KONINKLIJK NEDERLANDS
METEOROLOGISCH INSTITUUT

Wetenschappelijk Rapport W.R. 55-004 (III-150)

Dr C. Levert

# De persistentie van de dagelijkse hoeveelheid neerslag

De Bilt, 1955.

KONINKLIJK NEDERLANDS
METEOROLOGISCH INSTITUUT

Wetenschappelijk Rapport *)
W.R. 55-004 (III-150)
U.D.C.: 551.577.21 :
551.577.36 :
519.2

Dr C. Levert

# De persistentie van de dagelijkse hoeveelheid neerslag


| pag. | | INHOUD |
|---|---|---|
| 2 | 0 | Inleiding; overwegingen |
| 3 | 1 | Theoretische beschouwingen; afleiding van formules |
| 3 | 1.1 | Berekening van de interne (auto-)correlatiecoëfficiënten met behulp van de dagelijkse hoeveelheden neerslag in één maand |
| 6 | 1.2 | Berekening van de interne (auto-)correlatiecoëfficiënten met behulp van de veeljarige dagsommen in één maand |
| 13 | 2 | Uitvoering in de praktijk |
| 19 | 3 | Numerieke resultaten |
| 19 | 3.1 | De persistentiecoëfficiënten c en r |
| 19 | 3.1.1 | Eerste oriëntatie |
| 21 | 3.1.2 | Exacte berekening van $c_1$ voor 7 Februari- en 8 Augustus-maanden |
| 23 | 3.1.3 | Berekening van $c_1$, $c_2$ ..... $c_5$ voor alle maanden uit alle jaren |
| 28 | 3.1.4 | Berekening van $C_1$, $C_2$, $C_3$ m.b.v. de veeljarige dagsommen |
| 28 | 3.1.5 | Berekening van de datum-persistentiecoëfficiënten r |
| 31 | 3.2 | De varianties en de standaarddeviaties |
| 34 | 4 | Enige voorbeelden van het gebruik van de persistentiecoëfficiënten |
| 34 | 4.1.1 | Per maand, voor de één-daagse hoeveelheden. Theorie |
| 34 | 4.1.2 | Per maand, voor de twee-daagse hoeveelheden. Theorie |
| 35 | 4.1.3 | Per maand, voor de vijf-daagse hoeveelheden. Theorie |
| 37 | 4.2.1 | Numerieke resultaten bij 4.1.1 |
| 37 | 4.2.2 | Numerieke resultaten bij 4.1.2 |
| 37 | 4.2.3 | Numerieke resultaten bij 4.1.3 |
| 38 | 5 | Betrouwbaarheidsmarges |
| 39 | 6 | Addendum |
| 39 | 6.1 | De normale waarde van de dagelijkse hoeveelheid per datum |
| 40 | 6.2 | Regressie |
| 42 | 6.3 | De relatie $\rho_i = (\rho_1)^i$ |
| 43 | 7 | Opsomming der problemen, waartoe dit onderzoek leidde |
| 45 | 8 | Summary |


*) Dit rapport werd ook ten behoeve van de leden van de Werkgroep Regenwaarnemingen van de Commissie voor Hydrologisch Onderzoek T.N.O. samengesteld.

0 INLEIDING; OVERWEGINGEN

Voor bepaalde doeleinden [1] wilden wij de interne correlatiecoëfficiënt (ook wel genoemd: persistentiecoëfficiënt, autocorrelatiecoëfficiënt) tussen de hoeveelheid neerslag op dag i en die op dag i + j numeriek kennen. Wij hebben getracht de berekening ervan voor het grootste deel machinaal te doen uitvoeren en wel zo, dat tevens gegevens, die voor verder onderzoek gewenst waren, geleverd konden worden. Welke gegevens bedoeld worden, wordt uiteengezet in 2.

Wij gingen uit van de geponste daghoeveelheden neerslag van het regenstation Hoofddorp der jaren 1867 t/m 1953. Een jaarlijkse gang in bedoelde persistentiecoëfficiënten mocht a priori niet uitgesloten geacht worden.

Bovendien overwogen wij, dat er twee betekenissen aan het begrip interne correlatiecoëfficiënt (afgekort tot c.c.) van bijv. de eerste orde gehecht kunnen worden. Bijgaande schema's I en II bedoelen deze en andere grootheden, die in de verdere beschouwingen besproken zullen worden, te verduidelijken.

De beschikking hebbende over n Januari-maanden (en wel steeds de dagen 1 t/m 26, om een reden, die in 2 toegelicht wordt) kunnen wij in elke daarvan de interne (auto-)correlatiecoëfficiënt tussen de 25 paren getallen $h_{p.i}$; $h_{p.i+1}$ ($i = 1, 2, \ldots 25$) berekenen. Zij deze $c_{p1}$ (de p slaat op het nummer van de maand; $p = 1, 2, \ldots n$; de 1 duidt

(1) op de eerste orde). Deze n waarden kunnen wij middelen: $\bar{c}_1 = \frac{1}{n} \sum_{p=1}^{n} c_{p1}$.

Men kan echter ook de correlatiecoëfficiënt tussen de n paren $h_{q.i}$; $h_{q.i+1}$ ($q = 1, 2, \ldots n$) berekenen, d.i. $r_{i.i+1}$ en wel voor $i = 1, 2, 3 \ldots 25$. Daarna kunnen wij deze 25 r-waarden middelen tot

(2) $$\bar{r}_1 = \frac{1}{25} \sum_{i=1}^{25} r_{i.i+1} .$$

---

1) Wij verwijzen naar 1° onze beschouwing (Jan. 1953)."Interne correlatiecoëfficiënten en voortschrijdende sommen", geschreven voor de leden van de Werkgroep Regenwaarnemingen; 2° een nota voor dezelfde werkgroep: "Het werken met voortlopende n-sommen", Jan. 1953; 3° punt A van het "Ontwerp-werkprogramma voor het onderzoek ten behoeve van de Werkgroep", bij het Interim-rapport van Mei 1953; 4° ons artikel: "The persistence of daily rainfall and the reliability of estimates of the probability distribution of rainfalls in periods of different lengths", verschenen in Geophysica 4 No. 4 p.180 (1955).

De statisticus, deze berekeningen uitvoerende, stelt zich onmiddellijk verschillende vragen:

(3) a. Is $\bar{c}_1$ gelijk aan $\bar{r}_1$? Zo niet, zijn er misschien voorwaarden te noemen onder welke $\bar{c}$ wel gelijk aan $\bar{r}$ wordt? Men bedenke, dat de c's de naam auto-correlatiecoëfficiënten verdienen; de r's niet.

b. Mag men binnen een maand $r_{1.2} = r_{2.3} = r_{3.4} \ldots\ldots = r_{25.26}$ onderstellen of is een maand misschien een te groot deel van het jaar?

c. Wanneer heeft men meer aan $\bar{c}_1$ en wanneer meer aan $\bar{r}_1$?

Alvorens een werkschema voor de tabelleermachine te ontwerpen was het gewenst het vraagstuk van de persistentiecoëfficiënten van de theoretische zijde te bestuderen. Dit doen wij in 1, waarna in 2 en 3 de numerieke aanpak volgt, in 4 de resultaten genoemd worden en in 5 het gebruik van de persistentiecoëfficiënten met voorbeelden toegelicht wordt.

## 1 THEORETISCHE BESCHOUWINGEN; FORMULES

### 1.1 Berekening van de autocorrelatiecoëfficiënten met behulp van de dagelijkse hoeveelheden uit één maand.

Beschouw in bijv. elke maand Januari de etmaalhoeveelheden der eerste 26 dagen (in 2 wordt uitgelegd waar deze 26 vandaan komt). Wij laten gemakshalve de index i (het nummer van de maand - zie schema I -) in $h_{i.j}$ ($j = 1, 2, \ldots\ldots 26$) weg.

De interne correlatiecoëfficiënt $c_1$ van de eerste orde in deze reeks van 26 h-waarden is gedefinieerd door:

(4)
$$c_1 = \frac{\sum_1^{25} h_i h_{i+1} - \frac{\sum_1^{25} h_i \sum_2^{26} h_i}{25}}{\sqrt{\left[\sum_1^{25} h_i^2 - \frac{\left(\sum_1^{25} h_i\right)^2}{25}\right]\left[\sum_2^{26} h_i^2 - \frac{\left(\sum_2^{26} h_i\right)^2}{25}\right]}} = \frac{\frac{\sum_1^{25} h_i h_{i+1}}{24} - \frac{\sum_1^{25} h_i \sum_2^{26} h_i}{24.25}}{\sqrt{v.v'}}$$

Er komt voor elke maand Januari (d.i. voor ieder nummer k) één $c_{k.1}$ en één stel $v_k$, $v_k^1$ ($k = 1, 2, \ldots\ldots n$) [1]. Hierbij is $v$ = variantie in de dagsommen $h_1, h_2 \ldots\ldots h_{25}$ en $v^1$ die in de dagsommen $h_2, h_3 \ldots\ldots h_{26}$ (variantie = kwadraat van standaarddeviatie).

De berekening van (4) zou sterk vereenvoudigd worden als wij $\sum_1^{25} h_i$ gelijk aan $\sum_2^{26} h_i$ mochten stellen en $v$ gelijk aan $v^1$. Wat is hiervan te vertellen?

Wanneer de jaarlijkse gang in de dagelijkse hoeveelheden binnen een maand niet markant is, althans verwaarloosd zou mogen worden, zal, hoewel maand voor maand (meestal) $\sum_1^{25} h_i \neq \sum_2^{26} h_i$ gemiddeld over vele jaren het verschil tussen de twee sommen mogen worden verwaarloosd (d.w.z. $\bar{h}_1 \approx \bar{h}_{26}$). En voor $v$ en $v^1$? Hoe verhouden zich de verwachtingswaarden $\mathcal{E}$ van $v$ en $v^1$? Er komt:

(5) $$24\,\mathcal{E}v = \mathcal{E}\left(\sum_1^{25} h_i^2\right) - \frac{1}{25}\mathcal{E}\left(\sum_1^{25} h_i\right)^2$$

(6) $$24\,\mathcal{E}v' = \mathcal{E}\left(\sum_2^{26} h_i^2\right) - \frac{1}{25}\mathcal{E}\left(\sum_2^{26} h_i\right)^2$$

---

1) Om exact te zijn: gegeven zij een tweedimensionale verzameling van N elementen (d.w.z. N paren $x_i$, $y_i$). De centra (middens) van de x- en y-universa zullen heten $\xi$ en $\eta$; de varianties $\sigma_x^2$ en $\sigma_y^2$. Bij iedere $x_i$ behoort slechts één $y_i$ en omgekeerd. De correlatiecoëfficiënt $\rho$ is dan gedefinieerd door

$$\rho = \frac{\sum_1^N (x_i-\xi)(y_i-\eta)}{\sqrt{\sum_1^N (x_i-\xi)^2 \sum_1^N (y_i-\eta)^2}} = \frac{\sum_1^N (x_i-\xi)(y_i-\eta)}{N\sigma_x\sigma_y} = \frac{\mathcal{E}x.y - (\mathcal{E}x)(\mathcal{E}y)}{\sigma_x . \sigma_y}$$

$\mathcal{E}$ betekent: verwachtingswaarde van of universumgemiddelde van. Als wij noch $\xi$ noch $\eta$ noch $\sigma_x$ noch $\sigma_y$ kennen, zijn

$\bar{x} = \frac{1}{N}\sum_1^N x_i$ en $v = \frac{1}{N-1}\sum_1^N (x_i-\bar{x})^2$ de beste schattingen van

$\xi$ resp. $\sigma_x^2$ en idem voor de y.

In de tekst is N = 25.

Het is mogelijk [2] de uitdrukkingen (5) en (6) te schrijven in de vorm

(7) $$24\mathcal{E}_V = 24\mathcal{E}_{V'} = 25\sigma^2 - \frac{1}{25}\sigma^2\left[25 + 2\cdot 24\rho_I + 2\cdot 23\rho_{II} + \dots\ 2\rho_{XXIV}\right] =$$

$$24\sigma^2 - \frac{2}{25}\sigma^2\left[24\rho_I + 23\rho_{II} + \dots\ \rho_{XXIV}\right]$$

Hierbij moet echter aan de volgende voorwaarden (waaraan in genoemde noot een mathematische formulering gegeven wordt) voldaan zijn: de jaarlijkse gang mag binnen de maand worden verwaarloosd;

(8) én voor de normale waarde van de dagelijkse hoeveelheid neerslag
én voor de variantie per datum-hoeveelheid
én voor de datum-interne correlatiecoëfficiënt van iedere orde.

---

2) Het exacte bewijs verloopt a.v.: het symbool $\mathcal{E}$ door de middelingsstreep vervangende, komt er $24\bar{\bar{v}} = \overline{h_1^2 + h_2^2 + \dots\ h_{25}^2} - \frac{1}{25}\overline{(h_1 + h_2 \dots\ h_{25})^2} = A - \frac{1}{25}B$; Eerst A. Als $\bar{p}$ = veeljarig (universum) -gemiddelde van de dagsom op de p-de Jan. en $\overline{p^2}$ = universum-gem. van het kwadraat van de hoeveelheid op p Jan., geldt $\sigma_p^2 = \overline{p^2} - (\bar{p})^2$, zodat

$A = \overline{h_1^2} + \overline{h_2^2} + \dots\ \overline{h_{25}^2} = (\sigma_1^2 + \sigma_2^2 + \dots\ \sigma_{25}^2) + (\bar{h}_1)^2 + (\bar{h}_2)^2 + \dots\ (\bar{h}_{25})^2$.

$B = \overline{h_1^1 + h_2^2 + \dots\ h_{25}^2} + 2(\overline{h_1h_2 + h_2h_3 + \dots\ h_{24}h_{25}}) + 2(\overline{h_1h_3 + h_2h_4 + \dots\ h_{23}h_{25}}) + \dots\ 2h_1 \cdot h_{25} =$

$(\sigma_1^2 + \sigma_2^2 + \dots\ \sigma_{25}^2) + (\bar{h}_1)^2 + (\bar{h}_2)^2 + \dots\ (\bar{h}_{25})^2 + 2\{\rho_{1.2}\,\sigma_1 \cdot \sigma_2 + \bar{h}_1\bar{h}_2 + \rho_{2.3}\,\sigma_2\,\sigma_3 + \bar{h}_2\bar{h}_3 + \dots\ \rho_{24.25}\,\sigma_{24}\,\sigma_{25} + \bar{h}_{24}\bar{h}_{25}\} + 2\{\rho_{1.3}\,\sigma_1\,\sigma_3 + \bar{h}_1\bar{h}_3 + \rho_{2.4}\,\sigma_2\,\sigma_4 + \bar{h}_2\bar{h}_4 + \dots\ \rho_{23.25}\,\sigma_{23}\,\sigma_{25} + \bar{h}_{23}\bar{h}_{25}\} + \dots\ 2\{\rho_{1.25}\,\sigma_1\,\sigma_{25} + \bar{h}_1\bar{h}_{25}\}$.

Hierin is bijv. $\rho_{1.2}$ de (universum) correlatiecoëfficiënt - d.i. de datum-persistentiecoëfficiënt van 1 op 2 Januari - tussen de paren $h_{i.1}$; $h_{i.2}$ met i = 1, 2, 3 ....., zodat (per definitie) $\rho_{1.2} = \dfrac{\overline{h_1h_2} - (\bar{h}_1)(\bar{h}_2)}{\sigma_1\,\sigma_2}$. Analoog zijn gedefinieerd $\rho_{2.3}$, $\rho_{3.4}$ enz. (alle van van eerste orde) en $\rho_{1.3}$, $\rho_{2.4}$, $\rho_{3.5}$ ..... (alle van tweede orde), enz. Wanneer wij onderstellen:

(zie vervolg op pag. 6)

(vervolg noot 2) van pag. 5)

$\sigma_i^2 = \sigma^2$, voor elke i; $\bar{h}_i = \xi$, voor elke i; $\rho_{1.2} = \rho_{2.3} = \ldots\ldots \rho_{24.25} = \rho_I$; $\rho_{1.3} = \rho_{2.4} = \ldots\ldots = \rho_{II}$ enz., dan is

$24\,\bar{v} = 24\sigma^2 + 24\,\xi^2 - \frac{2}{25}\,\sigma^2\,(24\,\rho_I + 23\,\rho_{II} + \ldots\ldots\,\rho_{XXIV}) - \frac{2}{25}$

$(24 + 23 + 22 + \ldots\ldots 1)\,\xi^2 =$

$24\,\sigma^2 - \frac{2}{25}\left\{24\,\rho_I + 23\,\rho_{II} + \ldots\,\rho_{XXIV}\right\}\sigma^2$ q.e.d.

---

Wij constateren dus $\mathcal{E}_v = \mathcal{E}_v{}^1$.

Tevens hebben wij gezien - zelfs als iedere jaarlijkse gang afwezig is -, dat de gemiddelde waarde van de variantie binnen de maand kleiner is dan de variantie van de hoeveelheid neerslag op dag i, want in (7) staat:

(9) $\mathcal{E}_v = \sigma^2 F$ met $F = 1 - \frac{2}{24.25}\,(24\,\rho_I + 23\,\rho_{II} + \ldots\ldots\,\rho_{XXIV}) \leqslant 1$

Veelal blijken $\rho_{IV} = \rho_V = \rho_{VI} = \ldots\ldots\ 0$ en $\rho_{III} < \rho_{II} < \rho_I < 0.5$,

(10) zodat $F \approx 1 - \frac{2}{25}\,(\rho_I + \rho_{II} + \rho_{III})$.

Vanzelfsprekend is $\mathcal{E}_v = \sigma^2$ als er geen persistentie is (alle $\rho$'s = 0) en $\mathcal{E}_v = 0$, als er de sterkst mogelijke persistentie is (alle $\rho$'s = 1). Het is natuurlijk de vraag of wij dit verschil tussen $\bar{v}$ en $\sigma^2$ significant kunnen aantonen (zie aan slot van 3.2), aangezien het zo klein is bij kleine waarden van de $\rho$'s (als $\rho_I = 0.25$; $\rho_{II} = 0.06$; $\rho_{III} = 0.01$ is $F \approx 0.98$), terwijl nog bovendien $\bar{v}$ en $\sigma^2$ berusten op een materiaal van een betrekkelijk klein aantal jaren.

## 1.2 Berekening van de interne correlatiecoëfficiënten via de veeljarige dagsommen.

Beschouw opnieuw in een reeks Januari-maanden de hoeveelheden neerslag der dagen 1 t/m 26 en hun n-jarige sommen $H_{1.1}$, $H_{1.2}$, $H_{1.3}$ ..... $H_{1.26}$ (zie schema II). Een tweede groep van n Januari-maanden levert de sommen $H_{2.1}$, $H_{2.2}$, $H_{2.3}$ ..... $H_{2.26}$ enz. Aldus voor m groepen van n jaren.

De variantie in de sommen $H_{k.1}$, $H_{k.2}$ t/m $H_{k.25}$ van rij k, zij $V_k$; die in de sommen $H_{k.2}$, $H_{k.3}$ t/m $H_{k.26}$ zij $V_k^1$. Wij kunnen vragen naar de verwachtingswaarde van $V_k$. Deze is

$$(11) \qquad \mathcal{E} V = \tau^2 - \frac{2}{24.25} \tau^2 \left[ 24\, P_I + 23\, P_{II} + \ldots\ldots P_{XXIV} \right]$$

Wij hebben (11) volkomen analoog aan (7) gevormd, doch moeten nu goed vaststellen wat $\tau^2$, $P_I$, $P_{II}$ enz. betekenen.

Zij $t_1^2$ de variantie in $H_{1.1}$, $H_{2.1}$, $H_{3.1}$ ..... $H_{m.1}$, dan is $\lim_{m \to \infty} t_1^2 = \tau_1^2$; analoog $t_2^2$ en $\tau_2^2$ enz. Wij onderstellen, dat $\tau_1^2 = \tau_2^2 = \ldots\ldots \tau_{26}^2$ $(= \tau^2)$

Aangezien iedere $H_{k.1}$ (k = 1, 2, 3 ..... m; zie schema II) de som is van n van elkaar onafhankelijke termen - n.l. de 1 Jan.-dagsommen -

(12) is $\tau^2 = n\, \sigma^2$

Verder is $R_{1.2}$ = corr. coëff. tussen de m paren $H_{k.1}$; $H_{k.2}$; k = 1,2 ..... m en $R_{2.3}$ die tussen de paren $H_{k.2}$; $H_{k.3}$ enz. Hun universumwaarden heten (met Griekse letters) $P_{1.2}$, $P_{2.3}$ enz. Weer onderstellen wij iedere jaarlijkse gang verwaarloosbaar klein, zodat alle eerste-orde c.c. gelijk $P_I$, alle tweede-orde c.c. gelijk $P_{II}$ enz. gezet mogen worden.

Verderop bewijzen wij, dat $P_I = \rho_I$; $P_{II} = \rho_{II}$ enz., zodat (11) geschreven kan worden als

$$(13) \qquad \mathcal{E} V = n\, \sigma^2 \left[ 1 - \frac{2}{24.25} (24\, \rho_I + 23\, \rho_{II} + \ldots\ldots \rho_{XXIV}) \right]$$

Men komt er licht toe te redeneren, dat men via een berekening van $\bar{v}$ en $\bar{V}$ (de n is bekend), zowel de $\sigma^2$ als de term $24\,\rho_I + 23\,\rho_{II} + \ldots\ldots \rho_{XXIV}$ met behulp van (9) en (13) zou kunnen berekenen. Tegen deze redenering lijkt ons niets in te brengen, maar de door vele auteurs gemaakte "fout" is deze, dat zij niet met de $\bar{V}$, doch met slechts één enkele V (de V behorende bij de beschouwde n jaren)werken. Men moet echter goed het verschil tussen V en $\bar{V}$ in het oog houden. Toch plaatsen wij "fout" tussen aanhalingstekens, daarmee aangevende dat het mogelijk is, dat de fout niet groot is. Wij vermoeden namelijk, dat de V's (in schema II

$V_1$, $V_2$ ..... $V_m$) lang zo sterk niet zullen uiteenlopen als de V's in schema I (anders gezegd: de standaarddeviatie in de V's is vermoedelijk zeer klein t.o.v. de gemiddelde V). Wij zouden eigenlijk de variabiliteit in de V-reeks moeten kunnen vergelijken met die in de v-reeks; dit is ons nog niet gelukt.

Een andere fout, die men diezelfde auteurs ziet maken, is deze, dat zij (13) lezen alsof $V = n\,\sigma^2$, d.w.z. zij stellen in (13) de factor [.....] gelijk aan 1 en doen dus alsof er geen autocorrelatie meer is tussen de veeljarige dagsommen. Het laatste is beslist fout.

In dit verband zij nog gewezen op de uitdrukking $\sigma^2(H_i - H_{i+1}) = 2\,n\,\sigma^2 \times (1 - \rho_{i.i+1})$. Wij kunnen n.l. ieder jaar het verschil $\Delta_i$ tussen $h_i$ en $h_{i+1}$ beschouwen; dan is de variantie van deze $\Delta$'s

$$\sigma^2(\Delta_i) = \mathcal{E}\Delta_i^2 - (\mathcal{E}\Delta_i)^2 = \mathcal{E}(h_i - h_{i+1})^2 - [\mathcal{E}(h_i - h_{i+1})]^2 =$$

$$\sigma_i^2 + \sigma_{i+1}^2 - 2\,\rho_{i.i+1}\,\sigma_i . \sigma_{i+1}$$

(zie ook schema I) en wel vanwege de definitie van $\rho_{i.i+1}$.

Als $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = ..... = \sigma^2$ (d.w.z. de "datumvariantie"
(14) is datumonafhankelijk) en $\rho_{1.2} = \rho_{2.3} = ..... = \rho_I$, geldt $\sigma^2(\Delta_i) = 2\,\sigma^2(1 - \rho_I)$, voor elke i.

Vervolgens beschouwen wij, zie schema II, vele groepen van n jaren, en in elke groep $\square_i = H_i - H_{i+1} = (h_{1.i} - h_{1.i+1}) + (h_{2.1} - h_{2.i+1}) + ..... (h_{n.i} - h_{n.i+1})$.

Aangezien de afzonderlijke termen in het rechter lid onafhankelijk zijn (de successieve Jan.-maanden zijn niet gecorreleerd), geldt voor de variantie in het universum van deze $\square$'s

$$\sigma^2(\square_i) = n\,\sigma^2(\Delta_i) = 2\,n\,\sigma^2(1 - \rho_I) \text{ , voor elke i.} \tag{15}$$

Met nadruk zij erop gewezen, dat met deze $\sigma^2(\square)$ <u>niet</u> bedoeld wordt de variantie $V(H_i - H_{i+1})$ in de 25 jaren $H_{1.j} - H_{1.j+1}$, met j = 1, 2, ..... 25, bijv. in de eerste groep (schema II), zelfs niet de gemiddelde waarde van de m verschillende varianties (in schema II zijn m groepen gedacht).

Wij zijn hierover weer zo uitvoerig omdat sommigen uit déze variantie $V(H_i - H_{i+1})$ - bovendien slechts aan één stel $H_{1.j}$-waarden

($j = 1, 2, 3 \ldots\ldots$) gemeten - de $\sigma^2$ berekenen, uitgaande van de onjuiste gedachte, dat $V(H_i - H_{i+1})$ gelijk $n\sigma^2$ zou zijn.

Teneinde de verschillen duidelijk te doen uitkomen leiden wij ook nog de uitdrukking voor de verwachtingswaarde van deze $V(H_i - H_{i+1})$ af. Wij doen dit opzettelijk eerst voor $v(h_i - h_{i+1})$, de variantie van de 25 paren $h_i - h_{i+1}$) ($i = 1, 2, 3 \ldots\ldots 25$) per maand.

Het is natuurlijk duidelijk, dat

(16) $v(h_i - h_{i+1}) = -2c_1\sqrt{v.v^1} + v + v^1$. Zie voor c, v en $v^1$ in (4).

Dan is ook

(17) $\mathcal{E}v(h_i - h_{i+1}) = -2\,\mathcal{E}\left\{c_1\sqrt{v.v^1}\right\} + \mathcal{E}v + \mathcal{E}v^1$. Wij hebben reeds bewezen, zie (7), dat $\mathcal{E}v = \mathcal{E}v^1$. Voor $\mathcal{E}\left\{c_1\sqrt{v.v^1}\right\}$ mag alleen

(18) dan $(\mathcal{E}c_1)(\mathcal{E}\sqrt{v\,v^1})$ geschreven worden als $\sqrt{v.v^1}$ ongecorreleerd is met c.

(19) Wanneer ook nog $\mathcal{E}\sqrt{v.v^1} = (\mathcal{E}\sqrt{v}).(\mathcal{E}\sqrt{v^1}) = \mathcal{E}v$ zou gelden (zie noot blz 12 ), zou (16) geschreven mogen worden als

(20) $\sigma^2(h_i - h_{i+1}) \equiv \mathcal{E}v(h_i - h_{i+1}) = 2(\mathcal{E}v)(1 - \mathcal{E}c_1) = 2(1 - \mathcal{E}c_1).$

$\sigma^2.\left\{1 - \frac{2}{24.25}(24\rho_I + 23\rho_{II} + \ldots\ldots \rho_{XXIV})\right\}$, waarbij gebruikt gemaakt is van (7).

Geheel analoog aan (20) vormen wij nu $\tau^2(H_i - H_{i+1}) \equiv \mathcal{E}V^2(H_i - H_{i+1})$, daarbij $\mathcal{E}c_1$ door $\mathcal{E}C_1$ vervangende; de $\sigma^2$ door $n\sigma^2$ en $\rho_I$ door $P_I$, $\rho_{II}$ door $P_{II}$ enz. De $\mathcal{E}C_1$ stelt voor de universumwaarde van de autocorrelatiecoëfficiënt van de eerste orde tussen de veeljarige dagsommen. Er komt:

(21) $$\tau^2(H_i - H_{i+1}) \equiv \mathcal{E}V(H_i - H_{i+1}) \equiv 2(1 - \mathcal{E}C_1)n\sigma^2\left\{1 - \frac{2}{24.25}\left(24P_I + 23P_{II} + \ldots\ldots P_{XXIV}\right)\right\}$$

Men vergelijke thans (21) met (15) en lette op het grote verschil.

Thans volgt het bewijs van $\mathcal{P}_I = \rho_I$; $\mathcal{P}_{II} = \rho_{II}$ enz., waarvan wij reeds in (13) gebruik maakten.

Wij toonden aan, dat $\sigma^2 (h_i - h_{i+1}) = 2\sigma^2 (1 - \rho_{i,\ i+1})$, wanneer $\sigma_i^2 = \sigma^2$ gesteld wordt, voor elke i (zie 14). Wij beschouwen weer in een reeks groepen, elk van n Januari-maanden, het verschil $\square_i$ tussen $H_i$ en $H_{i+1}$ (gegeven i), zie schema II, en vragen naar de variantie $\sigma^2 (\square_i)$ in het universum der $\square_i$'s. De bekende stelling $\sigma^2(a \pm b) = \sigma^2 (a) + \sigma^2 (b) \pm 2\rho\,\sigma(a)\,\sigma(b)$ toepassende (waarin $\rho$ = c.c tussen a en b), komt er

$$\sigma^2(\square_i) = \sigma^2(H_i) + \sigma^2(H_{i+1}) - 2\mathcal{P}_{i.i+1}\sigma(H_i)\sigma(H_{i+1}) \quad ;$$

maar $\sigma^2 (H_i) = n\, \sigma^2 (h_i)$ en $\sigma^2 (H_{i+1}) = n\, \sigma^2 (h_{i+1})$, omdat de $h_i$-waarden der verschillende jaren ongecorreleerd zijn. Zodat

(22) $$\sigma^2(\square_i) = 2n\sigma^2(1 - \mathcal{P}_{i.i+1})$$

Maar ook geldt $\sigma^2 (\square_i) = n\, \sigma^2 (h_i - h_{i+1})$, omdat ook de $h_i - h_{i+1}$ - waarden der verschillende jaren ongecorreleerd zijn, en dus $= 2n\sigma^2$.

(23) $(1 - \rho_{i,i+1})$. Thans (22) met (23) vergelijkende, zien wij $P_{i,i+1} = \rho_{i,i+1}$. Bijgevolg is ook $P_I = \rho_I$ ; analoog $P_{II} = \rho_{II}$, enz.

Het lukte ons niet om te bewijzen, dat $\mathcal{E}$c gelijk aan $\mathcal{E}$C zou zijn. Wij kwamen tot het vermoeden $\mathcal{E}$c = $\mathcal{E}$C vanwege een stelling, die R. Waldo Lewis en D.Mc. Intosh noemen in hun artikel "Some effects of the coherence of meteorological time series" (in Met. Mag. <u>81</u> 242 1952). Deze stelling luidt: "de autocorrelatiecoëfficiënt in de rij gemiddelden is gelijk aan het gemiddelde der autocorrelatiecoëfficiënten in alle rijen". Ter toelichting: zie ons schema I. Aangezien de c.c. niet verandert wanneer alle H's door n gedeeld worden (waardoor "n-jarige gemiddelden" verkregen worden) houdt de bewering van genoemde auteurs dus in, dat $C_1 = \frac{1}{n} \sum c_1$.

Zij leveren geen bewijs. Het is ons (ook?) niet gelukt de stelling te bewijzen. Het is natuurlijk duidelijk, dat $\mathcal{E}$C = $\mathcal{E}$c zodra de stelling juist is.

Tenslotte nog de vraag a, pg. 3 . Is er een relatie tussen c en r of misschien tussen $\gamma \equiv \mathcal{E}$c en $\rho \equiv \mathcal{E}$r? Het ligt o.i. voor de hand, dat ook de datum - $\sigma$'s en de persistentiecoëfficiënten van verschillende orden $\gamma_{i,\ i+j}$ ($i = 1, 2, \ldots\ldots; j = 1, 2, \ldots\ldots$) in deze betrekking zullen voorkomen. Het is ons tot nu toe niet mogen gelukken, zelfs niet voor het eenvoudige geval, dat de $\sigma$, de veeljarige-gemiddelde-dagsom en de persistentiecoëfficiënt datum-onafhankelijk zouden zijn, een relatie te vinden. Om toch in dit rapport zo volledig mogelijk te zijn, willen wij aangeven waar de moeilijkheden liggen.

Wij schrijven nogeens de uitdrukking voor de autocorrelatiecoëfficiënt van de eerste orde ($c_1$) in de eerste maand Januari uit (zie schema I en formule (4))

$$c_1 = \frac{\frac{1}{24}\sum_{1}^{25} h_i h_{i+1} - \frac{1}{24 \times 25}\sum_{1}^{25} h_i \sum_{2}^{26} h_i}{\sqrt{v . v'}}$$

Zodat $c_1 . \sqrt{v . v^1} = (h_1 . h_2 + h_2 . h_3 + \ldots\ldots h_{25} . h_{26}) / 24 -$

$$(h_1 + h_2 + \ldots\ldots h_{25})\ (h_2 + h_3 + \ldots\ldots h_{26})\ /\ 24.25$$

Vervolgens nemen wij aan weerszijden van het = teken, de verwachtingswaarde $\mathcal{E}$. Er komt:

(24) $\mathcal{E}(c_1 . \sqrt{v} . \sqrt{v^1}) = \frac{24}{25}\ \mathcal{E}(h_1 . h_2) - \frac{1}{24.25}\ \mathcal{E}\,[\ h_1 + (h_2 + h_3 + \ldots\ldots h_{25})$

$(h_2 + h_3 + \ldots\ldots h_{25}) + h_{26}\ ]$

Hierbij is reeds ondersteld $\mathcal{E} h_i . h_{i+1} = \mathcal{E} h_j . h_{j+1}$, onverschillig i en j; wij werken de tweede term in het rechter lid verder uit a.v.:

$\frac{1}{24.25}\ \mathcal{E}\,[\,h_1 . h_{26} + (h_1 h_2 + h_1 h_3 + \ldots\ldots h_1 h_{25}) +$

$(h_2 h_{26} + h_3 h_{26} + \ldots\ldots h_{25} h_{26}) + (h_2^2 + h_3^2 + \ldots\ldots h_{25}^2) +$

$2\,(h_2 h_3 + h_3 h_4 + \ldots\ldots h_{24} h_{25}) + 2\,(h_2 h_4 + h_3 h_5 + \ldots\ldots h_{23} h_{25}) + \ldots\ldots$

$\ldots\ldots 2\, h_2 h_{25}\ ]$ .

Eveneens onderstellende $\mathcal{E} h_i^2 = \mathcal{E} h_j^2$ en $\mathcal{E} h_i h_{i+p} = \mathcal{E} h_j h_{j+p}$, onverschillig de i, j en p, wordt de grote veelterm:

$$\frac{1}{24.25}\,\mathcal{E}\,[\,24\,h_1^2 + h_1h_{26} + 2.24\,h_1h_2 + 2.23\,h_1h_4 + \ldots\ldots + 2.2\,h_1h_{24} + 2\,h_1h_{25}\,]$$

Wij maken het rekenwerk eenvoudiger door $\sigma_i = \sigma = 1$ en $\mathcal{E} h_i = \xi = 0$ voor elke i te onderstellen, zodat $\mathcal{E} h_i^2 = \sigma^2 = 1$ en $\rho_{i,i+1} = (\mathcal{E} h_i h_{i+1} - \mathcal{E} h_i . \mathcal{E} h_{i+1}) / \sigma_i . \sigma_{i+1} = \mathcal{E} h_i . h_{i+1}$ (analoog $\rho_{i,i+p}$; $p \geq 2$). Dan komt er $\mathcal{E}(c_1 \sqrt{v}.\sqrt{v^1}) = \frac{25}{24}\rho_1$

$$\frac{1}{24.25}\,[\,24 + \rho_{25} + 2.24\,\rho_1 + 2.23\,\rho_3 + \ldots\ldots 2.2\,\rho_{23} + 2.\rho_{24}\,]$$ of

(25) $$\mathcal{E}(c.\sqrt{v}.\sqrt{v^1}) = 1 - \frac{23}{24.25})\rho_1 - [\,1 + 2\frac{23}{24}\rho_2 + 2\frac{22}{24}\rho_3 + 2\frac{21}{24}\rho_4 + \ldots\ldots 2\frac{2}{24}\rho_{23} + 2\frac{1}{24}\rho_{24} + \frac{1}{24}\rho_{25}\,] / 25$$

Verdere vereenvoudiging is helaas niet mogelijk, tenzij wij zouden kunnen bewijzen, dat c en $\sqrt{v.v^1}$ niet of zo goed als niet gecorreleerd zouden zijn. Als dan óók nog $\mathcal{E}\, v.v^1$ gelijk $\mathcal{E}\, v = 1 - \frac{2}{24.25} \times (24\,\rho_1 + 23\,\rho_2 + \ldots\ldots \rho_{24})$ (zie (7)) gezet zou mogen worden [1], zou gelden:

$$\mathcal{E}(c\,\sqrt{v}\,\sqrt{v^1}) = \mathcal{E} c.\,\mathcal{E}\sqrt{v.v^1} \qquad \text{en}$$

---

1) Wij hebben dit trachten te toetsen. In fig. 6 hebben wij voor 7 Februari- en 8 Augustusmaanden de $c_1$ tegen $\sqrt{v.v^1}$ uitgezet. Hoewel het materiaal klein is, komt het ons voor, dat $c_1$ en $\sqrt{v.v^1}$ niet gecorreleerd zijn.
Anders is het natuurlijk gesteld met $\sqrt{v}$ en $\sqrt{v^1}$, zie de punten in fig. 7. De waarden $\sqrt{v}$ en $\sqrt{v^1}$ per maand zijn "zwaar lineair" gecorreleerd. Het kan ook niet anders, want van de 26 getallen, waarop v berust, behoren er 25 stuks tot de 26 getallen, waarop $v^1$ berust. Wij hebben dus de indruk, dat ~~geen~~ grote fout gemaakt wordt wanneer wij $\mathcal{E}(c_1\,\sqrt{v.v^1})$ door $(\mathcal{E} c_1).(\mathcal{E}\, v)$ vervangen.

(26) $$\varepsilon_c = \frac{-\frac{1}{25} - \left[-577\,\rho_1 + 2.23\,\rho_2 + 2.22\,\rho_3 + \ldots\ldots 2.2\,\rho_{23} + 2.\,\rho_{24}\right]/24.25}{1 - \left[2.24\,\rho_1 + 2.23\,\rho_2 + 2.22\,\rho_3 + \ldots\ldots 2.2\,\rho_{23} + 2.\,\rho_{24}\right]/24.25}$$

(Hierin is $577 = 25^2 - 2.24$)

(27) Deze uitdrukking levert het vreemde resultaat $\varepsilon_c = \gamma = -\frac{1}{25}$ als alle $\rho$'s nul zijn )totale afwezigheid van correlatie van $h_i$ op $h_{i+1}$), terwijl wij nul zouden verwachten.

(28) En als alleen $\rho_1 \neq 0$ en dan is $\varepsilon_c = \gamma = [577\,\rho_1 - \frac{1}{25}]$ : $[1 - \frac{2}{25}\rho_1] \approx (\rho_1 - \frac{1}{25}) : (1 - \frac{2}{25}\rho_1) \approx \rho_1$ als $\rho_1 \gg \frac{1}{25}$

Meer is er niet van te zeggen. De afleidingen na de misschien niet geoorloofde onderstelling, genoemd direct na (25), dienen met veel scepsis beschouwd te worden.

## 2 UITVOERING IN DE PRAKTIJK

Volkomen machinaal werd (voor Hoofddorp 1867 - 1953) voor bijv. de maand Januari 1877 met behulp van de dagelijkse hoeveelheden een overzicht vervaardigd en afgedrukt, hetwelk in tabel 1 wordt weergegeven. In deze tabel ziet men in de eerste kolom de nummers der dagen 1 t/m 31. In de tweede kolom staan de dagelijkse hoeveelheden neerslag $h_i = y_{1i}$ ($i = 1, 2, \ldots\ldots 31$), in tienden mm. Zo werd te 8 uur van 20 Jan. 1877 afgetapt 3.5 mm. In de derde kolom staan de kwadraten $h_i^2$ ; in de vierde kolom worden de producten $h_i.h_{i+1}$ genoemd ($i = 1, 2, \ldots\ldots 25$).

Daarna volgt een kolom met voortschuivende tweedaagse sommen $y_{2i}$ ($i = 1, 2, 3, \ldots\ldots 25$). Zo is voor $i = 1$ $y_{21} = h_1 + h_2 = 0{,}2 + 14{,}4 = 14{,}6$ mm en voor $i = 2$ $y_{22} = h_2 + h_3 = 14{,}4 + 5{,}5 = 19{,}9$ mm enz. De kolom daarnaast geeft de voortschuivende driedaagse sommen $y_{3i}$ ($i = 1, 2, \ldots\ldots 25$). Zo is $y_{35} = h_5 + h_6 + h_7 = 1{,}0 + 3{,}4 + 1{,}7 = 6{,}1$ mm; dan volgen ook nog kolommen met $y_{4i}$, $y_{5i}$, $y_{6i}$ en $y_{7i}$. 1). Hiermede was de machine aan het einde van haar capaciteit.

1) Deze k-daagse sommen (k = 1, 2, ..... 7) waren natuurlijk niet direct nodig voor de berekening van de persistentiecoëfficiënten. Zij werden echter gelijk met de berekening der autocorrelatiecoëfficiënten berekend (en in nieuwe kaarten geponst) omdat de in noot 1) blz. 2 , bedoelde werkgroep óók om een frequentieverdeling van meerdaagse hoeveelheden vroeg. Al deze bewerkingen werden op het K.N.M.I. uitgevoerd met een Bull-tabelleermachine (B.S. 120), met aangekoppelde totaalkaartenmachine.

Zulk een tabel werd afgedrukt voor elke maand Januari. Langs de onderzijde staan de sommen der kolommen. In het bijzonder interesseren ons de totalen $S = \sum_1^{25} h_i$ ; $K = \sum_1^{25} h_i^2$ en $P_1 = \sum_1^{25} h_i . h_{i+1}$ .

Ook de producten $h_i . h_{i+1}$ met $j = 2, 3, 4$ en $5$ en $i = 1, 2, 3,$ .. ..... 25 werden gemaakt en afgedrukt, zie tabel 2. Wij zullen gebruik maken van de kolomtotalen $P_2$, $P_3$, $P_4$ en $P_5$.

N.B. Het feit, dat, althans in een maand van 31 dagen (als Jan. 1877), de 5 tweedaagse sommen $h_{26} + h_{27}$; ..... $h_{30} + h_{31}$; de 4 driedaagse sommen $h_{26} + h_{27} + h_{28}$; ..... $h_{29} + h_{30} + h_{31}$; de 3 vierdaagse sommen $h_{26} + h_{27} + h_{28} + h_{29}$; ..... $h_{28} + h_{29} + h_{30} + h_{31}$; de 2 vijfdaagse sommen $h_{26} + \ldots\ldots h_{30}$; $h_{27} + \ldots\ldots h_{31}$ en de ene zesdaagse som $h_{26} + \ldots\ldots h_{31}$ ontbreken [1] heeft een zuiver machine-technische oorzaak, die niet in een paar woorden uit te leggen is. Het zal duidelijk zijn hoe zulke tabellen er voor maanden van 30, 29 en 28 dagen uitzien. Tevens blijkt, dat wij alleen met compleet aanwezige maanden konden werken; zodra er ook maar één dag is, waarop niet werd afgetapt, rijzen er moeilijkheden. Zo waren wij genoodzaakt alleen de hiaat-vrije maanden te bewerken (later en wel voor de berekening van k-daagse sommen met $k \geqslant 30$ bleek het toch nodig tot het alsnog invullen van "open dagen" te besluiten en wel aan de hand van de aftappingen in naburige stations).

Tenslotte hebben wij ook nog door de machine laten uitrekenen de som over 84 Januari-maanden van $h_1$, $h_2$, $h_1 . h_2$; $h_{15}$, $h_{16}$, $h_{15} . h_{16}$ en deze ook voor de 83 Februari-maanden, de 80 Maart-maanden, enz. Zie tabel 3.

---

1) Echt ontbreken doen zij niet; de machine kon zo worden geschakeld, dat genoemde k-daagse sommen toch "geboekt" en in ponskaarten gebracht konden worden, zodat zij in de totale frequentieverdeling zeker meedoen. De frequentieverdeling der vijfdaagse sommen in Januari berust dus op 84 x 27 = 2268 getallen, aangezien wij konden beschikken over 84 Januari's en elke Januari 27 vijfdaagse sommen bevat.

Tabel 3

| | aantal jaren | $\sum x_1$ mm | $\sum x_2$ mm | $\sum x_1 x_2$ mm$^2$ | $\sum x_{15}$ mm | $\sum x_{16}$ mm | $\sum x_{15} \cdot x_{16}$ mm$^2$ | $\sum x_1^2$ mm$^2$ | $\sum x_2^2$ mm$^2$ | $\sum x_{15}^2$ mm$^2$ | $\sum x_{16}^2$ mm$^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Jan. | 84 | 155.9 | 243.1 | 695.14 | 129.3 | 140.1 | 381.84 | 1212.95 | 2695.15 | 916.61 | 1424.87 |
| Feb. | 83 | 155.0 | 175.5 | 538.05 | 93.5 | 151.6 | 460.73 | 1326.80 | 1287.59 | 694.89 | 1208.48 |
| Mrt. | 80 | 149.6 | 134.5 | 419.55 | 126.0 | 84.4 | 250.01 | 1041.20 | 645.79 | 711.18 | 395.18 |
| Apr. | 83 | 103.6 | 82.2 | 156.72 | 107.7 | 85.1 | 169.12 | 737.28 | 404.38 | 583.83 | 371.95 |
| Mei | 80 | 127.5 | 152.7 | 892.79 | 119.2 | 121.0 | 342.31 | 1022.02 | 1625.64 | 1012.01 | 1194.34 |
| Juni | 84 | 156.8 | 167.6 | 785.92 | 115.9 | 165.8 | 295.35 | 1446.26 | 1955.86 | 716.99 | 1160.42 |
| Juli | 81 | 185.9 | 147.0 | 642.50 | 236.4 | 217.8 | 900.64 | 2134.77 | 1263.22 | 3808.14 | 3244.90 |
| Aug. | 84 | 160.1 | 201.1 | 768.05 | 195.3 | 187.8 | 1266.47 | 1555.01 | 2663.49 | 2719.39 | 2935.36 |
| Sep. | 83 | 161.8 | 248.4 | 943.75 | 178.7 | 189.1 | 664.95 | 1005.14 | 3149.86 | 1805.27 | 2101.29 |
| Oct. | 81 | 197.6 | 190.7 | 962.69 | 252.4 | 336.9 | 2777.61 | 2072.50 | 2388.53 | 2884.66 | 5752.47 |
| Nov. | 79 | 139.9 | 220.8 | 720.10 | 206.2 | 213.5 | 813.52 | 963.35 | 2017.48 | 1846.80 | 2564.73 |
| Dec. | 81 | 131.8 | 169.5 | 518.20 | 152.5 | 168.6 | 436.09 | 891.40 | 1342.27 | 928.87 | 1721.32 |

Wanneer wij noemen $S = h_1 + h_2 + \ldots\ldots h_{25}$; $S^1 = h_2 + h_3 + \ldots\ldots h_{26}$; $S^{11} = h_3 + h_4 + \ldots\ldots h_{27}$, enz. $K = h_1^2 + h_2^2 + \ldots\ldots h_{25}^2$ en $K^1 = h_2^2 + h_3^2 + \ldots\ldots h_{26}^2$; $K^1 = h_3^2 + h_4^2 + \ldots\ldots h_{27}^2$, enz. (althans voor een maand van 31 dagen) willen wij de volgende uitdrukkingen berekenen:

a) per individuele maand

(29) serie-persistentiecoëfficiënt (autocorrelatiecoëfficiënt) van eerste orde =

$$c_1 = \frac{P_1 - \frac{1}{25} S.S^1}{\sqrt{(K - \frac{1}{25} S^2)\,(K^1 - \frac{1}{25} S^{1\,2})}}$$

en de

(30) serie-persistentiecoëfficiënt (autocorrelatiecoëfficiënt) van de tweede orde =

$$c_2 = \frac{P_2 - \frac{1}{25} S.S^{11}}{\sqrt{(K - \frac{1}{25} S^2)\,(K^1 - \frac{1}{25} S^{11\,2})}}$$

Idem $c_3$, $c_4$, $c_5$ enz.

Helaas levert de machine ons niet $S^1$, $S^{11}$ enz. en $K^1$, $K^{11}$ enz. Het is echter duidelijk, dat $S^1 = S + h_{26} - h_1$ en $K^1 = K + h_{26}^2 - h_1^2$, zodat wij zelf (en dat elke maand) de $S^1$ en $K^1$ kunnen berekenen. Vervolgens kunnen wij de 84 $c_1$-waarden rekenkundig middelen (als bewezen is, dat zij niet significant verschillen) tot een nieuwe $\bar{c}$ (eigenlijk zou dit "middelen" via de Fisher-se z-transformatie moeten gebeuren, maar veel verschil maakt dit niet, althans niet bij zulk een groot aantal correlatiecoëfficiënten, wier absolute waarde meestal niet dicht bij één gelegen is)

b) per individuele maand de benadering

(31) $$c_1^{*} = \frac{P - \frac{1}{25} S^2}{K - \frac{1}{25} S^{1\,2}}$$ ; idem $c_2^{*}$, enz. Wij hebben dus in $c_1$ $S^1$ door S en $K^1$ door K vervangen. Over vele maanden gemiddeld is zeker $K^1 \cong K$ en $S^1 \cong S$, omdat $\bar{h}_{26} \cong h_1$, althans als wij de jaarlijkse gang binnen de maand mogen verwaarlozen (en zo doen wij immers steeds)

c) voor alle Januari-maanden tezamen de benadering

(32) $$\hat{c}_1 = \frac{\sum\limits_1^{84} P - \frac{1}{25} \sum\limits_1^{84} S^2}{\sum\limits_1^{84} K - \frac{1}{25} \sum\limits_1^{84} S^2} = \frac{\bar{P}_1 - \frac{1}{25} \overline{S_1^2}}{\bar{K}_1 - \frac{1}{25} \overline{S_1^2}}$$ ; idem $\hat{c}_2$ enz.

Wat stelt deze $\hat{c}$ voor? Dit is moeilijk te zeggen. Wij voeren hem in omdat hij zo gemakkelijk berekend kan worden. Men slaat namelijk de berekening per maand over, maar werkt regelrecht met de totaalsommen $\sum P$, enz. (of de gemiddelde $P_1$, de gemiddelde $S^2$, de gemiddelde K, wat op hetzelfde neerkomt). Men houde echter in het oog, dat zeker

(33) $$\frac{1}{84} \sum_1^{84} c^{*} = \overline{c^{*}} \neq \hat{c}$$ is en wel vanwege de stelling: "het gemiddelde der quotiënten is niet gelijk aan het quotiënt der gemiddelden" (Zie ons K.N.M.I.-rapport R III 90 1952)

d) Wij kunnen de Januari-dagen chronologisch achter elkaar plaatsen, d.w.z. $h_1$, $h_2$, $h_3$ ..... $h_{25}$ van Januari 1867 laten volgen direct door $h_1$, $h_2$ ..... $h_{25}$ van Januari 1868, enz. Vervolgens correleren wij deze met de één dag verschoven reeks $h_2$, $h_3$ ..... $h_{26}$, $h_2$, $h_3$ ..... $h_{26}$, $h_2$, $h_3$ ..... $h_{26}$ enz.

Ter verduidelijking:

| Jan. 1867 | 1868 | 1869 | | 1953 | |
|---|---|---|---|---|---|
| $h_1h_2$ ..... $h_{25}$ | $h_1h_2$ ..... $h_{25}$ | $h_1h_2$ ..... $h_{25}$ | ..... | $h_1h_2$ ..... $h_{26}$; | som S |
| $h_2h_3$ ..... $h_{26}$ | $h_2h_3$ ..... $h_{26}$ | $h_2h_3$ ..... $h_{26}$ | | $h_2h_3$ ..... $h_{26}$; | som $S^1$ |

Wij zouden nu de correlatiecoëfficiënt tussen de 84 x 25 boven elkaar staande paren kunnen berekenen, doch wij moeten goed beseffen, dat deze eigenlijk niet de autocorrelatiecoëfficiënt van de eerste orde in de bovenste rij is, omdat dan $h_1$ door $h_2$, $h_2$ door $h_3$ enz. ..... en tenslotte ook $h_{25}$ door $h_1$ gevolgd zou moeten worden, terwijl juist dit laatste niet het geval is. De tabelleermachine geeft ons de subsommen $S_1 = h_1 + h_2$ ..... $h_{25}$ (Jan. 1867); $S_2 = h_1 + h_2 +$ ..... $h_{25}$ (1868) ..... $S_{84}$; ook $S_1^{\ 1} = h_2 + h_3 +$ ..... $h_{26}$ (1867) ..... $S_{84}^{\ 1}$ ('53); ook $P_1 = h_1 h_2 + h_2 h_3 +$ ..... $h_{25} h_{26}$ (1867); $P_2 = h_1 h_2 +$ ..... $h_{25} h_{26}$ (1868) ..... $P_{84}$ ('53); ook $K_1 = h_1^{\ 2} + h_2^{\ 2} +$ ..... $h_{25}^{\ 2}$ (1867); $K_2 = h_1^{\ 2} +$ ..... $h_{25}^{\ 2}$ ('68) ..... $K_{84}$ ('53) en $K_1^{\ 1} = h_2^{\ 2} + h_3^{\ 2} +$ ..... $h_{26}^{\ 2}$ (1867); $K_2^{\ 1}$ (1868) ..... $K_{84}^{\ 1}$ ('53).

De correlatiecoëfficiënt $\mathring{c}$ tussen de 84 x 25 boven elkaar staande paren is dan

$$(34)\qquad \hat{c} = \frac{\sum_1^{84} P_i - \left(\sum_1^{84} S_i\right)\left(\sum_1^{84} S_i'\right)/25\times 84}{\sqrt{\left[\sum_1^{84} K_i - \left(\sum_1^{84} S_i\right)^2/25\times 84\right]\cdot\left[\sum_1^{84} K_i' - \left(\sum_1^{84} S_i'\right)^2/25\times 84\right]}}$$

Aangezien het verschil tussen $\sum_1^{84} S_i$ en $\sum_1^{84} S_i^{\ 1}$ mag worden verwaarloosd, (het is ongeveer 4 % ), evenals dat tussen $\sum_1^{84} K_i$ en $\sum_1^{84} K_i^{\ 1}$, wordt

$$(35)\qquad \mathring{c} = \frac{\sum_1^{84} P_i - \left(\sum_1^{84} S_i\right)^2/25\times 84}{\sum_1^{84} K_i - \left(\sum_1^{84} S_i\right)^2/25\times 84}$$

; men vergelijke $\mathring{c}$ met $\hat{c}$.

Wij kunnen nog trachten dit verschil tussen $\mathring{c}$ en $\hat{c}$ nader te bestuderen. Daarvoor gaat het om het verschil tussen

$$\frac{1}{84}(S_1 + S_2 + \ldots\ldots S_{84})^2 \text{ en } (S_1^{\,2} + S_2^{\,2} + \ldots\ldots S_{84}^{\,2}).$$

Wat doet dit verschil, gemiddeld beschouwd, als wij de vergelijking talloos vele keren maken, m.a.w. welke zijn de verwachtingswaarden van de twee uitdrukkingen?

Nu is $\mathcal{E}(S_1^{\,2} + S_2^{\,2} + \ldots\ldots S_{84}^{\,2}) = 84\,\mathcal{E}S^2$, de index weglatende, omdat de verschillende S-waarden ongecorreleerd zijn. Verder is

$$\frac{1}{84}\mathcal{E}(S_1 + S_2 + \ldots\ldots S_{84})^2 = \frac{1}{84}\left[84\,\mathcal{E}S^2 + 2\,\mathcal{E}(S_1S_2 + S_2S_3 + \ldots\ldots\right.$$
$$\ldots S_{83}S_{84}) + 2\,\mathcal{E}(S_1S_3 + S_2S_4 + \ldots\ldots S_{82}S_{84}) + 2\,\mathcal{E}(S_1S_4 + S_2S_5 + \ldots$$
$$\left.\ldots S_{81}S_{84}) + \ldots\ldots 2\,\mathcal{E}S_1S_{84}\right].$$

Echter geldt $\mathcal{E}S_pS_q = \mathcal{E}S_p\,\mathcal{E}S_q$ voor elke p en q, aangezien de S-waarden ongecorreleerd zijn. Zodat

$$\frac{1}{84}\mathcal{E}(S_1 + \ldots\ldots S_{84})^2 = \mathcal{E}S^2 + (2.83 + 2.82 + 2.81 + \ldots\ldots 2)\,(\mathcal{E}S)^2 =$$
$$\mathcal{E}S^2 + 83\,(\mathcal{E}S)^2 .$$

Dit moeten wij dus vergelijken met $84\,\mathcal{E}S^2$. Nu is zeker $(\mathcal{E}S)^2 \neq \mathcal{E}S^2$, immers $0 < \sigma_S^2 = \mathcal{E}S^2 - (\mathcal{E}S)^2$. Wanneer wij invoeren de variabiliteit $k = \sigma_s : \mathcal{E}s$, is $\mathcal{E}S^2 + 83\,(\mathcal{E}S)^2 = (84 + k^2)\,(\mathcal{E}S)^2$, zodat het verschil tussen $\hat{c}$ en $\mathring{c}$ ongeveer dat is tussen

$$\frac{\sum_1^{84} P_i - \frac{84}{25}\mathcal{E}S^2}{\sum_1^{84} K_i - \frac{84}{25}\mathcal{E}S^2} \quad\text{en}\quad \frac{\sum_1^{84} P_i - \frac{84+k^2}{25}\mathcal{E}S^2}{\sum_1^{84} K_i - \frac{84+k^2}{25}\mathcal{E}S^2} \tag{36}$$

Dit verschil is niet noemenswaardig als $k \ll 1$; hieraan is inderdaad voldaan 1° omdat de standaarddeviatie in de Januari-maandsommen (25 dagen vormen bijna een maand) klein is t.o.v. de gemiddelde J.-som (dit feit moge bekend ondersteld worden) en 2° omdat bovendien de standaarddeviatie in de sommen S nog weer ongeveer $\sqrt{84}$ keer zo klein is (S bestaat uit ongeveer 84 van elkaar onafhankelijke Januari's).

Vervolgens kunnen wij $\overset{o}{c}$ ook voor de tweede, derde, vierde en vijfde orde berekenen: $\overset{o}{c}_2$, $\overset{o}{c}_3$, $\overset{o}{c}_4$ en $\overset{o}{c}_5$. Wij vervangen daartoe slechts iedere $P = h_1h_2 + h_2h_3 + \ldots\ldots h_{25}h_{26}$ in de $\sum_1^{84} P_i$ in de teller van (34) door een nieuwe $P = h_1h_3 + h_2h_4 + \ldots\ldots h_{25}h_{27}$ en komen aldus tot $\overset{o}{c}_1$. Met $P = h_1h_4 + h_2h_5 + \ldots\ldots h_{25}h_{28}$ komen wij tot $\overset{o}{c}_2$ enz.

e) Wij lieten ook bepalen, geheel automatisch, door de machine, de som van alle $h_1$'s, van alle $h_2$'s, van alle $h_{15}$'s, van alle $h_{16}$'s en van alle $h_1.h_2$- en van alle $h_{15}.h_{16}$-waarden over alle Januari-maanden, alle Februari-maanden, etc. Langs deze weg konden wij zelf berekenen de variantie in de $h_1$-, de $h_2$-, de $h_{15}$- en de $h_{16}$-waarden en tevens de datumcorrelatiecoëfficiënt $r_{1.2}$ van $h_1$ op $h_2$ en $r_{15.16}$, die van $h_{15}$ op $h_{16}$, voor Januari, Februari, ..... December.

f) Ook berekenden wij tussen de 31 n-jarige dagsommen van Januari $H_1$, $H_2$, ..... $H_{31}$ de variantie V (zie blz. 7 ) en dan weer uit deze V via de onjuiste relatie $V = n\,\sigma^2$ de datumvariantie $\sigma^2$, n bekend zijnde (Januari-maanden, n = 84; Februari-maanden, n = 83; enz.) Zie voor de onjuistheid van deze relatie onder 1.2.

g) Tenslotte berekenden wij met behulp van de 31 48-jarige Januari-dagsommen $H_1$, $H_2$ ..... $H_{31}$ de autocorrelatiecoëfficiënt $C_1$ van de eerste orde tussen de 30 paren $H_1$, $H_2$; $H_2$, $H_3$ ..... $H_{30}$, $H_{31}$.

Ook $C_2$ tussen de 29 paren $H_1$, $H_3$; $H_2$, $H_4$ ..... $H_{29}$, $H_{31}$; idem $C_3$. Dit zijn de C's die wij ondermeer in (21) bedoelen. Zoals gezegd deden wij dit (om tijd te sparen) maar één keer (aan het materiaal der jaren 1901 - 1950; twee jaren ontbreken). Wij hadden het beter meermalen kunnen doen, bijv. aan "vele" groepen van 25 jaren. Wij krijgen dus een vrij onnauwkeurige persistentiecoëfficiënt. Idem voor de Febr.-, de Mrt. - ..... Dec.-maanden.

## 3 NUMERIEKE RESULTATEN

### 3.1 De persistentiecoëfficiënten c en r.

#### 3.1.1 Eerste oriëntatie

Ter oriëntatie hadden wij, lang voor de berekeningen met de tabelleermachines zouden plaats vinden, voor De Bilt voor een drietal droge en een drietal natte Jan.-maanden, eveneens April-maanden, Juli-en Octobermaanden de serie-interne correlatiecoëfficiënten van de orden 1, 2, 3 en 4 (dus $c_1$, $c_2$, $c_3$, $c_4$) berekend. Zie tabel 4

Tabel 4

De Bilt

| autocorrelatiecoëfficiënten c van orden 1, 2, 3 en 4. | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| droge maanden | | | | | natte maanden | | | | |
| | $c_1$ | $c_2$ | $c_3$ | $c_4$ | | $c_1$ | $c_2$ | $c_3$ | $c_4$ |
| Jan. 1929 18 mm | +0.333 | +0.088 | -0.052 | -0.200 | Jan. 1938 107 mm | +0.206 | -0.091 | +0.068 | +0.094 |
| Jan. 1933 25 mm | -0.070 | +0.218 | +0.153 | -0.174 | Jan. 1915 107 mm | +0.018 | +0.113 | -0.029 | -0.136 |
| Jan. 1901 22 mm | +0.555 | +0.190 | +0.011 | -0.153 | Jan. 1948 142 mm | +0.252 | +0.147 | +0.018 | +0.033 |
| Apr. 1933 19 mm | -0.021 | -0.051 | -0.169 | +0.529 | Apr. 1935 96 mm | -0.101 | +0.047 | -0.107 | -0.144 |
| Apr. 1913 20 mm | -0.147 | +0.153 | -0.100 | -0.159 | Apr. 1909 93 mm | +0.054 | +0.095 | +0.002 | +0.007 |
| Apr. 1946 16 mm | -0.086 | -0.121 | -0.134 | -0.129 | Apr. 1903 126 mm | +0.565 | +0.114 | -0.086 | -0.288 |
| Juli 1943 25 mm | +0.425 | +0.074 | +0.101 | +0.198 | Juli 1930 192 mm | +0.118 | -0.172 | -0.038 | -0.187 |
| Juli 1911 21 mm | +0.347 | +0.266 | -0.050 | -0.161 | Juli 1918 135 mm | +0.067 | 0.049 | -0.177 | +0.213 |
| Juli 1904 23 mm | +0.044 | -0.053 | -0.033 | -0.056 | Juli 1942 167 mm | +0.115 | -0.130 | -0.033 | -0.044 |
| Oct. 1910 20 mm | +0.385 | -0.199 | -0.184 | -0.158 | Oct. 1932 194 mm | -0.041 | -0.090 | -0.022 | -0.133 |
| Oct. 1940 27 mm | -0.003 | +0.246 | +0.066 | -0.147 | Oct. 1941 149 mm | +0.169 | -0.095 | -0.165 | -0.179 |
| Oct. 1948 26 mm | +0.192 | +0.213 | +0.157 | +0.084 | Oct. 1917 156 mm | +0.361 | +0.058 | -0.016 | +0.186 |
| 12 c's gemiddeld via z-transformatie | | | | | 12 c's gemiddeld via z-transformatie | | | | |
| | 0.173 | 0.087 | 0.020 | 0.039 | | 0.155 | 0.045 | 0.049 | 0.049 |
| Alle 24 c's gemiddeld via z-transformatie | | | | | | | | | |
| $c_1$ = 0.164 | | $c_2$ = 0.066 | | | $c_3$ = -0.034 | | $c_4$ = -0.044 | | |

De tabel biedt een onsystematisch beeld. De c's zijn blijkbaar maand voor maand zeer verschillend. Wij zien weinig of helemaal niets van een afneming van de c's met toenemende orde. Er zijn maanden met een positieve c, maar ook met een negatieve; of met een c dicht bij één, doch ook ongeveer nul. Louter statistisch gezien is het resultaat niet verwonderlijk, aangezien iedere c op slechts 26 à 30 paren gebaseerd is.

### 3.1.2 Exacte berekening van $c_1$; 7 Februari- en 8 Augustus-maanden

Wij kozen uit het ganse basismateriaal van Hoofddorp volkomen willekeurig 8 jaren met de bedoeling voor elk der maanden Februari en Augustus de autocorrelatiecoëfficiënt van de eerste orde $c_1$ te berekenen, op een wijze als beschreven werd onder a, b, c in 2. De volgende jaren kwamen uit de bus 1885 1941 1900 1923 1868 1896 1950 1892, met toevallig daarin twee schrikkeljaren. De maand Februari 1892 ontbrak. De resultaten der berekeningen werden opgenomen in de tabellen 5 en 6 en in fig. 1.

Tabel 5

Berekening van autocorrelatiecoëfficiënt $c_1$ in 7 Februari-maanden; Hoofddorp

| j = 23 of 24 | 1868 dagen 1 t/m 24 j = 24 | 1885 dagen 1 t/m 23 23 | 1896 dagen 1 t/m 24 24 | 1900 dagen 1 t/m 23 23 | 1923 dagen 1 t/m 23 23 | 1941 dagen 1 t/m 23 23 | 1950 dagen 1 t/m 23 23 | som |
|---|---|---|---|---|---|---|---|---|
| $h_j$ mm | 1.7 | 4.5 | 0.0 | 2.3 | 3.5 | 0.0 | 0.0 | 12.0 |
| $h_1$ mm | 9.5 | 9.1 | 0.0 | 0.0 | 3.0 | 0.0 | 3.9 | 25.5 |
| $\Delta = h_j - h_1$ | -7.1 | -4.6 | 0.0 | 2.3 | 0.5 | 0.0 | -3.9 | -13.5 |
| S mm | 43.3 | 65.6 | 4.6 | 55.4 | 26.6 | 36.4 | 65.8 | 297.5 |
| $q = h_j^2 - h_1^2$ | -87.36 | -62.56 | 0.00 | 5.29 | 3.25 | 0.00 | -15.21 | |
| K $mm^2$ | 259.63 | 652.70 | 7.96 | 457.88 | 92.40 | 165.92 | 351.82 | 1988.31 |
| $S^2$ $mm^2$ | 1874.89 | 4303.36 | 21.16 | 3069.16 | 707.56 | 1324.96 | 4329.64 | 15630.73 |
| $P_1$ $mm^2$ | 53.38 | 372.61 | 0.87 | 338.56 | 55.50 | 59.56 | 214.75 | 1095.23 |
| $c_1$ = T/n | -0.093 | 0.435 | -0.007 | 0.613 | 0.371 | -0.006 | 0.187 | 1.500 |
| $c_1^*$ | -0.158 | 0.387 | -0.007 | 0.625 | 0.387 | -0.006 | 0.116 | 1.344 |

Tabel 6

Berekening van autocorrelatiecoëfficiënt $c_1$ in 8 Augustus-maanden; Hoofddorp

| j = 26 | 1868 dagen 1 t/m 26 | 1885 dagen 1 t/m 26 | 1892 dagen 1 t/m 26 | 1896 dagen 1 t/m 26 | 1900 dagen 1 t/m 26 | 1923 dagen 1 t/m 26 | 1941 dagen 1 t/m 26 | 1950 dagen 1 t/m 26 | som |
|---|---|---|---|---|---|---|---|---|---|
| $h_{26}$ mm | 0.0 | 0.1 | 3.5 | 14.7 | 3.0 | 0.0 | 0.5 | 0.0 | 21.8 |
| $h_1$ mm | 0.0 | 0.0 | 6.3 | 0.9 | 0.0 | 14.7 | 1.6 | 4.2 | 27.7 |
| $\Delta = h_{26} - h_1$ | 0.0 | 0.1 | -2.8 | 13.8 | 3.0 | -14.7 | -0.9 | -4.2 | -5.9 |
| S mm | 7.89 | 5.80 | 4.64 | 9.10 | 7.42 | 6.56 | 11.19 | 7.57 | 59.17 |
| $q = h_{26}^2 - h_1^2$ | 0.0 | 0.01 | -27.44 | 215.28 | 9.0 | -216.09 | -2.31 | -17.64 | |
| K $mm^2$ | 882.33 | 581.98 | 307.40 | 825.92 | 592.58 | 1121.98 | 1122.97 | 718.73 | 6153.89 |
| $S^2$ $mm^2$ | 4747.21 | 3364.00 | 2152.96 | 8281.00 | 5505.64 | 4303.36 | 12521.61 | 5730.49 | 46606.27 |
| $P_1$ $mm^2$ | 524.01 | 262.31 | 128.64 | 448.43 | 320.21 | 27.14 | 559.94 | 397.33 | 2668.01 |
| $c_1 = T/n$ | 0.476 | 0.285 | 0.225 | 0.123 | 0.248 | -0.122 | 0.102 | 0.367 | 1.704 |
| $c_1^*$ | 0.476 | 0.286 | 0.192 | 0.237 | 0.269 | -0.153 | 0.095 | 0.343 | 1.745 |

Tabel 6a

| Samenvatting | | Februari | Augustus |
|---|---|---|---|
| Rekenkundig gemiddelde | $\bar{c}_1$ | 0.21 | 0.21 |
| Rekenkundig gemiddelde van benadering | $\overline{c_1^*}$ | 0.19 | 0.22 |
| Gemiddeld via z-transformatie 1) | $(\bar{c}_1)_z$ | 0.23 | 0.21 |
| Tweede benadering | $\hat{c}_1$ | 0.30 | 0.19 |

Toelichting $S^1 = S + \Delta$ ; $K^1 = K + q$; $T = P - \frac{S.S^1}{j-1}$ ; $N = K - \frac{S^2}{j-1}$ ;

$N^1 = K^1 - \frac{(S^1)^2}{j-1}$ ; $n = \sqrt{N.N^1}$; $\hat{c}_1 = \dfrac{\sum P_1 - \frac{1}{j-1}\sum S^2}{\sum K - \frac{1}{j-1}\sum S^2}$

1) Met deze Fisherse z-transformatie wordt bedoeld: wanneer de correlatiecoëfficiënt $r_i$ gebaseerd is op $p_i$ paren en aangetoond is, dat de n correlatiecoëfficiënten (i=1, 2,.....n) niet significant verschillen, vindt men het "beste gemiddelde" $r_o$ door oplossing uit $z_o = \frac{1}{2}\ln(1+r_o)/(1-r_o)$ , met $z_o = \sum(p_i-3)z_i / \sum(p_i-3)$ en $z_i = \frac{1}{2}\ln(1+r_i)/(1-r_i)$. Men kan bewijzen, dat het meer zin heeft deze $r_o$ te beschouwen dan het rekenkundig gemiddelde $\bar{z} = \frac{1}{n}\sum z_i$ . Over de moeilijkheden, die men ontmoet als men de onderlinge gelijkheid der correlatiecoëfficiënten toetsen wil, spreken wij later nog.

### 3.1.3 Berekening van $c_1$, $c_2$, ..... $c_5$ aan alle maanden uit alle jaren.

De tabellen 5 en 6 leren, dat - gelijk voor de hand ligt - het verschil tussen $c_1$ en $c_1^*$ relatief gering is als dat tussen $h_j$ en $h_1$ klein is ($h_j$ is $h_{23}$, $h_{24}$ of $h_{26}$). Natuurlijk is $c_1 = c_1^*$ als $h_j = h_1$. Hoe meer maanden wij in de berekening opnemen, hoe kleiner wordt het verschil tussen $\bar{h}_j$ en $\bar{h}_1$, want wij onderstellen de jaarlijkse gang binnen een maand verwaarloosbaar; juist hierom namen wij een maand; een halve maand zou misschien beter geweest zijn; echter komt dan niet alleen iedere $c_1$ op ten hoogste 14 paren te berusten, maar bovendien geraken wij in moeilijkheden met de autocorrelatiecoëfficiënt voor het geval alle 15 dagen droog zouden zijn en hierdoor de autoc.c. onbepaald is, terwijl volkomen droge perioden van 23 of meer dagen te Hoofddorp niet voor kwamen. De tweede benadering $\hat{c}_1$ is in tabel 6 weinig kleiner dan de exacte $\bar{c}_1$, maar in tabel 5 vrij veel groter.

Wij hebben erover gedacht op grond van dit resultaat voor alle maanden apart de $c_1^*$, $c_2^*$, $c_3^*$, $c_4^*$, $c_5^*$(volgens de benadering dus) te berekenen, maar achtten het doel niet zo veel werk (ongeveer 5000 correlatiecoëfficiënten) waard.

Wij geloofden te kunnen volstaan met de berekening van $\hat{c}_i$ ($i = 1, 2, 3, 4, 5$) voor alle maanden en die van $c_5$, op exacte wijze, alleen voor Januari, April, Juli en October (dit laatste in verband met een significantieprobleem, dat wij op deze plaats niet behandelen). De uitkomsten werden verenigd in de tabellen 7 en 8. De $c_5$'s en $c_5^*$'s werden ook nog in fig. 2 bijeengebracht.

Tabel 7

Autocorrelatiecoëfficiënten $c_i$, in procenten; Hoofddorp

| | | Jan. | Feb. | Mrt. | Apr. | Mei | Juni | Juli | Aug. | Sep. | Oct. | Nov. | Dec. | Jr. gem. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| orde i | aantal maanden | 84 | 83 | 80 | 83 | 80 | 84 | 81 | 84 | 83 | 81 | 79 | 81 | |
| 1 | exacte $\bar{c}$ | | 21 | | | | | | 21 | | | | | 21 |
| | benad. $\overline{c^*}$ | 16 | 19 | | | | | | 22 | | | | | 19 |
| | benad. $\hat{c}$ | 14 | 16 | 17 | 10 | 12 | 9 | 12 | 16 | 27 | 22 | 16 | 18 | 16 |
| | | | 30 | | | | | | 26 | | | | | 28 |
| | benad. $\mathring{c}$ | 22 | | | 16 | | | 20 | | | 30 | | | 22 |
| 2 | exacte $\bar{c}$ | | | | | | | | | | | | | |
| | benad. $\overline{c^*}$ | 0 | | | | | | | | | | | | |
| | benad. $\hat{c}$ | 0 | 1 | 5 | 3 | 2 | -2 | -2 | 2 | 8 | 8 | 7 | 11 | 4 |
| | benad. $\mathring{c}$ | 9 | | | 9 | | | 7 | | | 18 | | | 15 |
| 3 | exacte $\bar{c}$ | | | | | | | | | | | | | |
| | benad. $\overline{c^*}$ | -3 | | | | | | | | | | | | |
| | benad. $\hat{c}$ | -3 | -3 | 4 | 4 | -5 | -5 | -4 | 3 | 2 | 2 | -2 | 4 | -0 |
| | benad. $\mathring{c}$ | 6 | | | 10 | | | 6 | | | 12 | | | 11 |
| 4 | exacte $\bar{c}$ | | | | | | | | | | | | | |
| | benad. $\overline{c^*}$ | -5 | | | | | | | | | | | | |
| | benad. $\hat{c}$ | -5 | -5 | 4 | -2 | -5 | -1 | -5 | -4 | 1 | -5 | -1 | -4 | -3 |
| | benad. $\mathring{c}$ | 4 | | | 4 | | | 4 | | | 6 | | | 6 |
| 5 | exacte $\bar{c}$ | -0 | | | -5 | | | -4 | | | -7 | | | -4 |
| | benad. $\overline{c^*}$ | -6 | | | | | | | | | | | | |
| | benad. $\hat{c}$ | -2 | -7 | -3 | -4 | -4 | -8 | -6 | -5 | -1 | -6 | -6 | -2 | -5 |
| | benad. $\mathring{c}$ | 3 | | | 2 | | | 3 | | | 5 | | | 5 |

N.B. De $\bar{c}$, $\overline{c^*}$, $\hat{c}$ (tweede regel) berusten voor Feb. op 7, voor Aug. op 8 maanden.

Tabel 8

Autocorrelatiecoëfficiënten (in procenten); Hoofddorp

| | Januari | | | | | exacte $c_5$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $c_1^*$ | $c_2^*$ | $c_3^*$ | $c_4^*$ | $c_5^*$ | Jan. | Apr. | Jul. | Oct. |
| 1867 | - 6 | - 5 | - 5 | 21 | - 18 | - 16 | - 22 | - 9 | 10 |
| 68 | 12 | - 29 | 10 | 53 | - 1 | - 9 | 0 | - 12 | 31 |
| 69 | 16 | - 11 | - 16 | - 7 | 15 | 37 | 8 | 18 | - 10 |
| 70 | 22 | 18 | - 12 | - 14 | 2 | 15 | - 10 | - 19 | - 29 |
| 1871 | 3 | - 21 | - 4 | 17 | - 31 | - 33 | - 19 | - 23 | 26 |
| 72 | - 8 | - 10 | - 29 | 23 | - 01 | 8 | - 21 | - 3 | - 8 |
| 73 | 42 | 2 | 6 | - 6 | - 31 | - 20 | - 8 | - 6 | - 27 |
| 74 | 26 | - 0 | - 19 | - 29 | - 23 | 2 | 3 | 6 | - 22 |
| 75 | 34 | 1 | 33 | 20 | 1 | 4 | 15 | - 18 | - 4 |
| 76 | 16 | - 7 | - 4 | - 5 | - 30 | - 21 | - 23 | 50 | - 6 |
| 77 | 32 | - 0 | - 12 | - 16 | 2 | - 2 | - 21 | - 5 | 22 |
| 78 | 26 | - 4 | - 24 | - 22 | - 37 | - 38 | - 11 | - 9 | 9 |
| 79 | 1 | 42 | - 2 | - 15 | - 17 | - 8 | 22 | - 7 | - 9 |
| 80 | - 4 | - 10 | - 12 | - 13 | - 12 | - 11 | - 18 | 3 | - 12 |
| 1881 | 39 | - 4 | - 7 | - 10 | - 3 | - 3 | - 7 | - 29 | - 21 |
| 82 | - 2 | - 5 | 12 | - 8 | - 6 | - 3 | - 1 | - 6 | - 8 |
| 83 | - 7 | 1 | 5 | - 12 | - 12 | - 19 | - | 0 | - |
| 84 | 33 | 29 | 49 | 23 | - 4 | - 14 | 17 | - 20 | - 22 |
| 85 | 39 | 8 | - 8 | - 1 | - 1 | - 2 | - 14 | - 13 | - 5 |
| 86 | - 14 | 8 | - 15 | - 8 | 27 | 44 | - 15 | 4 | 16 |
| 87 | - 7 | - 17 | - 10 | - 10 | - 18 | - 18 | - 8 | - | - 25 |
| 88 | 0 | - 4 | 0 | 45 | 86 | 55 | - 15 | - | - 3 |
| 89 | - 5 | 8 | 49 | - 13 | 48 | 14 | 5 | - 26 | - 3 |
| 90 | 1 | 18 | 64 | 44 | 94 | 42 | - 24 | 1 | - |
| 1891 | 41 | - 12 | 12 | 15 | - 5 | - 6 | - | - | - |
| 92 | - | - | - | - | - | - | - 3 | - 14 | - 21 |
| 93 | 34 | 15 | - 7 | - 10 | - 17 | - 17 | - 1 | - 1 | 0 |
| 94 | - 6 | 31 | 92 | 12 | 56 | 27 | - 16 | - 16 | 6 |
| 95 | - 0 | 17 | - 1 | 11 | - 1 | 6 | 37 | - 6 | 10 |
| 96 | 1 | 32 | 41 | - 11 | - 22 | - 20 | - 7 | - 10 | - 4 |
| 97 | 16 | 2 | 28 | 54 | 62 | 33 | - 23 | - 3 | 4 |
| 98 | - 3 | 26 | - 9 | - 11 | - 11 | - 12 | - 11 | 12 | 23 |
| 99 | 21 | - 2 | 29 | 4 | - 20 | - 11 | - 2 | - 18 | - 24 |
| 00 | - 28 | 15 | 6 | 4 | 32 | 20 | - 17 | - 10 | - 18 |

Tabel 8 (vervolg)

Autocorrelatiecoëfficiënten (in procenten); Hoofddorp

| | Januari | | | | | exacte $c_5$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $c_1^*$ | $c_2^*$ | $c_3^*$ | $c_4^*$ | $c_5^*$ | Jan. | Apr. | Jul. | Oct. |
| 1901 | 14 | - 0 | - 1 | 5 | 16 | - 3 | - 9 | - 10 | - 22 |
| 02 | 14 | 10 | 17 | - 2 | - 26 | - 23 | - 9 | - 6 | 15 |
| 03 | - | - | - | - | - | - | - | - | - |
| 04 | - | - | - | - | - | - | - | - | - |
| 05 | 21 | 1 | 6 | 13 | 20 | 31 | 5 | 8 | - 9 |
| 06 | 18 | 12 | - 17 | - 23 | - 17 | - 10 | - 1 | - 9 | - 23 |
| 07 | - 2 | 30 | - 14 | 0 | - 7 | 7 | - 32 | 30 | 10 |
| 08 | 4 | - 9 | - 10 | - 13 | - 13 | - 18 | - 6 | - 13 | - 7 |
| 09 | 49 | 16 | - 11 | - 3 | 15 | 13 | 49 | - 18 | - 13 |
| 10 | 40 | 4 | - 6 | 21 | 34 | 22 | 52 | 1 | - 17 |
| 1911 | 17 | 12 | 19 | - 17 | - 20 | - 15 | 12 | - 7 | - 3 |
| 12 | 69 | 34 | 9 | - 12 | - 22 | - 16 | 34 | - 17 | 24 |
| 13 | 28 | - 13 | 17 | 48 | 16 | 12 | - 30 | - 9 | - 13 |
| 14 | 31 | 25 | 34 | - 1 | - 13 | - 15 | - 14 | 3 | - 13 |
| 15 | - 3 | - 5 | - 23 | - 34 | 11 | 23 | - 19 | - 11 | - 14 |
| 16 | 2 | 15 | - 39 | - 27 | - 9 | 24 | 32 | - 11 | - 33 |
| 17 | 29 | 3 | 22 | 5 | 1 | 16 | 40 | 37 | - 14 |
| 18 | 29 | - 3 | 29 | - 1 | - 18 | - 17 | - 15 | - 14 | 7 |
| 19 | 8 | 2 | - 6 | - 15 | - 19 | - 12 | - 20 | - 29 | - 14 |
| 20 | 48 | 21 | - 15 | - 41 | - 44 | - 52 | - 20 | - 1 | - 10 |
| 1921 | 22 | - 41 | - 41 | - 7 | 0 | 25 | - 2 | 10 | - 20 |
| 22 | - 3 | - 5 | - 13 | - 12 | - 1 | 25 | - 31 | 16 | - 1 |
| 23 | 9 | - 13 | 12 | - 27 | - 23 | - 24 | - 13 | - 1 | - 22 |
| 24 | 14 | - 14 | - 16 | - 25 | - 25 | - 20 | - 23 | 43 | - 6 |
| 25 | 52 | 17 | - 5 | - 25 | - 4 | 15 | 44 | - 8 | - 18 |
| 26 | 3 | 1 | 11 | - 19 | - 6 | 2 | - 15 | - 14 | 0 |
| 27 | - 13 | 31 | - 15 | 7 | - 6 | 0 | - 6 | 7 | - |
| 28 | - 17 | 9 | - 13 | - 28 | - 2 | 9 | - 31 | - 16 | - 9 |
| 29 | 34 | 22 | - 12 | - 7 | - 24 | - 27 | - 5 | 0 | - 24 |
| 30 | 32 | 7 | - 12 | - 16 | - 22 | - 22 | 22 | - | - 30 |
| 1931 | 8 | - 17 | - 11 | - 5 | 6 | 4 | - 19 | 6 | 5 |
| 32 | 30 | - 16 | - 8 | 25 | 34 | 42 | 16 | - 5 | 5 |
| 33 | 51 | 21 | 15 | 16 | 15 | 15 | - 29 | - 8 | - 15 |
| 34 | 37 | 48 | 32 | 3 | 3 | - 0 | - 15 | 69 | - 21 |
| 35 | 9 | - 21 | - 8 | 13 | 5 | 16 | - 22 | - 11 | - 42 |

Tabel 8 (vervolg)

Autocorrelatiecoëfficiënten (in procenten); Hoofddorp

| | Januari | | | | | exacte $c_5$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $c_1^*$ | $c_2^*$ | $c_3^*$ | $c_4^*$ | $c_5^*$ | Jan. | Apr. | Jul. | Oct. |
| 1936 | 13 | 3 | - 35 | - 16 | - 51 | - 44 | - 2 | - 21 | - 4 |
| 37 | - 10 | - 33 | - 26 | - 31 | 10 | 19 | 2 | - 11 | - 11 |
| 38 | 27 | - 9 | 14 | 29 | 28 | 10 | - 13 | 39 | 8 |
| 39 | - 16 | 0 | 10 | - 30 | - 52 | - 32 | - 24 | 5 | 8 |
| 40 | - 5 | 21 | 22 | 7 | 23 | - 18 | - 22 | - 14 | - 4 |
| 1941 | 19 | 30 | 25 | - 5 | 7 | 7 | - 6 | - 10 | 2 |
| 42 | 47 | - 4 | - 11 | 15 | 18 | 19 | - 6 | - 16 | - 28 |
| 43 | 33 | - 9 | - 19 | - 28 | - 22 | - 19 | - 18 | - 5 | - 13 |
| 44 | - 14 | - 14 | 0 | - 15 | - 19 | - 16 | - 3 | - 19 | - 16 |
| 45 | - 7 | - 8 | - 16 | 11 | - 15 | - 10 | - 6 | 9 | 23 |
| 46 | 27 | 21 | 2 | - 11 | - 2 | - 11 | 26 | - 16 | - 0 |
| 47 | 23 | 14 | 47 | - 5 | - 10 | - 11 | - 9 | 0 | - 3 |
| 48 | 33 | 12 | - 24 | - 1 | - 5 | 6 | 12 | 7 | - 15 |
| 49 | - 3 | - 24 | 4 | - 37 | - 19 | 2 | - 17 | - 15 | 40 |
| 50 | 8 | 41 | 9 | - 10 | - 8 | 23 | - 48 | - 12 | - 21 |
| 1951 | 6 | - 18 | - 4 | - 11 | - 11 | - 2 | 40 | - 11 | - 9 |
| 52 | 20 | - 4 | - 24 | - 35 | - 32 | - 24 | - 11 | - 11 | - 12 |
| 53 | - 12 | - 10 | - 12 | - 12 | - 12 | - 12 | - 17 | - 16 | - 13 |
| rek. gem. | 16 | 0 | - 3 | - 5 | - 6 | - 0 | - 5 | - 4 | - 7 |
| aantal keren > 0 | 60 | 41 | 36 | 29 | 29 | 38 | 22 | 25 | 22 |
| aantal keren < 0 | 24 | 43 | 48 | 55 | 55 | 46 | 61 | 56 | 57 |
| totale aantal | 84 | 84 | 84 | 84 | 84 | 84 | 83 | 81 | 79 |

### 3.1.4 Berekening van $C_1$, $C_2$, $C_3$ m.b.v. de veeljarige dagsommen

Thans volgt de berekening van de autocorrelatiecoëfficiënten $C_1$, $C_2$, $C_3$ met behulp van de veeljarige dagsommen, maand voor maand.

Ter verduidelijking: berekend werd bijv. voor Januari de autocorrelatiecoëfficiënt van de eerste orde ($C_1$) tussen de 30 paren $H_i$, $H_{i+1}$ (i = 1, 2 ..... 30), waarbij $H_i$ = 48 jarige som der hoeveelheden neerslag op de i-de Januari (om bepaalde redenen baseerden wij deze C op slechts 48 jaren (1901 - 1950; uitgezonderd 1903 en 1904), alhoewel wij over 80 à 85 jaren konden beschikken). Zo werd ook $C_2$, van de tweede orde, berekend, gebaseerd op 29 paren $H_i$, $H_{i+1}$; idem $C_3$. Zo ook voor Februari, Maart ..... December. De uitkomsten zijn bijeengenomen in tabel 9.

Tabel 9

Autocorrelatiecoëfficiënten c en C, in procenten; Hoofddorp

| orde | | Jan. | Feb. | Mrt. | Apr. | Mei | Juni | Juli | Aug. | Sep. | Oct. | Nov. | Dec. | Jaar |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | uit 48-j.som: C | 39 | 10 | 12 | 34 | 3 | 8 | 27 | 13 | 4 | 30 | 7 | 36 | 17 |
| | zie tabel 7; $c^*$ | 16 | 19 | | | | | | 22 | | | | | |
| | zie tabel 7; $\hat{c}$ | 14 | 16 | 17 | 10 | 12 | 9 | 12 | 16 | 27 | 22 | 16 | 18 | 16 |
| 2 | uit 48-j.som: C | 43 | 10 | 13 | 12 | -6 | -4 | 13 | -4 | 4 | 4 | -12 | 31 | 9 |
| | zie tabel 7: $c^*$ | 0 | | | | | | | | | | | | |
| | zie tabel 7; $\hat{c}$ | 0 | 1 | 5 | 3 | 2 | -2 | -2 | 2 | 8 | 8 | 7 | 11 | 4 |
| 3 | uit 48-j.som: C | 51 | -34 | 12 | -10 | -14 | -17 | -0 | 12 | 25 | -26 | 10 | 15 | 2 |
| | zie tabel 7: $c^*$ | | | | | | | | | | | | | |
| | zie tabel 7: $\hat{c}$ | -3 | -3 | 4 | 4 | -5 | -5 | -4 | 3 | 2 | 2 | -2 | 4 | -0 |

Het ligt voor de hand, dat wij de door Waldo Lewis en Mc. Intosh op blz. 10 genoemde stelling, volgens welke $C_j = \frac{1}{n} \sum_{i=1}^{n} c_{ij}$, verifieren. In ons geval in n = 48; echter hebben wij niet in iedere Januari gedurende 1901 - 1950 de exacte $c_1$ berekend, wel de benadering $c_1^*$. De 48 $c_1^*$-waarden hebben een gemiddelde $\overline{c_1^*}$ = 0.18, doch $C_1$ = 0.38; de 48 $c_2^*$'s leveren $\overline{c_2^*}$ = 0.06, echter $C_2$ = 0.43; voorts $\overline{c_3^*}$ = -0.02 maar $C_3$ = 0.51. De overeenkomst tussen $\overline{c_i^*}$ en $C_i$ is slecht (In de "stelling" dus niet waar?).

Trouwens wij zien ook al aan de onsystematische verandering van $C_1$ op $C_2$ op $C_3$, in elk der maanden, dat wij met deze C's niet kunnen werken. Toevallig (of niet toevallig?) is de overeenkomst tussen $C_1$, $C_2$, $C_3$, gemiddeld over het ganse jaar, t.w. 0.17 0.09 0.02 en $\hat{c}_1$, $\hat{c}_2$, $\hat{c}_3$ over het gehele jaar, t.w. 0.16 0.04 -0.00 (maar nu gemiddeld over 79 à 84 jaren en niet over 48 jaren) wel goed. Wij hadden deze C-waarden kunnen verzwijgen. Om zo volledig mogelijk te zijn en ook als waarschuwing om vooral niet met de op deze wijze berekende autocorrelatiecoëfficiënten te rekenen hebben wij er toch aandacht aan besteed.

### 3.1.5 Berekening van de datum-persistentiecoëfficiënten r

Men herleze noot 2) blz. 5 . Wij dachten te kunnen volstaan met een berekening van deze r voor twee plaatsen in iedere maand, n.l. van de 1ste op de 2de en van de 15de op de 16de. Tegelijk met de berekening van deze $r_{1.2}$ en $r_{15.16}$ konden de datumvarianties $s_1^2$ $s_2^2$ $s_{15}^2$ $s_{16}^2$ berekend worden. Zie tabel 10.

Tabel 10

Datum-standaarddeviaties s in mm; datum-persistentiecoëfficiënten r in %; Hoofddorp

| | Jan. | Feb. | Mrt. | Apr. | Mei | Jun. | Jul. | Aug. | Sep. | Oct. | Nov. | Dec. | Jaar |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| aantal maanden | 84 | 83 | 80 | 83 | 80 | 84 | 81 | 84 | 83 | 81 | 79 | 81 | |
| $s_1$ | 3.3 | 3.6 | 3.1 | 2.7 | 3.2 | 3.7 | 4.6 | 3.9 | 2.9 | 4.5 | 3.0 | 2.9 | 3.4 |
| $s_2$ | 4.9 | 3.3 | 2.3 | 2.0 | 4.1 | 4.4 | 3.5 | 5.1 | 5.7 | 4.9 | 4.2 | 3.5 | 4.0 |
| $s_{15}$ | 2.9 | 2.7 | 2.6 | 2.3 | 3.2 | 2.6 | 6.2 | 5.2 | 4.2 | 5.1 | 4.1 | 2.8 | 3.7 |
| $s_{16}$ | 3.8 | 3.4 | 2.0 | 1.9 | 3.6 | 3.2 | 5.8 | 5.5 | 4.5 | 7.4 | 5.0 | 4.1 | 4.2 |
| $\bar{s}_m$ | 3.7 | 3.1 | 2.5 | 2.8 | 3.8 | 3.8 | 4.9 | 4.8 | 4.5 | 5.0 | 3.9 | 3.7 | 4.0 |
| $r_{1.2}$ | 18 | 21 | 30 | 12 | 61 | 34 | 23 | 23 | 33 | 32 | 32 | 29 | 27 |
| $r_{15.16}$ | 18 | 39 | 29 | 16 | 17 | 10 | 9 | 34 | 17 | 16 | 16 | 14 | 23 |
| $\bar{r}_m$ | 19 | 30 | 24 | 30 | 38 | 22 | 18 | 30 | 26 | 26 | 26 | 20 | 27 |

Toelichting: $\bar{s}_m^2 = \frac{1}{6} \left[ (s_1^2 + s_2^2 + s_{15}^2 + s_{16}^2)_{Jan} + (s_1^2 + s_2^2)_{Feb} \right]$

voor Jan; idem Feb. enz.

$\bar{r}_m$ = gem. r per maand; voor Jan.: $\bar{r}_m = \left[ (r_{1.2} + r_{15.16})_{Jan} + (r_{1.2})_{Feb} \right] / 3$

idem Feb. enz.

De verschillende "soorten" persistentiecorrelatiecoëfficiënten, berekend in 3.1.2, 3.1.3, 3.1.4, 3.1.5, hebben wij verenigd in fig. 3.

De figuur leert, dat de 24 punten $r_{1.2}$ en $r_{15.16}$ rommelig liggen, hetgeen niet mag verwonderen, aangezien iedere r op maar ongeveer 80 paren berust. Wij weten bij ervaring dat al spoedig de $r_{1.2}$ "omhoog getrokken" wordt als het enige keren gebeurde, dat én op de 1ste én op de 2de veel regen viel. Wij konden controleren, dat zulks met $r_{1.2}$ in Mei inderdaad het geval was. Aangezien wij niet meer dan 84 jaren ter beschikking hebben, doen wij goed zoveel mogelijk $r_{i.i+1}$-waarden in iedere maand te berekenen. Eigenlijk voor i = 1, 2, 3 ..... 27 (29 of 30). Dit werk zou te veel tijd genomen hebben. Wij volstonden met i = 1 en i = 15 en berekenden per maand $\bar{r}_m = \frac{1}{3}(r_{1.2} + r_{15.16} + r_{1.2})$. Deze $r_m$'s liggen al minder grillig. Toch lijkt ons achteraf een tweetal r's per maand te weinig.

Verder leert de tabel, dat er tussen $c_5^*$, benaderd berekend, en $c_5$, exact berekend, een verschil is, dat sterk van de maand afhangt. Gemiddeld over alle Januari-maanden is $\bar{c}_5$ = -0% en $c_5^*$ = -6%, enz.

Aangezien de figuur niet op een jaarlijkse gang in de $r_{1.2}$- en $r_{15.16}$-waarden lijkt te wijzen (echter weten wij nog niet goed hoe wij de jaarlijkse gang statistisch zouden moeten toetsen) besloten wij de 24 r-waarden te middelen: $\bar{r}$ = 0.27 heet dan de beste schatting van de datumcorrelatiecoëfficiënt $\rho_I$ van de eerste orde.

In fig. 4 hebben wij alle persistentiecoëfficiënten in afhankelijkheid van hun orden uitgezet. De vraag is - want wij moeten aansturen op de datum-persistentiecoëfficiënten van de orden 1, 2, 3 ..... d.i. $\rho_I$, $\rho_{II}$, $\rho_{III}$ ....., waarbij wij een jaarlijkse gang verwaarloosbaar achten, - hoe wij op basis van de genoemde $\rho_I \cong 0.27$ en van wat verder berekend werd, tot een schatting van $\rho_{II}$, $\rho_{III}$ ..... zouden kunnen komen. Hier doet zich een nieuwe statistische moeilijkheid gelden, n.l. hoe toetsen wij of de universum-$\rho$ zeer waarschijnlijk van nul verschilt (of gelijk nul is) op basis van de steekproef-r, die natuurlijk meestal ongelijk nul is? De klassieke toets mag niet toegepast worden omdat het zeker is, dat elk der eventueel gecorreleerde universa niet een normale verdeling bezit. Het universum der dagelijkse hoeveelheden is immers uitermate scheef verdeeld (type "J-vorm"). Nu zijn er moderne, z.g. parametervrije, toetsen, die "geen" voorwaarden aan de verdelingen in de gecorreleerde universa stellen, doch dan mogen er niet te veel paren van gelijken voorkomen en juist in ons geval treden paren 0; 0 zeer vaak op. Wij zwijgen hier verder over deze kwestie, die wij nog in studie hebben (een criticus zou ook kunnen vragen waarom wij hier de gewone correlatiecoëfficiënt berekenen en niet de correlatie-ratio). Op grond van het verloop van de krommen in fig. 4 en onderstellende, dat $\rho_{IV} = \rho_{V} = \rho_{VI} = \rho_{VII} = \ldots$ $\ldots = 0$, lijkt ons tenslotte de kromme in fig. 4 een redelijke schatting. Er komt: $\rho_{I} = 0.25$ $\rho_{II} = 0.06$ $\rho_{III} = 0.02$ ;nogmaals: een jaarlijkse gang is verwaarloosbaar geacht. (de lezer zal bemerken, dat $0.06 = 0.25^2$; $0.02 = 0.25^3$; zie over dit verband in het addendum)

Hoewel wij zelf aangetoond hebben, dat het niet veel zin heeft de "relatie" (26) te toetsen, aangezien wij tot deze gekomen zijn via onderstellingen, die wellicht onhoudbaar zijn, gaan wij de bovengeschatte $\rho$'s substitueren in de uitdrukking voor $\mathcal{E}c_1$ en de aldus verkregen waarde vergelijken met de gemiddelde, exact berekende, $\bar{c}_1$'s van Februari en Augustus en de gemiddelde, benaderd berekende, $\overline{c_1^*}$ voor Januari. Met $\rho_{I} = 0.25$ $\rho_{II} = 0.06$ $\rho_{III} = 0.02$ komt er $\mathcal{E}c_1 \approxeq 0.20$.

De $\bar{c}_1$ (7 Febr.-maanden) = 0.21; de $\bar{c}_1$ (8 Aug.-mnd.) = 0.21 en de $\overline{c_1^*}$ (84 Jan.-mnd.) = 0.16. De overeenkomst is niet slecht.

## 3.2 De varianties en de standaarddeviaties

Wij hebben de varianties (standaarddeviatie = $\sqrt{\text{variantie}}$) op verschillende manieren berekend.

a) Per maand uit de eerste 25 dagsommen $h_1, h_2, \ldots\ldots h_{25}$; de maanden Januari (84 stuks), Juli (81 stuks), October (81 stuks); ook uit de eerste 24 dagsommen der maanden April (83 stuks). Verder nog voor de vroeger genoemde 7 Februari- en 8 Augustusmaanden. Zie tabel 11.

Tabel 11

De $\sqrt{v}$ (in mm) in de eerste 25 daghoeveelheden in Januari, April, Juli, October en in de eerste 24 daghoeveelheden in April, alsmede voor enige Februari- en Augustusmaanden. In mm.

| | Jan. | Feb. | Apr. | Juli | Aug. | Oct. | | Jan. | Feb. | Apr. | Juli | Aug. | Oct. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1867 | 5.3 | | 2.5 | 8.3 | | 5.0 | 1878 | 3.0 | | 4.6 | 2.3 | | 5.8 |
| 68 | 2.8 | 2.8 | 2.0 | 1.2 | 4.3 | 5.0 | 79 | 3.8 | | 4.4 | 6.4 | | 5.1 |
| 69 | 4.0 | | 3.3 | 2.9 | | 5.8 | 80 | 3.3 | | 1.8 | 5.1 | | 4.1 |
| 1870 | 2.5 | | 2.4 | 2.8 | | 7.1 | 81 | 2.5 | | 1.4 | 1.9 | | 3.5 |
| 71 | 1.4 | | 2.5 | 5.1 | | 10.1 | 82 | 4.4 | | 2.7 | 4.7 | | 4.0 |
| 72 | 3.9 | | 2.0 | 5.6 | | 9.9 | 83 | 1.9 | | | 4.4 | | |
| 73 | 3.0 | | 1.3 | 4.8 | | 6.6 | 84 | 3.8 | | 1.2 | 4.8 | | 4.6 |
| 74 | 3.1 | | 2.2 | 3.3 | | 3.8 | 85 | 4.5 | 4.7 | 0.5 | 0.8 | 4.3 | 8.6 |
| 75 | 2.7 | | 3.9 | 11.4 | | 3.2 | 86 | 4.3 | | 2.0 | 4.0 | | 5.2 |
| 76 | 1.3 | | 2.4 | 2.4 | | 1.9 | 87 | 2.4 | | 2.8 | | | 4.5 |
| 77 | 5.3 | | 2.0 | 2.2 | | 3.8 | 88 | 1.2 | | 2.3 | | | 3.4 |

Tabel 11 (vervolg)

De $\sqrt{v}$ (in mm) in de eerste 25 daghoeveelheden in Januari, April, Juli, October en in de eerste 24 daghoeveelheden in April, alsmede voor enige Februari- en Augustusmaanden. In mm.

| | Jan. | Feb. | Apr. | Juli | Aug. | Oct. | | Jan. | Feb. | Apr. | Juli | Aug. | Oct. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1889 | 0.5 | | 2.6 | 6.7 | | 3.0 | 1922 | 3.5 | | 2.6 | 5.7 | | 1.3 |
| 1890 | 2.7 | | 2.6 | 5.8 | | | 23 | 1.7 | 1.7 | 0.9 | 3.5 | 6.3 | 4.3 |
| 91 | 5.0 | | | | | | 24 | 2.9 | | 2.0 | 3.4 | | 7.1 |
| 92 | | | 1.4 | 3.3 | 3.0 | 6.1 | 25 | 2.2 | | 2.2 | 1.0 | | 4.4 |
| 93 | 3.8 | | 0.0 | 3.0 | | 5.6 | 26 | 3.2 | | 3.9 | 5.6 | | 5.9 |
| 94 | 2.2 | | 1.8 | 5.0 | | 2.0 | 27 | 1.7 | | 4.1 | 2.9 | | |
| 95 | 4.0 | | 1.9 | 4.6 | | 4.1 | 28 | 3.6 | | 2.5 | 1.8 | | 4.2 |
| 96 | 1.6 | 0.5 | 1.2 | 3.2 | 4.5 | 5.6 | 29 | 0.8 | | 1.4 | 2.4 | | 6.9 |
| 97 | 1.0 | | 4.0 | 3.3 | | 4.1 | 1930 | 1.9 | | 2.2 | | | 8.4 |
| 98 | 2.6 | | 1.9 | 3.6 | | 3.9 | 31 | 2.9 | | 2.4 | 3.4 | | 1.8 |
| 99 | 3.6 | | 3.1 | 5.6 | | 2.8 | 32 | 2.9 | | 4.8 | 3.0 | | 8.4 |
| 1900 | 3.1 | 3.9 | 1.9 | 2.1 | 3.9 | 4.3 | 33 | 1.2 | | 0.5 | 2.6 | | 3.4 |
| 01 | 1.1 | | 4.5 | 4.1 | | 6.8 | 34 | 3.6 | | 1.7 | 0.8 | | 4.6 |
| 02 | 1.6 | | 3.9 | 2.2 | | 2.6 | 35 | 2.6 | | 2.8 | 2.0 | | 2.8 |
| 03 | | | | | | | 36 | 2.4 | | 2.5 | 4.9 | | 6.2 |
| 04 | | | | | | | 37 | 3.5 | | 6.6 | 6.6 | | 1.2 |
| 05 | 2.7 | | 4.3 | 5.1 | | 7.8 | 38 | 4.2 | | 2.0 | 5.0 | | 7.8 |
| 06 | 5.3 | | 1.7 | 4.8 | | 6.0 | 39 | 4.4 | | 2.6 | 5.1 | | 4.7 |
| 07 | 2.0 | | 2.8 | 1.1 | | 4.1 | 1940 | 1.6 | | 3.1 | 5.2 | | 3.6 |
| 08 | 3.4 | | 1.4 | 4.7 | | 0.7 | 41 | 1.5 | 2.2 | 2.0 | 1.5 | 5.1 | 9.5 |
| 09 | 1.4 | | 1.7 | 4.4 | | 3.0 | 42 | 2.7 | | 2.3 | 10.3 | | 3.9 |
| 1910 | 2.2 | | 3.1 | 5.5 | | 1.9 | 43 | 6.2 | | 2.7 | 1.5 | | 1.8 |
| 11 | 2.2 | | 0.4 | 1.5 | | 6.2 | 44 | 5.3 | | 2.1 | 1.6 | | 4.6 |
| 12 | 2.8 | | 1.7 | 4.3 | | 4.1 | 45 | 4.1 | | 1.0 | 3.3 | | 0.8 |
| 13 | 2.1 | | 1.1 | 4.6 | | 5.0 | 46 | 1.5 | | 1.4 | 3.9 | | 2.5 |
| 14 | 3.6 | | 3.4 | 4.3 | | 2.1 | 47 | 1.1 | | 2.1 | 3.8 | | 1.2 |
| 15 | 4.3 | | 2.3 | 5.5 | | 2.6 | 48 | 4.4 | | 3.0 | 3.3 | | 4.0 |
| 16 | 3.3 | | 3.1 | 1.1 | | 3.1 | 49 | 3.0 | | 4.0 | 1.3 | | 6.2 |
| 17 | 1.4 | | 2.3 | 8.1 | | 7.5 | 50 | 2.1 | 2.7 | 1.7 | 5.3 | 4.5 | 2.9 |
| 18 | 4.8 | | 2.4 | 4.4 | | 8.0 | 51 | 3.4 | | 4.2 | 4.2 | | 0.5 |
| 19 | 3.1 | | 2.0 | 6.3 | | 5.6 | 52 | 2.7 | | 2.3 | 4.5 | | 4.5 |
| 1920 | 2.8 | | 4.0 | 4.3 | | 1.4 | 53 | 1.1 | | 3.8 | 2.2 | | 0.4 |
| 21 | 4.6 | | 2.0 | 0.3 | | 2.6 | $\sqrt{\overline{v}}$ | 3.2 | 2.9 | 2.8 | 4.4 | 4.7 | 5.0 |

b) Wij hebben ook de variantie V in de 31 48 jarige dagsommen $H_1, H_2, \ldots\ldots H_{31}$ van Januari berekend en met deze V weer de $\sigma^2$ geschat, via $V = 48\ \sigma^2$. Zie onder 1.2, formule (13). Men leest er, dat $\mathcal{E} V = 48\ \sigma^2 . f$, met $f$ een functie van $\rho_I, \rho_{II} \ldots\ldots, < 1$. Wij maken dus twee fouten door $\sigma^2$ te berekenen als V : 48, n.l. 1ste, wij moeten niet met één enkele V, maar met $\mathcal{E} V$ werken, welke benaderd kan worden als de gemiddelde waarde van V over vele groepen van 48 jaren; 2de, wij onderstellen $f = 1$. De uitkomsten hebben wij verenigd in tabel 12.

Tabel 12

Berekeningen van $\sigma$ (in mm) op verschillende wijzen; Hoofddorp

| | Jan. | Feb. | Mrt. | Apr. | Mei | Jun. | Jul. | Aug. | Sep. | Oct. | Nov. | Dec. | Jaar |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| V uit 48 j. | 802 | 413 | 159 | 320 | 622 | 489 | 1251 | 1219 | 1321 | 806 | 1113 | 955 | |
| V : 48 | 36.9 | 9.0 | 3.5 | 7.0 | 13.6 | 10.6 | 26.6 | 26.3 | 28.7 | 17.5 | 23.2 | 20.8 | |
| $\sigma = \sqrt{V : 48}$ | 6.1 | 3.0 | 1.9 | 2.6 | 3.7 | 3.3 | 5.2 | 5.1 | 5.4 | 4.2 | 4.8 | 4.6 | 4.0 |
| $\bar{s}_m$ tabel 10 | 3.7 | 3.1 | 2.5 | 2.8 | 3.8 | 3.8 | 4.9 | 4.8 | 4.5 | 5.0 | 3.9 | 3.7 | 4.0 |
| $\sqrt{\bar{v}}$ tabel 11 | 3.2 | 2.9 | | 2.8 | | | 4.4 | 4.7 | | 5.0 | | | 3.9 |

Al deze standaarddeviaties zijn in fig. 5 bijeengebracht

Tenslotte meenden wij nog, dat het mogelijk zou zijn om, bijv. voor de maand Januari, met behulp van relatie (7), de $\bar{v}$ (zie tabel 12), de $\bar{s}_m$ (als beste schatting van $\sigma$, zie tabel 10) de waarde van de term

$$\frac{24}{25} \rho_I + \frac{23}{25} \rho_{II} + \frac{22}{25} \rho_{III} + \ldots\ldots \text{ (als } \rho_{IV} = \rho_V = \rho_{VI} \ldots\ldots = 0,$$

ongeveer $\rho_I + \rho_{II} + \rho_{III}$) te kunnen berekenen. Deze berekeningen (ook voor Feb., Mrt. enz.) liepen op niets uit.

De waarden van deze term lopen uiteen tussen 0.01 (October) en 3.1 (Januari) met een jaargemiddelde 1.2. De 0.01 is uitermate klein, de 3.1 uitermate groot, doch zelfs 1.2 is te groot om de som van redelijke waarden van $\rho_I$, $\rho_{II}$ en $\rho_{IIII}$ te kunnen zijn (met $\rho_I = 0.25$, $\rho_{II} = 0.06$ en $\rho_{III} = 0.02$ is de som 0.33). Het grote steekproefeffect doet ons hier de das om. Een directe berekening van deze $\rho$'s, gevolgd door de sommering $\rho_I + \rho_{II} + \rho_{III}$, schijnt effectiever te zijn.

# 4. ENIGE VOORBEELDEN VAN HET GEBRUIK VAN DE PERSISTENTIECOËFFICIËNTEN

4.1.1 In de inleiding werd reeds verwezen naar ons rapport "Interne correlatiecoëfficiënten en voortschrijdende sommen" en ons artikel "The persistence of daily rainfall and the reliability of estimates of the probability distribution of rainfalls in periods of different lengths" in Geophysica 4:4; 1954. Zo kan men in een verzameling van 50 x 30 dagsommen in Juni (50 Juni-maanden) bijv. 10 etmaalhoeveelheden groter dan een bepaald bedrag tellen. De relatieve frequentie heet dan $\frac{10}{50 \times 30}$ 100 = $\frac{2}{3}$ %. Dit is weliswaar juist gezegd, maar als wij de betrouwbaarheidsmarge moeten berekenen (via de binomiale verdeling; zie de Stevens-tabel), behorende bij zulk een relatieve frequentie, moeten wij te doen hebben met onafhankelijke (dus zeker persistentieloze) elementen en dit is hier nu juist niet het geval.

Hoevele elementen telt zulk een steekproef van 30 stuks, welke een maand Juni biedt, effectief? In het rapport werd uiteengezet [1], dat

(37) $N_{eff} = N : w$, met $w_1 = 1 + 2 \left[\frac{N-1}{N}\rho_1 + \frac{N-2}{N}\rho_2 + \ldots\ldots \frac{1}{N}\rho_{N-1}\right]$. Hier is $N = 50 \times 30$; per maand is $n_{eff} = n : w_1$, met $n = 30$ en

(38) $w_1 = 1 + 2\left(\frac{29}{30}\rho_1 + \frac{28}{30}\rho_2 + \ldots\ldots \frac{1}{30}\rho_{29}\right)$.

Wanneer de $\rho_i$'s met toenemende i snel afnemen en reeds $\rho_1$ klein is,

(39) mogen we schrijven [2] $w_1 \approx 1 + 2(\rho_1 + \rho_2 + \rho_3 + \ldots\ldots)$

4.1.2 Wij kunnen ook de 29 met één dag voortschuivende tweedaagse sommen per maand van 30 dagen beschouwen (n.l. $x_1 + x_2$; $x_2 + x_3$; $x_3 + x_4$ .. ... $x_{29} + x_{30}$). Deze 29 stuks zijn equivalent met 29 : $w_2$ onafhankelijke elementen, waarin

(40) $w_2 = 1 + 2\left(\frac{28}{29}B_1 + \frac{27}{29}B_2 + \ldots\ldots \frac{1}{29}B_{28}\right)$, als $B_1$, $B_2$ enz.

de persistentiecoëfficiënten tussen de successieve tweedaagse sommen zijn. In bedoeld artikel lieten wij zien, dat deze B's functies van de $\rho$ 's zijn.

---

1) Zie voor deze formule o.a. bij J. Bartels: "Gesetz und Zufall in der Geophysik" in Naturwissenschaften 31 421 1943.

2) Wanneer de simpele relatie $\rho_i = \rho_1^{\,i}$ geldt (zie onder 6.3) en N zo groot is, dat $N(1-\rho) \gg \rho^{N+1}$, dan geldt in zeer goede benadering $w = \frac{1+\rho}{1-\rho}$ ; $\rho = 0.6$; $w = 3.0$ en $\rho = 0.3$; $w = 1.9$.

Wij rekenen $B_1$, $B_2$ ..... uit a.v.

$$(41)\quad B_1 = \frac{\sum(x_1 + x_2)(x_2 + x_3)}{\sqrt{\sum(x_1 + x_2)^2 \sum(x_2 + x_3)^2}} = \frac{\sum x_1x_2 + \sum x_1x_3 + \sum x_2^2 + \sum x_2x_3}{\sum(x_1^2) + 2\sum x_1x_2 + \sum x_2^2} = \frac{1+2\rho_1+\rho_2}{2(1+\rho_1)}$$

$$B_2 = \frac{\sum(x_1 + x_2)(x_3 + x_4)}{\sqrt{\sum(x_1 + x_2)^2 \cdot \sum(x_3 + x_4)^2}} = \frac{\sum x_1x_3 + \sum x_1x_4 + \sum x_2x_3 + \sum x_2x_4}{2\,(1 + \rho_1)} = \frac{\rho_1 + 2\rho_2 + \rho_3}{2\,(1+\rho_1)}$$

$$B_3 = \frac{\sum(x_1 + x_2)(x_4 + x_5)}{\sqrt{\sum(x_1 + x_2)^2 \cdot \sum(x_4 + x_5)^2}} = \frac{\rho_2 + 2\rho_3 + \rho_4}{2(1 + \rho_1)}$$

$$B_4 = \frac{\rho_3 + 2\rho_4 + \rho_5}{2\,(1 + \rho_1)}$$

$$B_5 = \frac{\rho_4 + 2\rho_5 + \rho_6}{2\,(1 + \rho_1)} \qquad \text{etc.}$$

Deze B's moeten worden berekend m.b.v. de $\rho$'s, waarna (40) kan worden berekend.

4.1.3 De met één dag voortschuivende 5-daagse sommen.
Er liggen er 27 stuks in een maand van 30 dagen.
Nu is het effectieve aantal 27 : $w_5$ met

$$(42)\quad w_5 = 1 + 2\left\{\frac{26}{27} D_1 + \frac{25}{27} D_2 + \ldots\ldots \frac{1}{27} D_{26}\right\}$$

De D's zijn thans ingewikkelder functies van de $\rho$'s. Welke? Wij berekenen:

$$(43)\quad D_1 = \frac{\sum(x_1 + x_2 + x_3 + x_4 + x_5)(x_2 + x_3 + x_4 + x_5 + x_6)}{\sqrt{\sum(x_1 + x_2 + x_3 + x_4 + x_5)^2 \cdot \sum(x_2 + \ldots\ldots x_6)^2}} = \frac{\text{teller}}{5+8\rho_1+6\rho_2+4\rho_3+2\rho_4}$$

$$\text{Teller: } \sum x_1x_2 + \sum x_1x_3 + \sum x_1x_4 + \ldots\ldots \sum x_5x_6 = 4 + 8\rho_1 + 6\rho_2 + 4\rho_3 + {}$$
$$+ 2\rho_4 + \rho_5\,;$$

$$D_2 = \frac{\mathcal{E}(x_1 + \ldots\ldots x_5)\ (x_3 + \ldots\ldots x_7)}{\mathcal{E}(x_1 + \ldots\ldots x_5)^2} = \frac{\text{teller}}{5 + 8\ {}_1 + 6\ {}_2 + 4\rho_3 + 2\rho_4}$$

Teller: $\mathcal{E}(x_1x_3) + \mathcal{E}(x_1x_4) + \mathcal{E}(x_1x_5) + \ldots\ldots \mathcal{E}x_5x_7 =$

$$3 + 6\rho_1 + 6\rho_2 + 4\rho_3 + 3\rho_4 + 2\rho_5 + \rho_6$$

$$D_3 = \frac{\mathcal{E}(x_1 + \ldots\ldots x_5)(x_4 + \ldots\ldots x_8)}{5 + 8\rho_1 + 6\rho_2 + 4\rho_3 + 2\rho_4} \ ;$$

teller $= 2 + 4\rho_1 + 4\rho_2 + 5\rho_3 + 4\rho_4 + 3\rho_5 + 2\rho_6 + \rho_7$

$$D_4 = \frac{\mathcal{E}(x_1 + \ldots\ldots x_5)\ (x_5 + \ldots\ldots x_9)}{5 + 8\rho_1 + 6\rho_2 + 4\rho_3 + 2\rho_4} \ ;$$

teller $= 1 + 2\rho_1 + 3\rho_2 + 4\rho_3 + 5\rho_4 + 4\rho_5 + 3\rho_6 + 2\rho_7 + \rho_8$

$$D_5 = \frac{\mathcal{E}(x_1 + \ldots\ldots x_5)\ (x_6 + \ldots\ldots x_{10})}{5 + 8\rho_1 + 6\rho_2 + 4\rho_3 + 2\rho_4} \ ;$$

teller $= \rho_1 + 2\rho_2 + 3\rho_3 + 4\rho_4 + 5\rho_5 + 4\rho_6 + 3\rho_7 + 2\rho_8 + \rho_9$

$$D_6 = \frac{\mathcal{E}(x_1 + \ldots\ldots x_5\ (x_7 + \ldots\ldots x_{11})}{5 + 8\rho_1 + 6\rho_2 + 4\rho_3 + 2\rho_4} =$$

$$\frac{\rho_2 + 2\rho_3 + 3\rho_4 + 4\rho_5 + 5\rho_6 + 4\rho_7 + 3\rho_8 + 2\rho_9 + \rho_{10}}{5 + 8\rho_1 + 6\rho_2 + 4\rho_3 + 2\rho_4}$$

$$D_7 = \frac{\rho_3 + 2\rho_4 + 3\rho_5 + 4\rho_6 + 5\rho_7 + 4\rho_8 + 3\rho_9 + 2\rho_{10} + \rho_{11}}{5 + 8\rho_1 + 6\rho_2 + 4\rho_3 + 2\rho_4} \ ;$$

$$D_8 = \frac{\rho_4 + 2\rho_5 + 3\rho_6 + 4\rho_7 + 5\rho_8 + 4\rho_9 + 3\rho_{10} + 2\rho_{11} + \rho_{12}}{5 + 8\rho_1 + 6\rho_2 + 4\rho_3 + 2\rho_4}$$

enz. enz. t/m D 26.

Wij moeten deze D's berekenen m.b.v. de $\rho$'s en daarna (42) berekenen.

4.2.1 Numerieke resultaten

Als $\rho_1 = 0.25$ $\rho_2 = 0.06$ ($\rho_3 = 0.02$) $\rho_4 = \rho_5 = \rho_6 = \ldots\ldots = 0$, dan is $w_1 = 1 + 2\,(\frac{29}{30}\,0.25 + \frac{28}{30}\,0.06 + \frac{27}{30}\,0.02) = 1 + 2\,(0.316) = 1.63.$ Dus een maand van 30 dagen is effectief equivalent met $\frac{30}{1.63} = 18$ dagen.

4.2.2 Voorts is $w_3 = 1 + 2\,(\frac{28}{29}\,B_1 + \frac{27}{29}\,B_2 + \frac{26}{29}\,B_3 + \frac{25}{29}\,B_4)$, waarin

$$B_1 = \frac{1 + 2\rho_1 + \rho_2}{2(1 + \rho_1)} = \frac{1.56}{2.50} = 0{,}624\ ;\ B_2 = \frac{\rho_1 + 2\rho_2 + \rho_2}{2.54} = 0.154;$$

$$B_3 = \frac{\rho_2 + 2\rho_3}{2.54} = 0{,}0394\ ;\ B_4 = \frac{\rho_3}{2.54} = 0{,}0079$$

zodat

$w_2 = 1 + 2\,(0.604 + 0.143 + 0.0353 + 0.0068) = 2.58$. Dus zijn 29 tweedaagse sommen per maand equivalent met slechts $\frac{29}{2.58} = 11$ à 12 stuks onafhankelijke.

4.2.3 Verder $w_3 = 1 + 2\,(\frac{26}{27}\,D_1 + \frac{25}{27}\,D_2 + \frac{24}{27}\,D_3 + \frac{23}{27}\,D_4 + \frac{22}{27}\,D_5 + \frac{21}{27}\,D_6 + \frac{20}{27}\,D_7 + \frac{19}{27}\,D_8)$

$$D_1 = \frac{4 + 8\rho_1 + 6\rho_2 + 4\rho_3}{5 + 8\rho_1 + 6\rho_2 + 4\rho_3} = \frac{6.44}{7.44} = 0.865\ ;$$

$$D_2 = \frac{3 + 6\rho_1 + 6\rho_2 + 4\rho_3}{7.44} = \frac{4.94}{7.44} = 0.664\ ;$$

$$D_3 = \frac{2 + 4\rho_1 + 4\rho_2 + 5\rho_3}{7.44} = \frac{3.34}{7.44} = 0.449\ ;$$

$$D_4 = \frac{1 + 2\rho_1 + 3\rho_2 + 4\rho_3}{7.44} = \frac{1.76}{7.44} = 0.237 \quad ;$$

$$D_5 = \frac{\rho_1 + 2\rho_2 + 3\rho_3}{7.44} = \frac{0.43}{7.44} = 0.0579 \quad ;$$

$$D_6 = \frac{\rho_2 + 2\rho_3}{7.44} = \frac{0.10}{7.44} = 0.01345 \quad ;$$

$$D_7 = \frac{\rho_3}{7.44} = \frac{0.02}{7.44} = 0.00269 \quad ;$$

zodat

$$w_5 = 1 + 2\,(0.834 + 0.615 + 0.399 + 0.202 + 0.0471 + 0.01046 + 0.00199) = 5.24$$

zodat 27 met 1 dag voortschuivende vijfdaagse sommen per maand equivalent zijn met slechts 27/5.24 = 5 à 6 stuks onafhankelijke.

5 BETROUWBAARHEIDSMARGES

Wij maken aan een voorbeeld duidelijk welke gevolgen dit feit heeft op de breedte van een marge $j_1$ à $j_2$ in een uitspraak, dat men met 95% zekerheid een vijfdaagse Juni-som verwachten mag in gemiddeld $j_1$ à $j_2$ jaren. De verzameling aller Juni- 5d.sommen (2184 stuks, n.l. 84 Juni's elk met 26, met één dag voortschuivende, vijfdaagse sommen) bevat, zoals wij voor Hoofddorp konden vaststellen, 10 stuks, gelijk aan of groter dan 51 mm. De relatieve frequentie van zulke vijfdaagse sommen is dus $\frac{10}{2184}$ . 100 = 0.352 %. De uitspraak, dat dús zulk een vijfdaagse hoeveelheid (op de een of andere wijze in Juni gelegen) verwacht mag worden 1 keer in gemiddeld (2184/30) : 10 = 7.3 jaren is, statistisch gesproken, echter onvolledig. Hoe (on)betrouwbaar is deze 7.3? Wij beantwoorden deze vraag m.b.v. de theorie van de binomiaalverdelingen en wel met behulp van Table VIII 1 "Binomial and Poisson Distributions Limits of the Expectertion (based on W.L. Stevens)", zie in de "Tables of Fisher and Yates". In de notatie bij deze tabel zal voor de gezochte relatieve frequentië p (waarvan 0.352 % een eerste schatting is) gelden, dat, met een zekerheid van 95 %, de waarde van p.2184 tussen 4.80 en 18.36 gelegen is; dit geldt in de onderstelling, dat de elementen onafhankelijk zijn. M.a.w. de uitspraak, dat een 5-d.som $\geq$ 51 mm voorkomt 1 keer in gemiddeld (2184 : 30) /18.36 = 5.1 à (2184 : 30)/4.80 = 19.3 jaren, heeft een betrouwbaarheid van 95 %. Met lette op deze brede marge (die natuurlijk de 7.3 bevat). Doch de elementen zijn niet onafhankelijk. Wij zeggen wel: 10 op 2184, maar effectief is het 10/5.24 = 1.9 op 2184/5.24 = 417. Weer met de Stevens-tabel werkende komt er nu een marge van (417 : 30)/722 = 1.9 à (417 : 30)/0.242 = 57.4 jaren. Nu een veel bredere marge. Eerst 5 à 19 jaren, nu 2 à 57 jaren.

Deze verbreding is een gevolg van de feiten:
1°. de vijfdaagse sommen overlappen elkaar en zelfs maximaal (voor 4 dagen)
2°. de successieve dagsommen zijn gepersisteerd en wel tot in de orde drie.
Ook zonder persistentie van dag op dag zouden de overlappende 5-daagse sommen zwaar gepersisteerd zijn. Zie de formules voor D; dan zou
$D_1 = \frac{4}{5}$ ; $D_2 = \frac{3}{5}$ ; $D_3 = \frac{2}{5}$ ; $D_4 = \frac{1}{5}$ en $w_5 = 1 + 2 \left(\frac{26}{27} \cdot \frac{4}{5} + \frac{25}{27} \cdot \frac{3}{5} + \frac{26}{27} \cdot \frac{2}{5} + \frac{25}{27} \cdot \frac{1}{5}\right) = 1 + 2\ (1.852) = 4.714$. Deze $w_5$ is wel is waar kleiner dan 5.24 (met $\rho_1 = 0.25$; $\rho_2 = 0.06$; $\rho_3 = 0.02$), maar toch nog groot.

Door de overlapping te verkleinen of zelfs de opvolgende vijfdaagse sommen in het geheel geen gemeenschappelijke dagen te geven, verkleinen wij de persistentie in deze vijfdaagse sommen (de D's worden kleiner), maar tegelijkertijd leveren ons de zelfde 84 Juni-maanden minder informatie. Men wint en verliest derhalve met zulk een handeling. De consequenties behandelden wij in ons artikel, zie in noot 1) blz. 2 .

# 6. ADDENDUM

## 6.1 De normale waarde van de dagsom per datum

De jaarlijkse gang van de gemiddelde dagsom (gemiddeld over 81 jaren) ziet men in fig. 6. Nog niet afgesloten berekeningen suggereren, dat de jaarlijkse gang binnen iedere maand verwaarloosbaar klein is (d.w.z. de grillige sprongen van dag op dag en de pieken en dalen kunnen louter toevallig zijn; zij zullen meer en meer verdwijnen naarmate wij over meer dan 81 jaren middelen).

## 6.2 Regressie

Het is wellicht gewenst op nog een andere betekenis van de persistentiecoëfficiënt te wijzen, een die meer tot de verbeelding spreekt, n.l. de betekenis i.v.m. de regressierechte.

Wanneer twee grootheden y en x gelieerd zijn en men zet in een Cartesisch coördinatenstelsel bij alle gemeten x-waarden de bijbehorende y-waarde af en men construeert de "beste" rechte door de punten, dan heeft de rechte "y t.o.v. x" een richtingscoëfficiënt (de hoek met de x-as heet $\alpha$) $b_{yx} = \mathrm{tg}\,\alpha = r\,\sigma_y/\sigma_x$, als $r$ = corr. coëff.; $\sigma$ = standaarddeviatie. De rechte "x t.o.v. y" heeft een $b_{xy} = \mathrm{tg}\,\beta = r\,\sigma_x/\sigma_y$ (de hoek, die deze rechte met de positieve richting van de y-as maakt heet $\beta$), zodat $r = \sqrt{\mathrm{tg}\,\alpha . \mathrm{tg}\,\beta}$

Met "y t.o.v. x" wordt bedoeld, dat voor de te zoeken rechte $y = a\,x + b$ zulke a en b-waarden gekozen worden, dat $\sum_1^N (y_g - y_b)^2$ = minimum, waarin $y_g$ = gemeten y (bij $x_g$) en $y_b$ is berekende y (bij zelfde $x_g$); N paren. In ons geval is de situatie zeer eenvoudig, omdat de x kan voorstellen de hoeveelheid neerslag op dag i en y die op dag i + 1 (als wij de persistentiecoëfficiënt van de orde één, $\rho_1$, beschouwen).

In dit geval is $\sigma_x = \sigma_y$ (de gemiddelde waarde van de standaarddeviatie $\sqrt{v}$ - zie tabel 11 - in een maand geeft er een schatting van) en is de regressierechte $y = \bar{y} + a\,(x - \bar{x})$, met $a = \rho_1$ en $\bar{x} = \bar{y}$ = gemiddelde hoeveelheid neerslag per dag. Dus $y = \rho_1\, x + \bar{x}\,(1 - \rho_1)$ voor elke maand. Deze $\bar{x}$ = gemiddelde dagsom binnen een maand is in tabel 13 opgenomen.

Tabel 13

| Jan. | Feb. | Mrt. | Apr. | Mei | Jun. | Jul. | Aug. | Sep. | Oct. | Nov. | Dec. | Jaar |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.7 | 1.5 | 1.5 | 1.4 | 1.5 | 1.8 | 2.2 | 3.0 | 2.8 | 3.0 | 2.5 | 2.2 | 2.1 mm |

De drie regressierechten, n.l. van dag i op dag i + 1, of op dag i + 2, of op dag i + 3 kunnen wij schrijven als $y = \rho_1\, x + A_1$; $y = \rho_2\, x + A_2$; $y = \rho_3\, x + A_3$, met $A_i = (1 - \rho_i)\,\bar{x}$ ; als wij $\rho_1 = 0.25$ $\rho_2 = 0.06$ en $\rho_3 = 0.02$ substitueren komen er, met tabel 13, A-waarden als in tabel 14.

Tabel 14

| | Jan. | Feb. | Mrt. | Apr. | Mei | Jun. | Jul. | Aug. | Sep. | Oct. | Nov. | Dec. | Jaar |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1$ | 1.27 | 1.12 | 1.12 | 1.05 | 1.12 | 1.35 | 1.65 | 2.25 | 2.10 | 2.25 | 1.87 | 1.65 | 1.58 mm |
| $A_2$ | 1.60 | 1.41 | 1.41 | 1.32 | 1.41 | 1.69 | 2.07 | 2.82 | 2.63 | 2.82 | 2.35 | 2.07 | 1.97 mm |
| $A_3$ | 1.66 | 1.47 | 1.47 | 1.37 | 1.47 | 1.76 | 2.16 | 2.94 | 2.74 | 2.94 | 2.45 | 2.16 | 2.06 mm |

Wij geven een toelichting bij Februari en Augustus, zie fig. 9. De hoeveelheden neerslag op de eerste, tweede, derde en vierde dag na een droge dag in Februari bedragen (gebruik de kromme x = 0), gemiddeld, resp. 1.12; 1.41; 1.47 en 1.50 mm. In Augustus: 2.25; 2.82; 2.94 en 3.00 mm. Gezien het feit, dat de persistentie doorwerkt tot en met de orde drie, ongeacht de maand, is de gemiddelde hoeveelheid neerslag op de $i^{de}$ dag na een dag met h mm in Februari 1.5 mm en in Augustus 3.0 mm, voor elke $i \geqslant 4$ en iedere h.

Men kan, zuiver statistisch gezien, met deze gemiddeld te verwachten dagelijkse hoeveelheid, die een functie is van x, de i en de maand, niet tevreden zijn. Wij zijn inderdaad meer gebaat bij een uitspraak van het volgende type: "de bewering, dat de dagsom op de i-de dag na een dag met $h_o$ mm (bijv. in Februari) gelegen is tussen $h_o^{1}$ en $h_o^{11}$ mm ($h_o^{1} < h_o^{*}$ en $h_o^{11} > h_o^{*}$) heeft een zekerheid van P = 95 %".

Hierin is $h_o^{*} = \rho_i \cdot h_o + A_i$. Deze grenzen $h_o^{1}$ en $h_o^{11}$ hangen samen o.a. met i en P. In vele problemen is de bij iedere $x = h_o$ behorende y normaal verdeeld rondom een gemiddelde, gegeven door de lineaire regressie $y = a\,x + b$. In zulke gevallen is $h_o^{11} - h_o^{*} = h_o^{*} - h_o^{1}$ en wel $2\,\sigma_y\sqrt{1-\rho_i^2}$ (door de $P \doteq 95$ % wordt de factor 2 bepaald). Men tekent dan twee rechten parallel aan de regressierechte op deze afstand. Binnen deze band liggen dan gemiddeld 95 % der punten x, y. In ons geval is van zulk een normale verdeling volstrekt geen sprake. De frequentieverdeling der hoeveelheden neerslag op de i-de dag na een dag met h mm (of beter: met een hoeveelheid tussen $h_1$ en $h_2$ mm) is zeer scheef, welke ook i is of welke ook de waarde(n) van h (of van $h_1$ en $h_2$) is (zijn). De vraag hoe wij dan wel de betrouwbaarheidsband moeten trekken hebben wij in studie. Bovendien is het vermoedelijk beter naar de beste kromlijnige regressie te zoeken.

## 6.3 De relatie $\rho_i = (\rho_1)^1$

Sommige auteurs gaan uit van $\rho_i = (\rho_1)^i$. Hoe komen zij daaraan en hoe komt het, dat er in de praktijk niet veel blijkt van deze simpele relatie? Zij de neerslag op dag 1: $h_1$, met $\mathcal{E} h_1 = n_1$ (normale waarde

(1) op bijv. 1 Jan.). Dan is hij op dag 2: $h_2 = n_2 + b_{2.1}(h_1 - n_1) + \varepsilon_2$;

$b_{2.1}$ is de regressiecoëfficiënt bij de rechtlijnige regressie tussen alle paren $h_1$, $h_2$ (dus de hoeveelheden neerslag op 1 en 2 Januari van vele jaren); $\varepsilon_2$ = toevallige variabele, onafhankelijk van $h_1$. Deze regressiecoëfficiënt is zo berekend, dat $\mathcal{E}\varepsilon_2^2$ = minimum (tegelijkertijd is dan $\mathcal{E}\varepsilon_2 = 0$), zodat $b_{2.1} = \rho_{1.2}\frac{s_2}{s_1}$, als de correlatiecoëfficiënt is:

(2) $$\rho_{1.2} = \left[\Sigma(h_1 - n_1)(h_2 - n_2)\right] \Big/ \sqrt{\Sigma(h_1 - n_1)^2\,\Sigma(h_2 - n_2)^2}$$

(gedacht is aan een sommatie over vele jaren) en $s_2^2$ = variantie tussen de hoeveelheden $h_2$ (in universum: $\sigma_2^2$); $s_1^2$ idem voor $h_1$.

Onderstel vervolgens gemakshalve de jaarlijkse gang zo gering, dat $s_2 \approx s_1$; dus $b_{2.1} = \rho_{1.2}$.

(3) Verder blijkt bij berekening $\mathcal{E}\varepsilon_2^2 = (1 - \rho_{1.2}^2)\,\sigma_2^2$.

Als dan $h_2$ door (1) gegeven wordt, is analoog $h_3 = n_3 + b_{3.2}(h_2 - n_2) + \varepsilon_3$, wanneer $b_{3.2}$ de regressiecoëfficiënt is in het universum der $h_2$; $h_3$-waarden, zodat ook $\mathcal{E}\varepsilon_3 = 0$ en $\mathcal{E}\varepsilon_3^2 = (1 - \rho_{2.3}^2)\,s_3^2$; $b_{3.2} = \rho_{2.3}\frac{s_3}{s_1} \cong \rho_{2.3}$; $\varepsilon_3$ is weer een toevallige variabele, onafhankelijk van $h_1$ en $h_2$.

(4) Uitgewerkt: $h_3 = n_3 + b\,(h_1 - n_1) + \varepsilon$, met $b = b_{3.2} = b_{2.1} = \rho_{1.2}$. $\rho_{2.3} = \rho_I^2$ en $\varepsilon = b_{3.2}\,\varepsilon_2 + \varepsilon_3$. Wij hebben ook de jaarlijkse gang in de correlatiecoëfficiënten $\rho_{1.2}$, $\rho_{2.3}$, $\rho_{3.4}$ ..... verwaarloosd (alle $= \rho_I$).

De correlatiecoëfficiënt $\rho_{II}$ tussen deze $h_3$ en $h_1$ is b als wij kunnen aantonen dat $\mathcal{E}\varepsilon^2 = (1 - b^2)\,s_3^2$. Dit gaat simpel a.v.:

$$b \cong \rho_I^2 ; \mathcal{E}\varepsilon^2 = \mathcal{E}(b_{3.2}\varepsilon_2 + \varepsilon_3)^2 = b_{3.2}^2 \mathcal{E}\varepsilon_2^2 + 2 b_{3.2} \mathcal{E}\varepsilon_2 \varepsilon_3 + \mathcal{E}\varepsilon_3^2 \cong$$
$$\rho_I^2 (1-\rho_I^2) s_2^2 + 2\rho_I \mathcal{E}\varepsilon_2\varepsilon_3 + (1-\rho_I^2) s_3^2 = (1-\rho_I^4) s_3^2 = (1-b^2) s_3^2$$

q.e.d.

Dus $\rho_{II} \approx \rho_I^2$ Idem bewijzen wij $\rho_{III} \approx \rho_I^3$ etc.

Onderstellingen zijn: 1) De toevallige "componenten" van $h_1$, $h_2$, $h_3$, ..... varieren onafhankelijk van elkaar en van $h_1$, $h_2$ .....

2) De datum-varianties $\sigma_1^2$, $\sigma_2^2$ ..... zijn gelijk.

3) De c.c. van dag i op i + 1 hangt niet met i samen.

Een en ander hierover kan men vinden o.a. bij Brooks: Handbook of Persistence in Meteorology, 1953, pg. 323. Brooks is niet duidelijk, wanneer hij "verklaart" hoe het komt, dat in de praktijk niet veel van $\rho_{II} \approx \rho_I^2$; $\rho_{III} \approx \rho_I^3$ etc. blijkt, want hij zegt: "In practice this very rarely happers; $\rho$ decreased rapidly at first, but owing to the varying incidence of the accidental element and generally also to the presence of short cyclic variations, it then fluctuates about zero". Hij noemt dus twee oorzaken, doch wat is de zin van de eerste? De zoeven genoemde voorwaarden 1, 2 en 3 beschouwende, lijkt het ons mogelijk, dat niet aan de eerste voorwaarde voldaan is. Wat betekent dit? Tweeërlei: de $\varepsilon_i$'s (vaste i) zijn wèl afhankelijk van $h_i$ (misschien zelfs van $h_{i-1}$, $h_{i-2}$ enz?) of (en) de $\varepsilon_i$ is wèl afhankelijk van de $\varepsilon_{i-1}$ (misschien zelfs van $\varepsilon_{i-2}$, $\varepsilon_{i-3}$, enz.?). Wat betekent dit meteorologisch? Wat <u>is</u> eigenlijk de $\varepsilon$ ? Wij weten daar geen antwoord op.

## 7 OPSOMMING DER PROBLEMEN, WAARTOE DIT ONDERZOEK LEIDDE.

1 Hoe bewijzen wij de stelling van R. Waldo Lewis en D.Mc. Intosh: "de autocorrelatiecoëfficiënt in de rij gemiddelden is gelijk aan het gemiddelde der autocorrelatiecoëfficiënten in alle rijen"? (zie pg. 10 )

2 Welke functie zal $\mathcal{E} c_1$ zijn van de datumcorrelatiecoëfficiënten $\rho_i$? Hierin is $c_1$ = autocorrelatiecoëfficiënt van de eerste orde bijv. tussen de 31 dagen in Januari. Is $\mathcal{E} c_1 = 0$ als alle $\rho_i = 0$? Is $\mathcal{E} c_1 = \rho_1$ als alle $\rho_i = 0$ voor i $\geqslant$ 2? En wat is $\mathcal{E} c_1$ als $\rho_1 \neq 0$, $\rho_2 \neq 0$ en $\rho_i = 0$ voor i $\geqslant$ 3? Enz.

3 Hoe toetsen wij of de 24 r-waarden ($r_{1.2}$ en $r_{15.16}$) al dan niet een significante jaarlijkse gang hebben? (zie pg. 29 ) Idem voor de waarden van $s_1$, $s_2$, $s_{15}$, $s_{16}$ (zie pg. 29 ).

4 Hoe toetsen wij de significantie der in dit onderzoek besproken typen correlatiecoëfficiënten, nu het zeker is, dat de klassieke toets van Fisher niet mag worden toegepast? (zie pg. 30 ).

5 Hoe vinden wij de beste vloeiende kromme tussen de grillig liggende punten van fig. 8? Welke vereffeningsformule verdient de voorkeur?. Deze vraag stellen wij ook in verband met het volgende. Wij hebben, zoals meermalen betoogd werd, de jaarlijkse gang binnen een maand verwaarloosbaar klein geacht. Natuurlijk is er een jaarlijkse gang (zie fig. 8). Hoe echter anders te handelen? Men zou alle dagsommen kunnen standaardiseren, bijv. door iedere hoeveelheid neerslag op de i-de van de maand m te verminderen met het veeljarig gemiddelde van die dag ($g_i$; af te lezen bijv. van bovenbedoelde nog te zoeken kromme) en daarna dit verschil te delen door de datumstandaarddeviatie in het universum der hoeveelheden op deze dagen i, d.i. $\sigma_i$; het quotiënt hete $q_i$; aldus levert ieder jaar 365 q-waarden. Het is zeker, dat, zo er in deze q's nog een jaarlijkse gang is, deze in elk geval kleiner is dan vóór deze bewerking. Vervolgens zou men misschien van één maand op drie successieve maanden mogen overgaan (seizoenen?) en in zulk een drietal de jaarlijkse gang mogen verwaarlozen. Echter zagen wij af van dit procédé 1° omdat de preciese vorm van de jaarlijkse gang zowel in $g_i$ als in $\sigma_i$ nog niet vast staat 2° omdat dit zeer veel meer rekenwerk gevergd zou hebben.

6 Is het wellicht beter helemaal niet met de "gewone" correlatiecoëfficiënt (eenleariteitscoëfficiënt) te werken, maar met bijv. de correlatieratio?

In verband hiermee, doen wij niet beter aan te sturen op de beste kromlijnige regressie?.

7 Hoe moeten wij de betrouwbaarheidsgordel rondom deze kromlijnige regressie leggen? (zie pg.41 )

8 Het is niet goed mogelijk te toetsen of aan de simpele relatie $\rho_i = (\rho_1)^i$ gehoorzaamd wordt (i.v.m. probleem 4), doch als wij zouden kunnen aantonen, dat deze betrekking niet geldt, waardoor dan niet? Aan welke voorwaarden wordt dan blijkbaar niet voldaan? Wat betekent dit? In het bijzonder physisch?

Aan verschillende van deze vragen (in het bijzonder 5, 6 en 7) wordt nog aandacht besteed als dit rapport afgesloten wordt.

## SUMMARY

The persistence of the daily amount of rainfall.

### 0 INTRODUCTION

In our paper "The persistence of daily rainfall and the reliability of estimates of the probability distribution of rainfalls in periods of different lengths" we dealt chiefly with the theoretical aspect of this rather complicated problem. The main purpose of the article was to derive a number of expressions, each a function of all the correlation coefficients in such a way that several practical questions could be answered systematically if one only substitutes the numerical values of the persistence coefficients.

In this report a survey is given of the way in which we computed these numerical values of the persistence coefficients by means of the Bull-tabulator (B.S. 120) connected with the Summary Punch. Since an exact computation appeared to be too elaborate we decided to make some approximations after having examined the influence of these approximations.

Moreover, it was necessary to make a clear distinction between two meanings of the persistence coefficient (alternative terms are coherence, conservation, auto regressive coefficients)

Suppose we have the disposal of n months of January. In each of these months we computed the auto correlation coefficient $c_{p.1}$ between the 30 pairs of elements $h_{p.i}$; $h_{p.i+1}$ ($i = 1, 2, \ldots\ldots 30$; $p = 1, 2, 3, \ldots\ldots n$). The symbol $h_{p.1}$ indicates the amount of rainfall on the i th day of the p th month of January. Than the average value is

$\bar{c}_1 = \frac{1}{n} \sum_{p=1}^{n} c_{p.1}$ . This $\bar{c}_1$ is called the average persistence (serial correlation) coefficient of the first order.

In the same way we computed the persistence coefficient of the second order $c_{p.2}$ between the 29 pairs of elements $h_{p.i}$; $h_{p.i+2}$ ($i = 1, 2, 3, \ldots\ldots 29$) and the average value

$$\bar{c}_2 = \frac{1}{n} \sum_{p=1}^{n} c_{p.2}, \text{ etc.}$$

But we can also calculate the correlation coefficient $r_{i.i+1}$ between the n pairs $h_{q.i}$; $h_{q.i+1}$ ($q = 1, 2, \ldots\ldots n$) and make this computation for $i = 1, 2, 3, \ldots\ldots 30$. This value $r_{i.i+1}$ may regarded as an estimation of the correlation coefficient between the rainfall on day i and that on the next day. Next we may average:

$$\bar{r}_1 = \frac{1}{29} \sum_{1}^{29} r_{i.i+1} \ .$$

The statistician is interested in several questions, for instance:

a) is $\bar{c}_1$ (nearly) equal to $\bar{r}_1$?

b) is $r_{1.2}$ (nearly) equal to $r_{2.3}$, (nearly) equal to $r_{3.4}$, $r_{4.5}$ etc.?

c) which coefficient shall we use, c or r and why?

## 1 THEORETICAL CONSIDERATIONS AND FORMULAS

1.1 The value of $c_1$ is computed per month of January on basis of the first 26 daily rainfalls. See (4). Here v = variance of the daily rainfalls $h_1$, $h_2$ ..... $h_{25}$ and $v^1$ the same for $h_2$, $h_3$ ..... $h_{26}$. We showed that $\mathcal{E}v \approx \mathcal{E}v'$ ($\mathcal{E}$ = expectation or "the average value of") In this way (we) arrived at (9). Here $\rho_I$ = the date correlation coefficient of day i to i+1, irrespective of i; $\rho_{II}$ = the date correlation coefficient of day i to i+2; etc.

The suppositions are: the annual course within the month may be neglected a) between the normal values of the daily rainfalls b) between the date variances (explanation: the n daily falls $h_{1.1}$, $h_{2.1}$ ..... $h_{n.1}$ have a variance $s_1^2$; the n daily falls $h_{1.2}$, $h_{2.2}$ ..... $h_{n.2}$ have $s_2^2$; we suppose $s_1^2 \approx s_2^2 \approx s_3^2$ ..... .) c) between the date correlation coefficients $\rho_{i.i+1}$, with i = 1, 2, 3, .....; and between the $\rho_{i.i+2}$'s; etc. See scheme I.

## 1.2 COMPUTATION OF THE PERSISTENCE COEFFICIENTS WITH THE AID OF THE DAILY RAINFALLS TOTALIZED OVER MANY YEARS

See scheme II.

We consider m groups of months of January, each group containing n months. The daily amounts gathered over n years for group k are $H_{k.1}$ $H_{k.2}$ $H_{k.3}$ ..... We compute the variance $V_k$ between the $H_{k.1}$, $H_{k.2}$ ...

... $H_{k.25}$ and $V_k^1$ between $H_{k.2}$, $H_{k.3}$ ..... $H_{k.26}$ and show that $\mathcal{E} V_k = \mathcal{E} V_k^1$; see (11). In this formula $\tau_1^2 = \lim\limits_{m\to\infty} t_1^2$, with $t_1^2$ = variance between $H_{1.1}$, $H_{2.1}$, $H_{3.1}$ ..... $H_{m.1}$; in the same way $t_2^2$ and $\tau_2^2$, etc.

Again we suppose $\tau_1^2 \approx \tau_2^2 \approx \tau_3^2$ ..... $= \tau^2$. Of course $\tau^2 = n\sigma^2$. Further $R_{1.2}$ = corr. coeff. between the m-pairs $H_{k.1}$, $H_{k.2}$ ($k = 1, 2, \ldots m$); $R_{2,3}$ between $H_{k.2}$, $H_{k.3}$ ($k = 1, 2, \ldots m$). The theoretical (universum) values are $P_{1.2}$, $P_{2.3}$ ..... We show that $P_I = \rho_I$; $P_{II} = \rho_{II}$ etc. In this way we arrive at (13).

Many authors overlook the fact that there is a difference between V and $\bar{V}$ ($= \mathcal{E}V$). They believe to be right in computing $\sigma$ by means of $\bar{V} = n\,\sigma^2$. We see that such a computation assumes that the factor between the brackets in (13) is nearly one and that the variability of the V's is rather small.

Next we draw the attention to the persistence coefficients C, analogous to the c's (see schemes I and II). R. Waldo Lewis and D. Mc Intosh [in their article "Some effects of the coherence of meteorological time series"; Met. Mag. 81 242 1952] say that $C_1 = \frac{1}{n} \sum c_1$ but they do not prove this statement. We have tried to show $\mathcal{E} c = \mathcal{E}\, C$ but did not succeed.

## 2 PRACTICAL COMPUTATION

By using the punched cards, containing the daily rainfalls of Hoofddorp (1867 - 1953), we made with the aid of the Bull-tabulator several data, which were printed together in a survey one per month; see for instance table 1. The products $h_i . h_{i+2}$, $h_i . h_{i+3}$, $h_i . h_{i+4}$ and $h_i . h_{i+5}$ were also computed mechanically and printed, see table 2.

Introducing $S = \sum\limits_1^{25} h_i$; $S^1 = \sum\limits_2^{26} h_i$; $S^{11} = \sum\limits_3^{27} h_i$ .....; $K = \sum\limits_1^{25} h_i^2$; $K^1 = \sum\limits_2^{26} h_i^2$ etc. we computed per month the expressions (29) (30) (31) and for all January months together: (32) and (35), giving $c_1$ ($c_2$, $c_3$ .....); $c_1^*$ ($c_2^*$, $c_3^*$ ..... ); $\hat{c}_2$, $\hat{c}_3$ .....) and $\mathring{c}_1$ ($\mathring{c}_2$, $\mathring{c}_3$ ... ...)

The c's are exact; the $c^*$, $\hat{c}$, $\mathring{c}$ 's are approximations. We also computed the values of $r_{1.2}$ and $r_{15.16}$ in each of the twelve months of the year.

## 3 NUMERICAL RESULTS

3.1 In table 4 a survey is given of the persistence coefficients $c_1$, $c_2$, $c_3$ and $c_4$ in some dry and some wet months (station De Bilt) chosen at random. The average values of these 24 c's are 0.164 0.066 -0.034 and -0.044.

3.1.2 In this section we compute in an exact way the $c_1$ for 7 months of February and 8 months of August. These months were chosen at random. See the tables 5 and 6; table 6 a presents a summary of the values of $\bar{c}_1$, $c_1^*$, $(\bar{c}_1)_z$ and $\hat{c}_1$. Here $(\bar{c}_1)_z$ indicates the value of $c_1$ averaged by means of the z-transformation of the correlation coefficient (according to Fisher).

### 3.1.3 COMPUTATION OF $c_1$, $c_2$, $c_3$, $c_4$ AND $c_5$ IN ALL MONTHS OF ALL YEARS.

As a result of the tables 5 and 6 we decided to compute only $\hat{c}_i$ (i = 1, 2, 3, 4, 5) for all months and $c_5$ only for all months of January, April, July and October. See the tables 7 and 8 and figure 2.

### 3.1.3 COMPUTATION OF $C_1$, $C_2$ AND $C_3$ WITH THE AID OF THE DAILY RAINFALLS TOTALIZED OVER MANY YEARS

$C_1$ indicates the autocorrelation coefficient of the first order between the 30 pairs $H_i$, $H_{i+1}$ (i = 1, 2, ..... 30), with $H_i$ = daily rainfall on the i-th January, gathered over 48 years etc. See table 9.

### 3.1.5 COMPUTATION OF THE DATE PERSISTENCE COEFFICIENT r

We computed only $r_{1.2}$ and $r_{15.16}$ per month. In the same time the variances $s_1$, $s_2$, $s_{15}$ and $s_{16}$ (of the daily rainfalls on 1, 2, 15 and 16 January, Febr. etc.) could be calculated. See table 10. All these persistence coefficients are drawn together in graph 3. Figure 4 gives the coefficients in dependence of the order. We decided to assume $\rho_{IV} = \rho_V = \rho_{VI} = \ldots\ldots = 0$ and to take the values 0.25 0.06 0.02 as the best approximations of $\rho_I$, $\rho_{II}$ and $\rho_{III}$ (note that $\rho_i = \rho_1^{\,i}$)

## 3.2 THE VARIANCES AND STANDARD DEVIATIONS

We have calculated the variance of the first 25 daily rainfalls in each month of January, April, July and October, see table 11. We also computed the standard deviation by means of the variance V between the 31 daily rainfalls in January, totalized over 48 years, using the approximation $V = 48\,\sigma^2$, see formula (13). The same for the other months. See table 12. All standard deviations have been drawn together in fig. 5.

## 4 SOME EXAMPLES THAT SHOW THE USE OF THE PERSISTENCE COEFFICIENTS

We refer to our article "The persistance of rainfall and the reliability of estimates of the probability distribution of rainfalls in periods of different lengths". Suppose we have the disposal of 50 x 30 daily rainfalls in June and suppose there are 10 daily falls larger than $h_0$ mm. The percentage frequency is $\frac{10}{50 \times 30}\,100 = \frac{2}{3}$. Next we want to know the reliability of this frequency which should be computed on the basis of the binomial distribution. The limits of the reliability interval pertaining to a result k "successes" in a sample of N trials may be read in the Stevens-table "Binomial and Poisson Distribution Limits of the Expectation" (see Statistical tables for biological, agriculture and medical research, by Fisher and Yates, 1949). This table is based on the assumption that all trials are independent, i.e. the elements are not persistent. Now it is a fact that a group of N elements taken from an universum, in which the successive persistence coefficients are $\rho_1, \rho_2, \rho_3 \ldots\ldots$, is equivalent with $N_{eff}$ independent elements, with $N_{eff} = N = w$ and

$$w = 1 + 2\left\{\frac{N-1}{N}\rho_1 + \frac{N-2}{N}\rho_2 + \cdots\cdots \frac{1}{N}\rho_{N-1}\right\}$$

In many cases the relation $\rho_i = \rho_1^{\,i}$ holds good. Then $w \approx \frac{1+\rho_1}{1-\rho_1}$, provided that $N(1-\rho_1) \gg \rho_1^{N+1}$ (large N and small $\rho_1$). Owing to the persistence the reliability interval becomes broader.

We also computed the effect of the persistence if we take overlapping k-sums, for instance k = 5. In such a case each month of 30 days contains 26 5-day sums. We counted in 84 x 26 = 2184 5-day sums (in all months of June out of a period of 84 years) 10 5-day sums equalling or exceeding 51 mm. So the relative frequency is 0.35 %; in other words: we may expect such an amount in a 5-day period in June once in 7,3 years.

Neglecting any persistence the reliability limits are 5 and 19 years. But there is persistence a) because the 5 day periods overlap each other for 4 days and b) because of the coherence between successive daily falls. The reliability limits become 2 and 57 years.

That means: There is a 95 % probability that the mean return period of a 5-day period of 51 or more mm in June is situated between 2 and 57 years.

## 6 ADDENDUM

6.1 In this section we treat the annual course of the normal value of the daily rainfall. In fig. 6 we have drawn the values of the daily rainfalls, averaged over 81 years. The irregular course of the 365 points is rather surprising but appeared not to be significant.

6.2 The persistence coefficient has also a significance which may be related to the effect of linear regression. In our case x denotes the daily rainfall on a day and y that on the next day. Then $y = \bar{y} + \rho_1 (x - \bar{x})$, with $\rho_1$ = persistence coefficient of the order 1; $\bar{x} \equiv \bar{y}$ = average daily rainfall. So $y = \rho_1 x + \bar{x}(1 - \rho_1)$ for each of the twelve months. The value of $\bar{x}$ within a month is shown in table 13. Of course we also consider the regression lines connecting the day i with i + 2 or with i + 3; so $y = \rho_1 x + A_1$; $y = \rho_2 x + A_2$ and $y = \rho_3 x + A_3$, with $A_i = (1 - \rho_i)\bar{x}$; we substitute $\rho_1 = 0.25$; $\rho_2 = 0.06$ and $\rho_3 = 0.02$ and come to table 14. See also fig. 9.

The two curves may be used in the following way: the daily rainfalls on the first, second and third day after a dry day in February is on an average 1.12; 1,41; 1.47 and 1.50 mm. In August: 2.25; 2.82; 2.94 and 3.00 mm. The statistician should not be satisfied with this statement, but should ask between which limits this rainfall is situated with a probability of say 95%. We are studying this question, which is a very difficult one, owing to the fact that the distribution of the daily rainfalls is not normal, but J-shaped. Moreover we had probably better search for the best curved regression (in stead of a linear regression).

### 6.3 DERIVATION OF THE SIMPLE RELATION $\rho_i = (\rho_1)^i$

The question arose why the numerical values of $\rho_1, \rho_2, \rho_3$ .. ... often do not satisfy this relation and what may be the meteorological cause.

7 In this chapter we summarize the unsolved problems, arising from the investigation which is treated in this report.

Tabel 1

Hoofddorp Januari 1877

| | i = 1 ..... 31 | | i = 1, 2, ..... 25 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| datum | $h_i = y_{1i}$ in tienden mm | $h_i^2$ | $h_i \cdot h_{i+1}$ | $y_{2i}$ | $y_{3i}$ | $y_{4i}$ | $y_{5i}$ | $y_{6i}$ | $y_{7i}$ |
| 31 | 26 | 676 | | | | | | | |
| 30 | 127 | 16129 | | | | | | | |
| 29 | 58 | 3364 | | | | | | | |
| 28 | 30 | 900 | | | | | | | |
| 27 | 2 | 4 | | | | | | | |
| 26 | 103 | 10609 | | | | | | | |
| 25 | 84 | 7056 | 8652 | 187 | 189 | 219 | 277 | 404 | 430 |
| 24 | 7 | 49 | 588 | 91 | 194 | 196 | 226 | 284 | 411 |
| 23 | 2 | 3 | 14 | 9 | 93 | 196 | 198 | 228 | 286 |
| 22 | | | | 2 | 9 | 93 | 196 | 198 | 227 |
| 21 | | | | | 2 | 9 | 93 | 196 | 198 |
| 20 | 35 | 1225 | | 35 | 35 | 37 | 44 | 128 | 231 |
| 19 | 28 | 784 | 980 | 63 | 63 | 63 | 65 | 72 | 156 |
| 18 | 1 | 1 | 28 | 29 | 64 | 64 | 64 | 66 | 73 |
| 17 | 8 | 64 | 8 | 9 | 37 | 72 | 72 | 72 | 74 |
| 16 | | | | 8 | 9 | 37 | 72 | 72 | 72 |
| 15 | 33 | 1089 | | 33 | 41 | 42 | 70 | 105 | 105 |
| 14 | 3 | 9 | 99 | 36 | 36 | 44 | 45 | 73 | 108 |
| 13 | | | | 3 | 36 | 36 | 44 | 45 | 73 |
| 12 | | | | | 3 | 36 | 36 | 44 | 45 |
| 11 | 53 | 2809 | | 53 | 53 | 56 | 89 | 89 | 97 |
| 10 | | | | 53 | 53 | 53 | 56 | 89 | 89 |
| 09 | 195 | 38025 | | 195 | 248 | 248 | 248 | 251 | 284 |
| 08 | 150 | 22500 | 29250 | 345 | 345 | 398 | 398 | 398 | 401 |
| 07 | 17 | 289 | 2550 | 167 | 362 | 362 | 415 | 415 | 415 |
| 06 | 34 | 1156 | 578 | 51 | 201 | 396 | 396 | 449 | 449 |
| 05 | 10 | 100 | 340 | 44 | 61 | 211 | 406 | 406 | 459 |
| 04 | 21 | 441 | 210 | 31 | 65 | 82 | 232 | 427 | 427 |
| 03 | 55 | 3025 | 1155 | 76 | 86 | 120 | 137 | 287 | 482 |
| 02 | 144 | 20736 | 7920 | 199 | 220 | 230 | 264 | 281 | 431 |
| 01 | 2 | 4 | 288 | 146 | 201 | 222 | 232 | 266 | 283 |
| som | 882<br>S | 99366<br>K | 52660<br>P | 1865 | 2706 | 3522 | 4375 | 5345 | 6307 |

Tabel 2

Hoofddorp Januari 1877

| | i=1 ..... 31 | i = 1 ..... 25 | | | |
|---|---|---|---|---|---|
| datum | $h_i$ tienden mm | $h_1h_{i+2}$ | $h_ih_{i+3}$ | $h_ih_{i+4}$ | $h_ih_{i+5}$ |
| 31 | 026 | | | | |
| 30 | 127 | | | | |
| 29 | 058 | | | | |
| 28 | 030 | | | | |
| 27 | 002 | | | | |
| 26 | 103 | | | | |
| 25 | 084 | 168 | 2520 | 4872 | 10668 |
| 24 | 007 | 721 | 14 | 210 | 406 |
| 23 | 002 | 168 | 206 | 4 | 60 |
| 22 | 0 | | | | |
| 21 | 0 | | | | |
| 20 | 035 | | 70 | 245 | 2940 |
| 19 | 028 | | | 56 | 196 |
| 18 | 001 | 35 | | | 2 |
| 17 | 008 | 224 | 280 | | |
| 16 | 0 | | | | |
| 15 | 033 | 264 | 33 | 924 | 1155 |
| 14 | 003 | | 24 | 3 | 84 |
| 13 | 0 | | | | |
| 12 | 0 | | | | |
| 11 | 053 | | 159 | 1749 | |
| 10 | 0 | | | | |
| 09 | 195 | 10335 | | | 585 |
| 08 | 150 | | 7950 | | |
| 07 | 017 | 3315 | | 901 | |
| 06 | 034 | 5100 | 6630 | | 1802 |
| 05 | 010 | 170 | 1500 | 1950 | |
| 04 | 021 | 714 | 357 | 3150 | 4095 |
| 03 | 055 | 550 | 1870 | 935 | 8250 |
| 02 | 144 | 3024 | 1440 | 4896 | 2448 |
| 01 | 002 | 110 | 42 | 20 | 68 |
| som | | 24898 $P_2$ | 23095 $P_3$ | 19915 $P_4$ | 32759 $P_5$ |

Schema I

Januari-maanden

| Jaar-nummer | dagelijkse hoeveelheden neerslag; dagen 1 t/m 26 | | | | | | variantie in de dagsommen der dagen: 1 t/m 25 | 2 t/m 26 | serie- of autocorrelatie-coëfficiënt: 1ste orde | 2de orde | 3de orde | enz. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $h_{1.1}$ | $h_{1.2}$ | $h_{1.3}$ | .......... | $h_{1.25}$ | $h_{1.26}$ | $v_1 \neq$ | $v_1^1$ | $c_{11} \gtrless$ | $c_{12} \gtrless$ | $c_{13}$ | enz. |
| 2 | $h_{2.1}$ | $h_{2.2}$ | $h_{2.3}$ | .......... | $h_{2.25}$ | $h_{2.26}$ | $v_2 \neq$ | $v_2^1$ | $c_{21} \gtrless$ | $c_{22} \gtrless$ | $c_{23}$ | enz. |
| 3 ..... | ..... | | | | | | ..... | | | | | |
| n | $h_{n.1}$ | $h_{n.2}$ | $h_{n.3}$ | .......... | $h_{n.25}$ | $h_{n.26}$ | $v_n \neq$ | $v_n^1$ | $c_{n1}$ | $c_{n2}$ | $c_{n3}$ | enz. |
| som | $H_1$ | $H_2$ | $H_3$ | .......... | $H_{25}$ | $H_{26}$ | gemiddeld | | gemiddeld | | | |
| variantie | $s_1^2 \neq$ | $s_2^2 \neq$ | $s_3^2 \neq$ | .......... $\neq$ | $s_{25}^2 \neq$ | $s_{26}^2$ | $\bar{v} \cong$ | $\overline{v^1}$ | $\bar{c}_1 >$ | $\bar{c}_2 >$ | $\bar{c}_3$ | enz. |
| corr.coëff. per datum: 1e orde | | $r_{1.2} \neq$ | $r_{2.3} \neq$ | $r_{3.4} \neq$ ..... $\neq$ | $r_{24.25} \neq$ | $r_{25.26}$ | $\overset{gem}{\bar{r}_1}$ | | | | | |
| 2e orde | | $r_{1.3} \neq$ | $r_{2.4} \neq$ | $r_{3.5}$ ...... | $r_{24.26}$ | $\bar{r}_2$ | | | | | | |
| enz. | | | | enz. | | enz. | | | | | | |
| Universum $n \rightarrow \infty$ | $\sigma_1^2$ | $= \sigma_2^2$ | $= \sigma_3^2$ | = .......... | $= \sigma_{25}^2$ | $= \sigma_{26}^2 = \sigma^2$ | $\mathcal{E}(v) = \mathcal{E}(v^1)$ | | $\mathcal{E}c_1 = \gamma_1$; | $\mathcal{E}c_2 = \gamma_2$; | $\mathcal{E}c_3 = \gamma_3$ | enz. |
| | $\rho_{1.2} =$ | $\rho_{2.3} =$ | $\rho_{3.4}$ ........ | | $= \rho_{25.26}$ | $= \rho_{I}$ | | | $\gamma_1$ | $> \gamma_2$ | $> \gamma_3$ | |
| | | $\rho_{1.3} =$ | $\rho_{2.4} =$ | $\rho_{3.5} =$ .... | $= \rho_{24.26}$ | $= \rho_{II}$ | | | | | | |
| | | | | enz. | | | | | | | | |

Vragen: $\bar{c}_p = \frac{1}{n} \sum_{i=1}^{n} c_{ip}$?; voor elke p ? $\gamma_1 = \gamma_1 (\rho_I; \rho_{II} \ldots)$ ?

Schema II

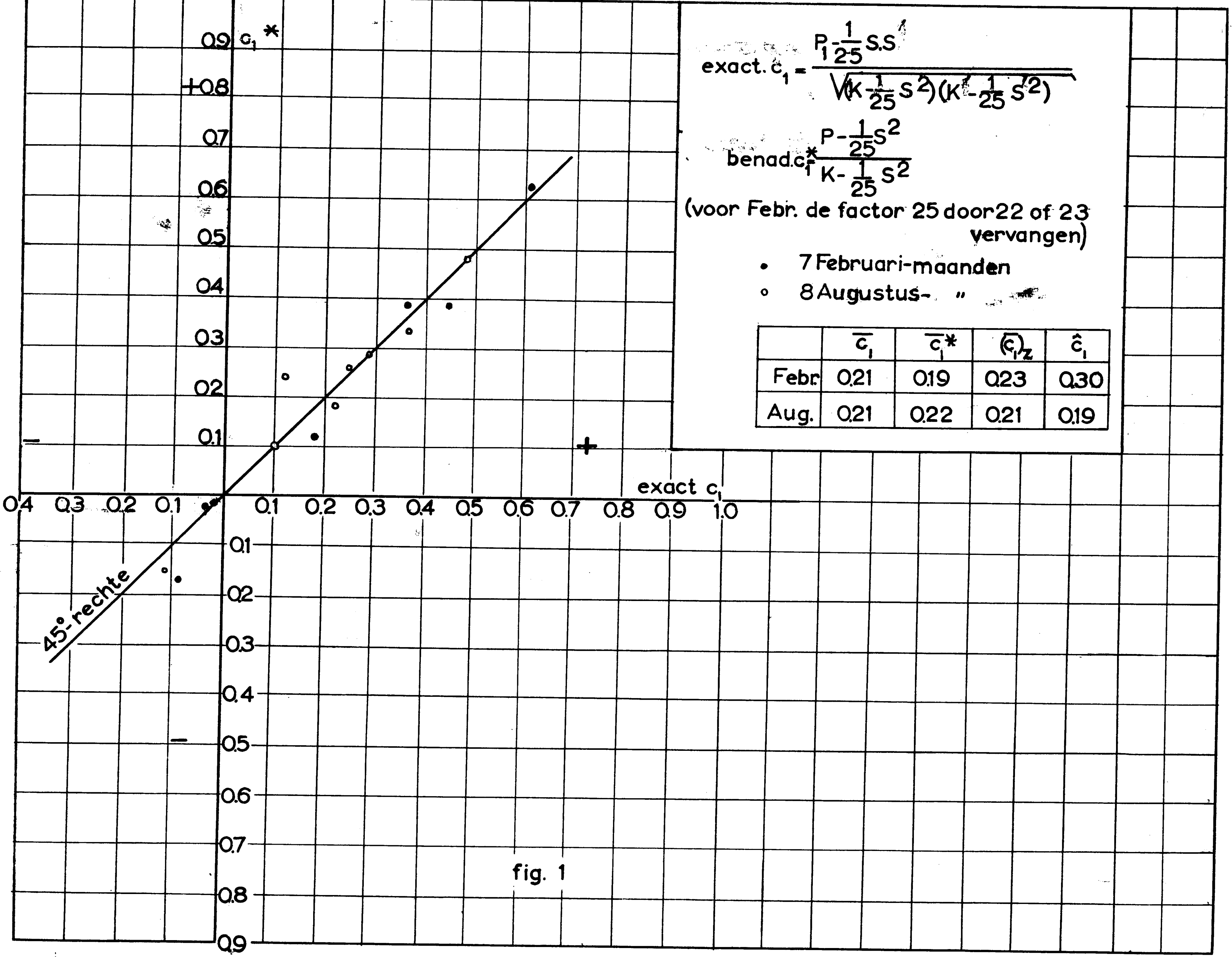
Januari - maanden

| nummer van groep van n jaren | n-jarige dagsommen; dagen 1 t/m 26 | | | | | | variantie in de n-jarige dagsommen der dagen | | serie- of autocorrelatie-coëfficiënt | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | 1 t/m 25 | 2 t/m 26 | 1ste orde | 2de orde | 3de orde | enz. |
| 1 | $H_{1.1}$ | $H_{1.2}$ | $H_{1.3}$ | .......... | $H_{1.25}$ | $H_{1.26}$ | $V_1 \neq$ | $V_1^1$ | $c_{1.1}$ | $\gtrless c_{1.2}$ | $\gtrless c_{1.3}$ | enz. |
| 2 | $H_{2.1}$ | $H_{2.2}$ | $H_{2.3}$ | .......... | $H_{2.25}$ | $H_{2.26}$ | $V_2 \neq$ | $V_2^1$ | | | | |
| 3..... | ..... | | | | | | ..... | | | | | |
| m | $H_{m.1}$ | $H_{m.2}$ | $H_{m.3}$ | .......... | $H_{m.25}$ | $H_{m.26}$ | $V_{25} \neq$ | $V_{25}^1$ | $c_{m.1}$ | $\gtrless c_{m.2}$ | $\gtrless c_{m.3}$ | enz. |
| variantie | $t_1^2 \neq$ | $t_2^2 \neq$ | $t_3^2$ | .......... $\neq$ | $t_{25}^2 \neq$ | $t_{26}^2$ | gemiddelde $\bar{V} \approx$ | $\overline{V^1}$ | gemiddelde $\bar{c}_1 >$ | $\bar{c}_2 >$ | $\bar{c}_3$ | enz. |
| datum - c.c. 1e orde | $R_{1.2} \neq$ | $R_{2.3} \neq$ | $R_{3.4} \neq$ | .......... | $R_{25.26}$ | | | | | | | |
| 2e orde | | $R_{1.3} \neq$ | $R_{2.4} \neq$ | .......... $\neq$ | $R_{24.26}$ | | | | | | | |
| enz. | | | | enz. | | | | | | | | |
| Universum $m \rightarrow \infty$ | $\tau_1^2 =$ | $\tau_2^2 =$ | $\tau_3^2 =$ | .................. | $= \tau^2$ | $= n\sigma^2$ | $\mathcal{E}_V =$ | $\mathcal{E}_{V^1}$ | $\mathcal{E}_{c_1} = \Gamma_1$ | $\mathcal{E}_{c_2} = \Gamma_2$ | $\mathcal{E}_{c_3} = \Gamma_3$ | |
| | $P_{1.2} =$ | $P_{2.3} =$ | $P_{3.4} =$ | | $= P_{24.26}$ | $= P_I$ | | | $\Gamma_1$ | $> \Gamma_2$ | $> \Gamma_3$ | |
| | | $P_{1.3} =$ | $P_{2.3} =$ | $P_{3.4} =$ | $P_{24.26}$ | $= P_{II}$ | | | | | | |
| | | | enz. | | | | | | | | | |

Vragen: $P_I \equiv \rho_I$; $P_{II} \equiv \rho_{II}$ ..... ?

benadering

Hoofddorp; autocorrelatiecoëff. 1ste orde




$$\text{exact. } c_1 = \frac{P_1 - \frac{1}{25} S.S'}{\sqrt{(K - \frac{1}{25} S^2)(K' - \frac{1}{25} S'^2)}}$$

$$\text{benad. } c_1^* = \frac{P - \frac{1}{25} S^2}{K - \frac{1}{25} S^2}$$

(voor Febr. de factor 25 door 22 of 23 vervangen)

• 7 Februari-maanden

∘ 8 Augustus- "

| | $\overline{c_1}$ | $\overline{c_1^*}$ | $(\overline{c_1})_z$ | $\hat{c}_1$ |
|---|---|---|---|---|
| Febr. | 0.21 | 0.19 | 0.23 | 0.30 |
| Aug. | 0.21 | 0.22 | 0.21 | 0.19 |

fig. 1

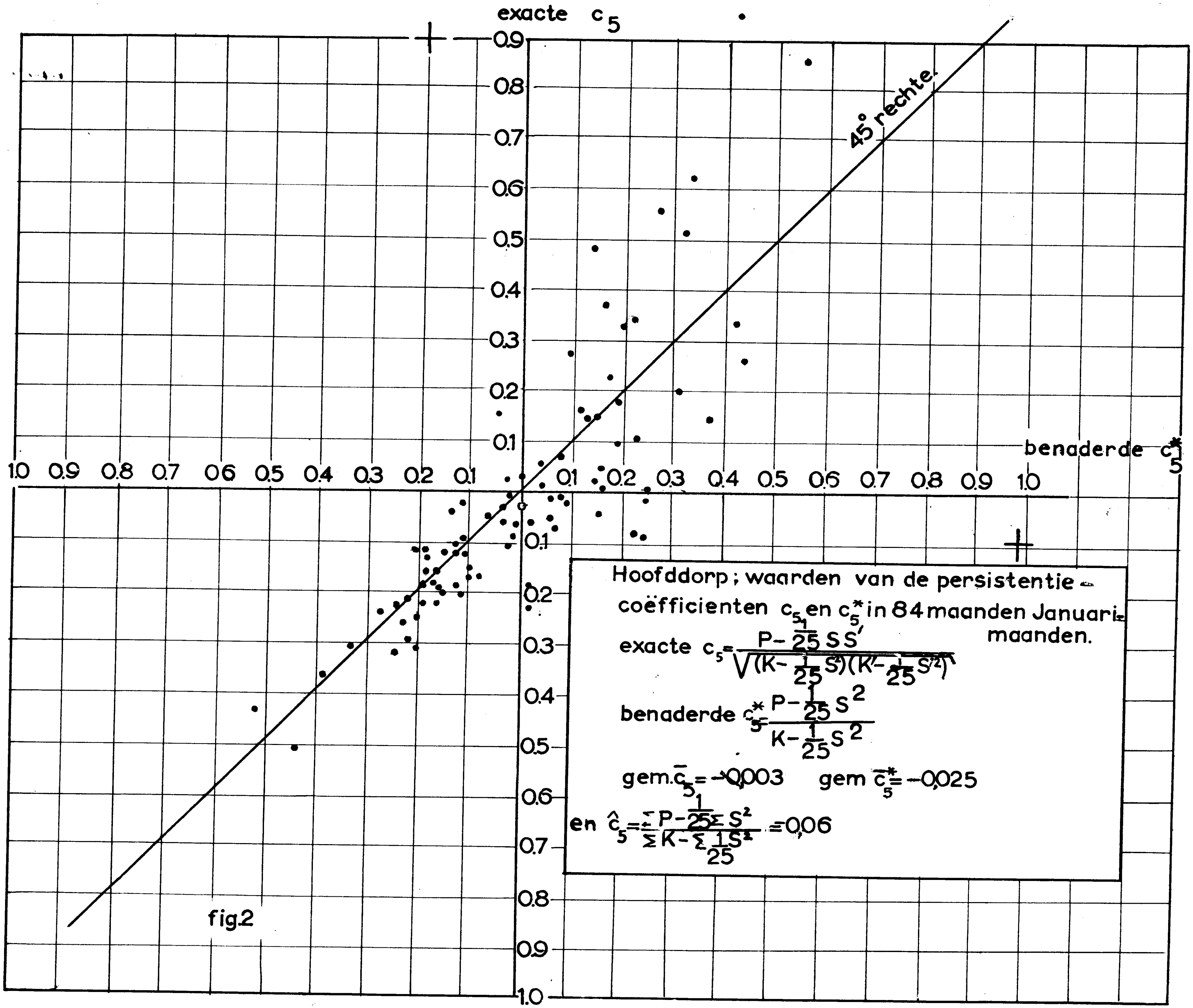

exacte $c_5$
benaderde $c_5^*$
45° rechte.
Hoofddorp; waarden van de persistentie-coëfficienten $c_5$ en $c_5^*$ in 84 maanden Januari-maanden.
exacte $c_5 = \dfrac{P - \frac{1}{25} S S'}{\sqrt{(K - \frac{1}{25} S^2)(K' - \frac{1}{25} S'^2)}}$
benaderde $c_5^* = \dfrac{P - \frac{1}{25} S^2}{K - \frac{1}{25} S^2}$
gem. $\bar{c}_5 = -0{,}003$ gem $\bar{c}_5^* = -0{,}025$
en $\hat{c}_5 = \dfrac{\Sigma P - \frac{1}{25}\Sigma S^2}{\Sigma K - \Sigma \frac{1}{25} S^2} = 0{,}06$

fig.2

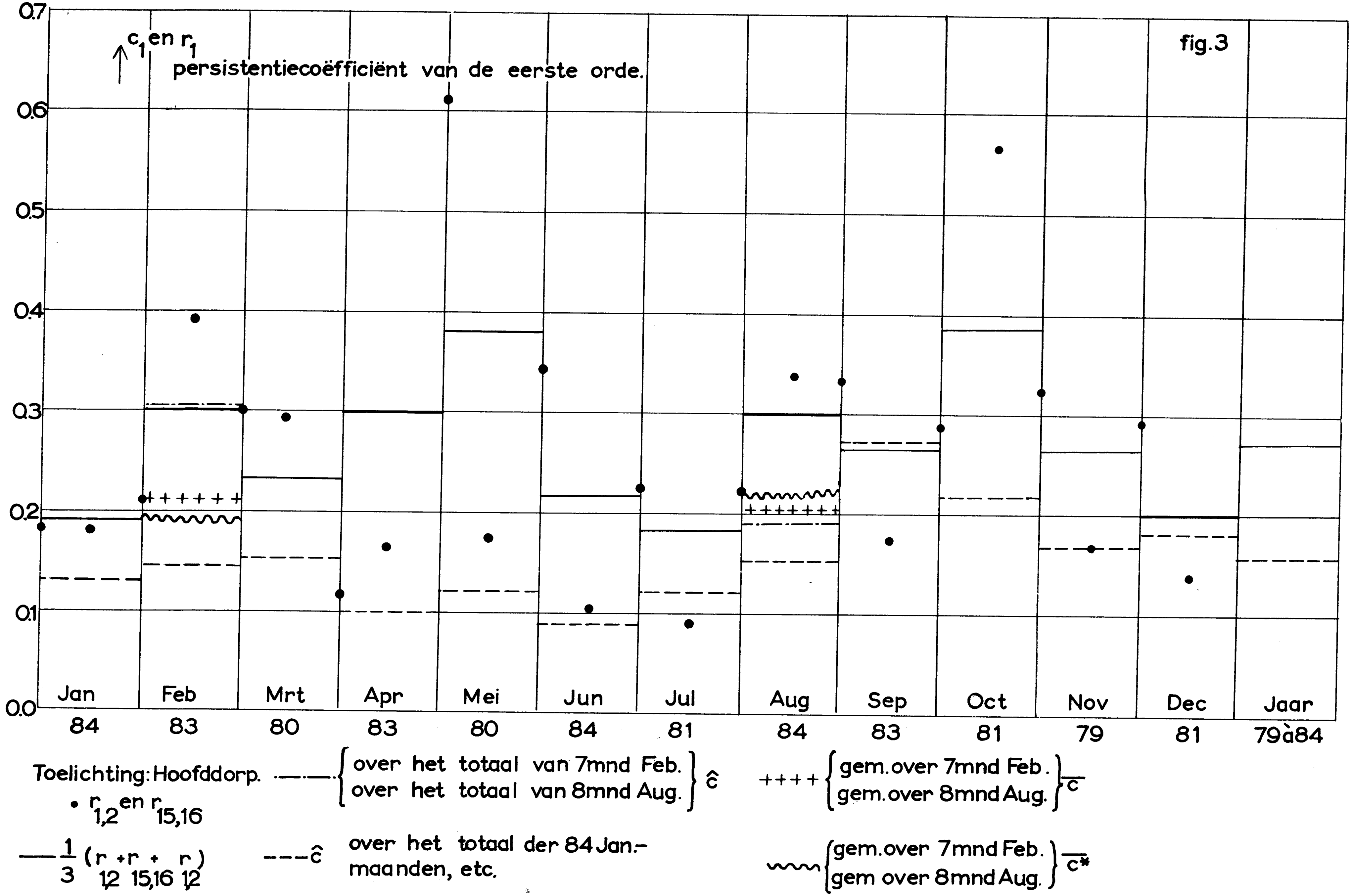
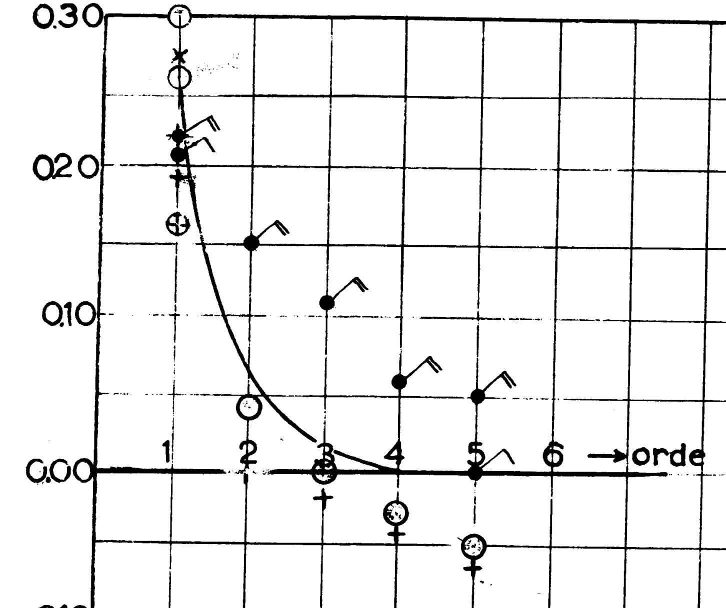


fig.3
$\uparrow c_1$ en $r_1$
persistentiecoëfficiënt van de eerste orde.
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Jan 84
Feb 83
Mrt 80
Apr 83
Mei 80
Jun 84
Jul 81
Aug 84
Sep 83
Oct 81
Nov 79
Dec 81
Jaar 79 à 84
Toelichting: Hoofddorp.
• $r_{1,2}$ en $r_{15,16}$
—— $\frac{1}{3}(r_{1,2} + r_{15,16} + r_{1,2})$
–·–·– {over het totaal van 7mnd Feb. / over het totaal van 8mnd Aug.} $\hat{c}$
- - - $\hat{c}$ over het totaal der 84 Jan.-maanden, etc.
++++ {gem. over 7mnd Feb. / gem. over 8mnd Aug.} $\overline{c}$
{gem. over 7mnd Feb. / gem over 8mnd Aug.} $\overline{c^*}$




**VERKLARING.**

| | ORDE → | 1 | | 2 | | 3 | | 4 | | 5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| • | $\bar{C}$ | 0.21 | 7 Feb | | | | | | | -0.00 | 84 Jan |
| • | | 0.21 | 8 Aug | | | | | | | | |
| + | $\bar{C}^*$ | 0.16 | 84 Jan. | -0.00 | 84 Jan | -0.03 | 84 Jan | -0.04 | 84 Jan. | -0.06 | 84 Jan |
| + | | 0.19 | 7 Febr. | | | | | | | | |
| + | | 0.22 | 8 Aug. | | | | | | | | |
| ○ | $\hat{C}$ | 0.16 | 82 jaren. | 0.04 | 82 j. | -0.00 | 82 j. | -0.03 | 82 j. | -0.05 | 82 j. |
| ⊙ | | 0.36 | 7 Febr. | | | | | | | | |
| ⊙ | | 0.26 | 8 Aug. | | | | | | | | |
| • | $\mathring{C}$ | 0.22 | 82 jaren | 0.15 | 82 j. | 0.11 | 82 j. | 0.06 | 82 j. | 0.05 | 82 j. |
| × | $\bar{r} = \bar{r}_{1.2 \; 15.16.}$ | 0.27 | 82 " | | | | | | | | |

fig. 4

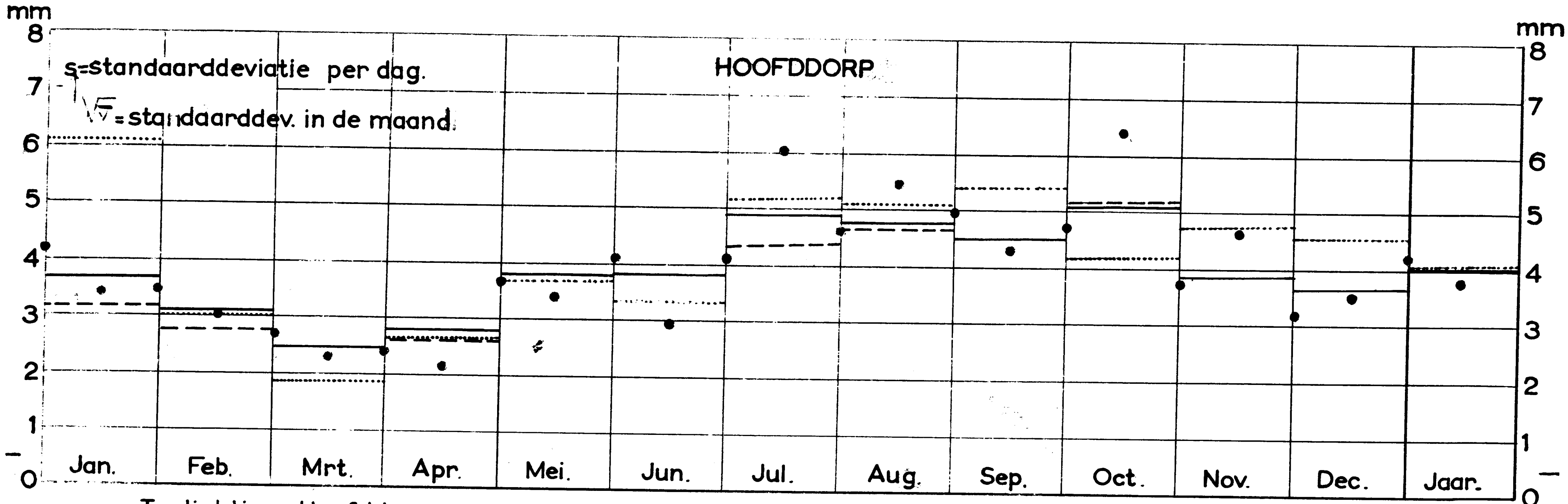



fig.5

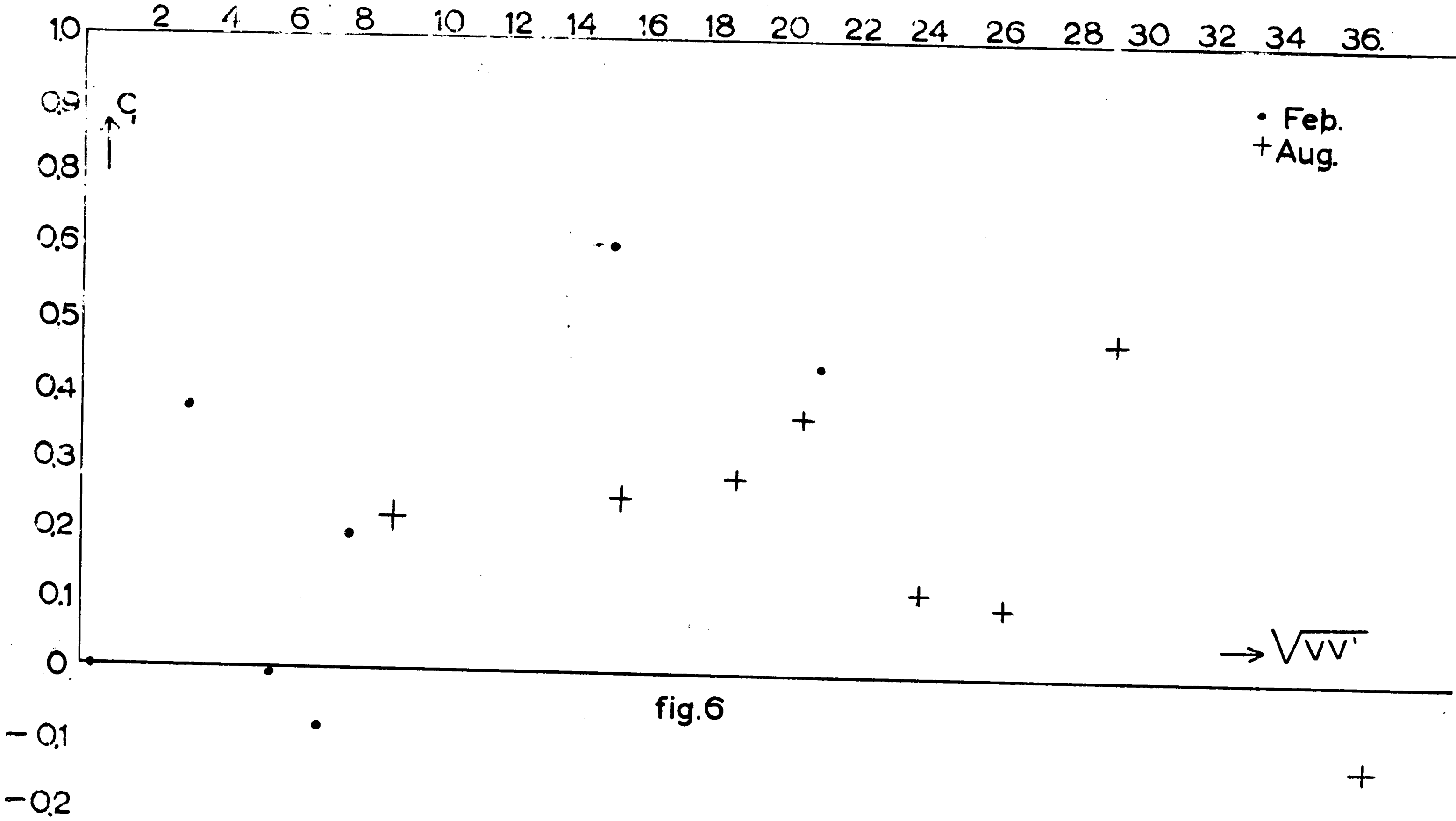
Toelichting : Hoofddorp

Elk punt is de combinatie van de datumvariantie van de 1$^{ste}$ en de 2$^{de}$ of van de 15de en de 16de van iedere maand, (tabel 10)

——— De gemiddelde s per maand, berekend uit die van de 1$^{ste}$, 2$^{de}$, 15$^{de}$, 16$^{de}$, 1$^{ste}$ en 2$^{de}$ (tabel 10)

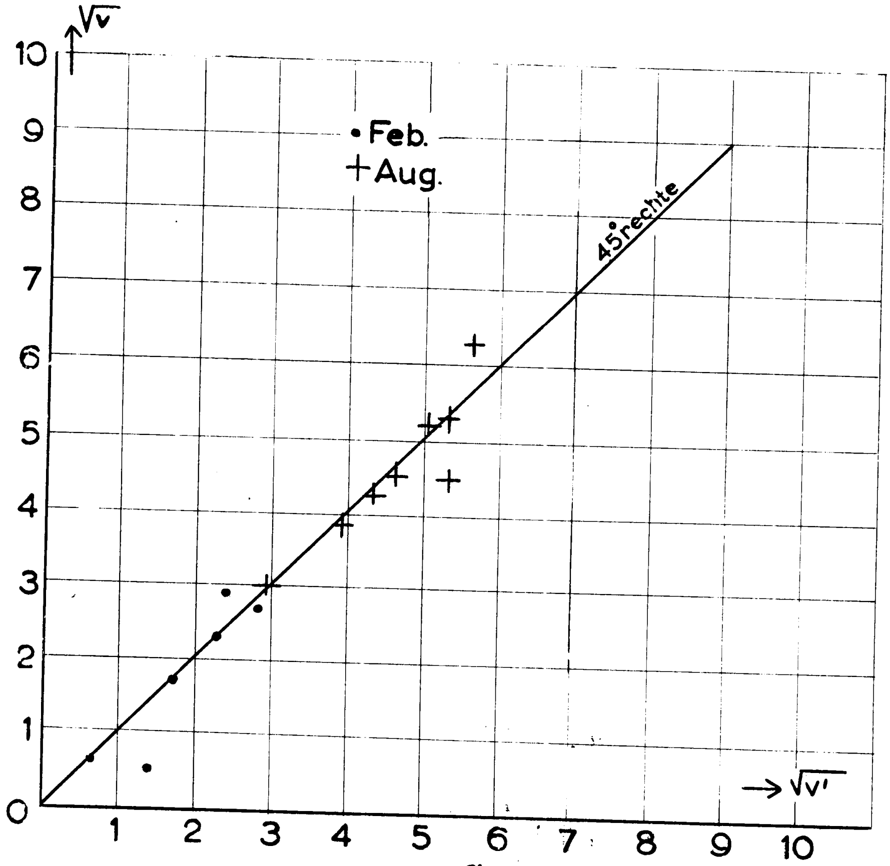
........ De $\sigma$ – waarde, berekend m.b.v. de 31 (30 of 28) 48 j-dagsommen (tabel 12)

– – – De $\sqrt{v}$ uit de maanden Jan, Feb, Apr, Jul, Aug, Oct (tabel 11)



10
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36.
0.9
0.8
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
C
• Feb.
+ Aug.
→ √vv'


fig.6



√v
10
9
8
7
6
5
4
3
2
1
0
1 2 3 4 5 6 7 8 9 10
• Feb.
+ Aug.
45° rechte
→ √v'


fig.7

Hoofddorp gem. daghoeveelheid neerslag
(gem. over 81 jaren)

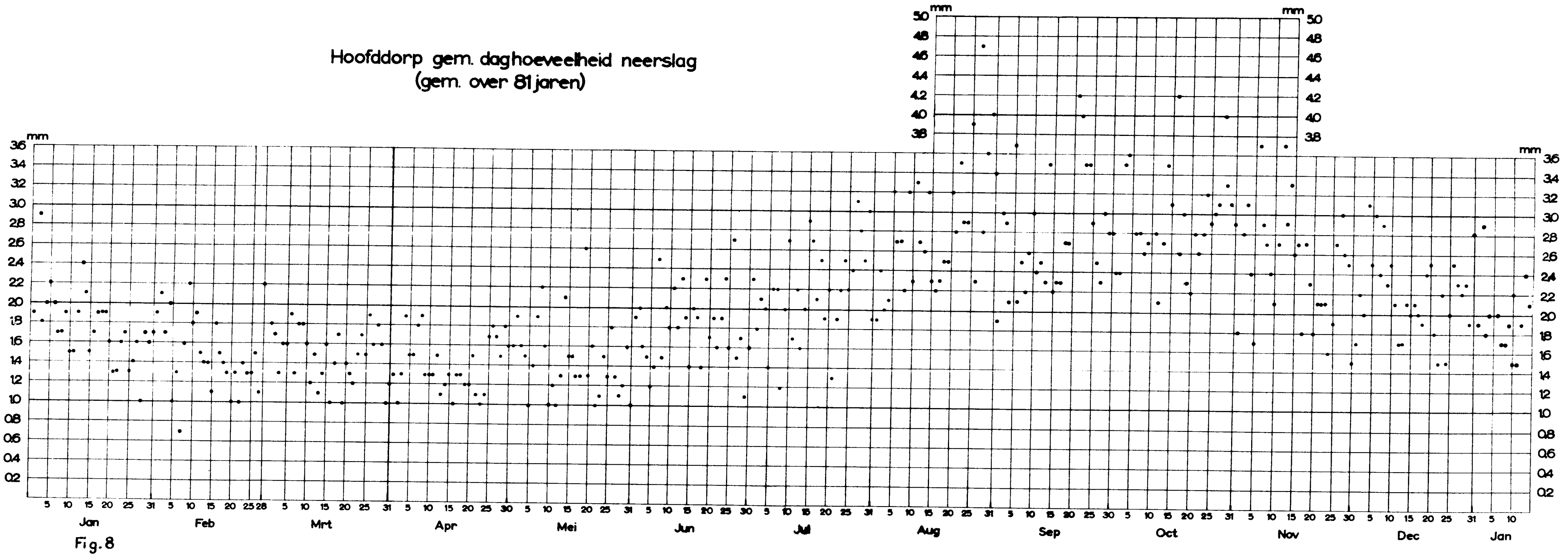



Fig. 8

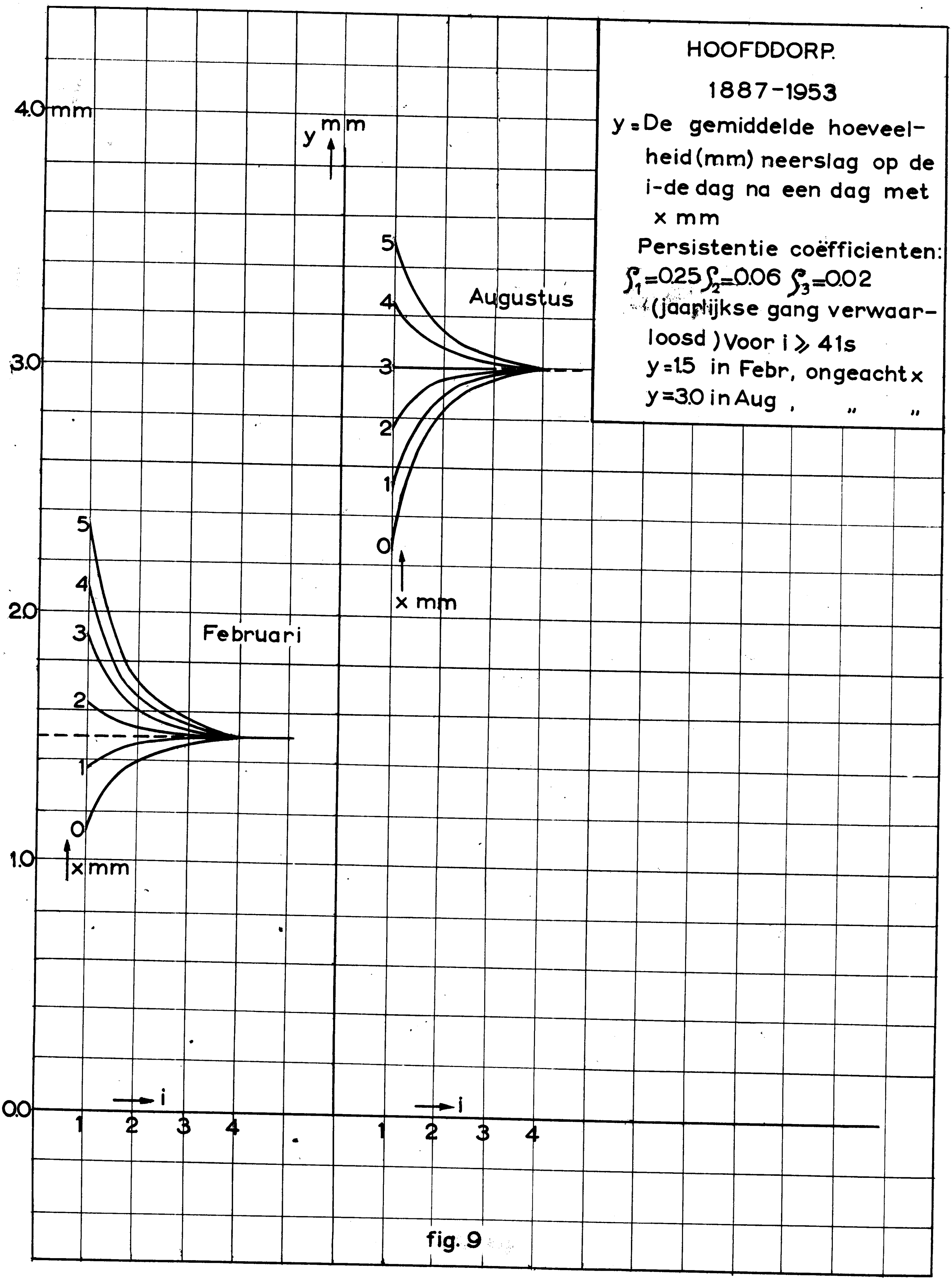

HOOFDDORP.
1887-1953
y = De gemiddelde hoeveelheid (mm) neerslag op de i-de dag na een dag met x mm
Persistentie coëfficienten:
$\rho_1$=0.25 $\rho_2$=0.06 $\rho_3$=0.02
(jaarlijkse gang verwaarloosd) Voor i ≫ 4 is
y = 1.5 in Febr, ongeacht x
y = 3.0 in Aug , " "
4.0 mm
3.0
2.0
1.0
0.0
y mm
x mm
Augustus
Februari
i
1 2 3 4
0 1 2 3 4 5

fig. 9