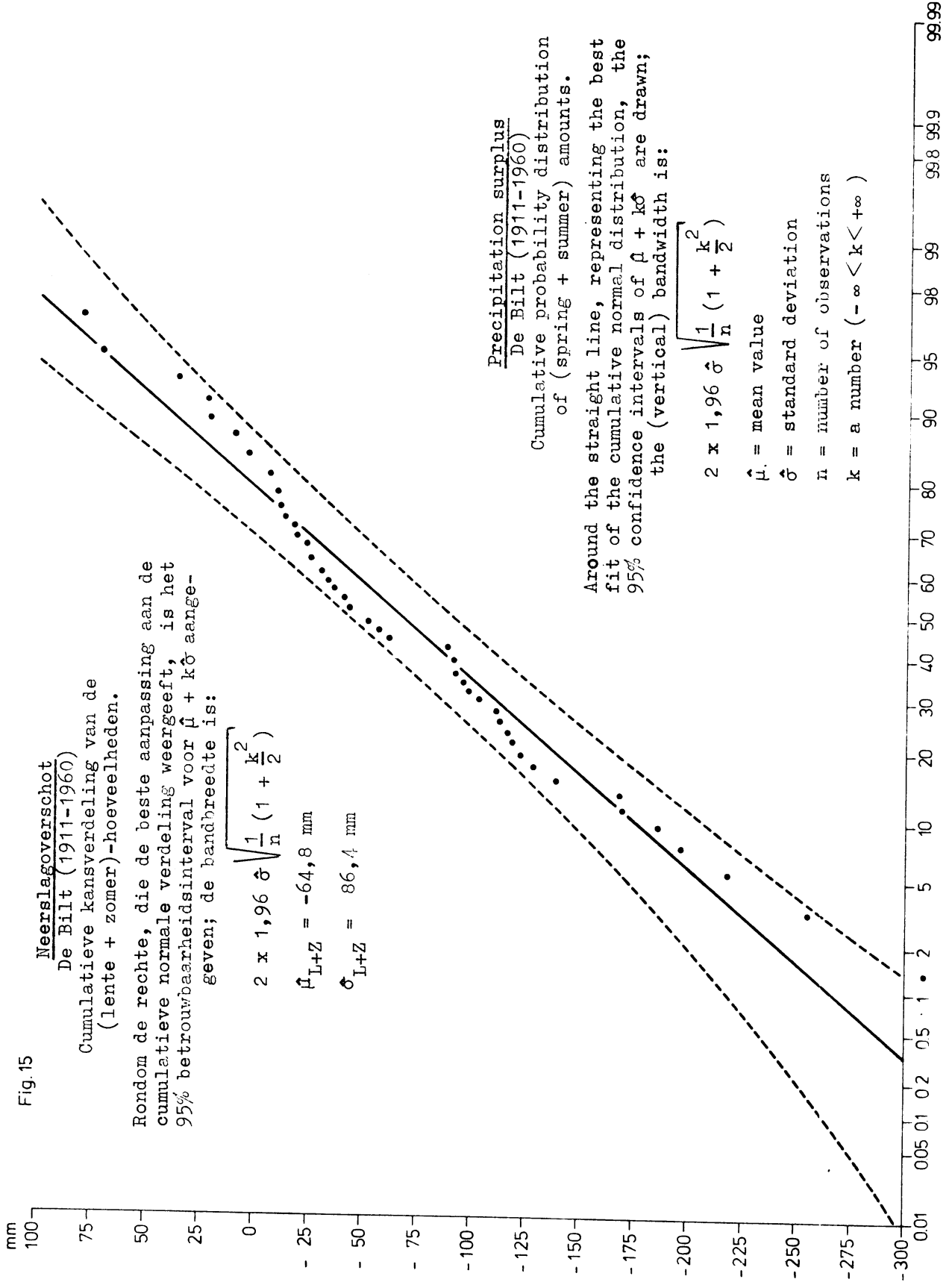


Fig. 15

Neerslagoverschot

De Bilt (1911-1960)

Cumulative kansverdeling van de (lente + zomer)-hoeveelheden.

Rondom de rechte, die de beste aanpassing aan de cumulatieve normale verdeling weergeeft, is het 95% betrouwbaarheidsinterval voor $\hat{\mu} + k\hat{\sigma}$ aangegeven; de bandbreedte is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$$\hat{\mu}_{L+Z} = -64,8 \text{ mm}$$

$$\hat{\sigma}_{L+Z} = 86,4 \text{ mm}$$

Precipitation surplus
De Bilt (1911-1960)

Cumulative probability distribution of (spring + summer) amounts.

Around the straight line, representing the best fit of the cumulative normal distribution, the 95% confidence intervals of $\hat{\mu} + k\hat{\sigma}$ are drawn; the (vertical) bandwidth is:

$$2 \times 1,96 \hat{\sigma} \sqrt{\frac{1}{n} \left(1 + \frac{k^2}{2}\right)}$$

$\hat{\mu}$ = mean value

$\hat{\sigma}$ = standard deviation

n = number of observations

k = a number $(-\infty < k < +\infty)$