

KONINKLIJK NEDERLANDS
METEOROLOGISCH INSTITUUT

D e B i l t

WETENSCHAPPELIJK RAPPORT

W.R. 75-2

H.A.R. de Bruin

Over het interpoleren
van de neerslaghoogte

De Bilt, 1975

Publikationsnummer: K.N.M.I. W.R. 75-2 (MO)

U.D.C.: 551.501.4 :
551.577.2 :
551.579

Summary

The sub-group B of the Commission on Study of the Water Management of Gelderland has developed a mathematical hydrological model, which can simulate the groundwater level, the moisture state of the unsaturated zone, the evaporation and the surface run-off. Therefore, with this model it is possible to assess the consequences of human intervention, such as groundwater withdrawal.

One of the input parameters of the model is the precipitation depth, which has to be known at a number of grid-points every ten days.

The model will be applied to the eastern part of the Province of Gelderland. This area is covered by about 750 grid-points. As, generally, no rain-gauge is situated at these points, the precipitation depth must be found by interpolation.

In this report the method of interpolation developed by Eliassen, Drozdov and Sepelevskij is described and applied to the precipitation data of 35 rain-gauge stations. These data were processed for the period June 1966 to March 1972.

The error of interpolation was related to the distances between the rain-gauges, to the co-ordinates of the point of interpolation and to the error of observation (which seems to be 10-20 per cent).

Also the more general problem of the optimum density of a rain-gauge network is taken into consideration.

One of the conclusions of the study is that for ten-days periods the error of interpolation is about 20-25 per cent. It appears further that the relatively poor quality of the data does not make possible a good understanding of the statistical structure of precipitation.

Legends to figures

1. The correlation coefficient plotted against the distance between the rain-gauges, November - March.
2. idem , September, October, April and May.
3. idem , June - August.
4. As fig. 1, but for selected stations.
5. As fig. 2, but for selected stations.
6. As fig. 3, but for selected stations.
7. Error of interpolation plotted against the distance (one rain-gauge involved with the 'interpolation').
8. Estimated precipitation depth plotted against the measured depth of station 'Mauribaan'.
9. Places within a square (formed by 4 rain-gauges) to which the interpolation method has been applied.
10. Error of interpolation to the centre of a square (formed by 4 rain-gauges) related to the edge (l) of the square.
Curve I: error of observation 22 per cent,
Curve II: error of observation 15 per cent,
Curve III: error of observation 5 per cent.

Inhoud

1. Inleiding.
 2. Beschrijving van de interpolatiemethode.
 3. De correlatiefunctie.
 4. Controle van het interpolatiemodel.
 5. Enkele voorbeelden van berekende gewichtsfactoren.
 6. De optimale dichtheid van een regenmeternetwerk m.b.t. interpolatie van decadesommen.
 7. Uitbreiding naar gebiedsgemiddelde neerslag en naar andere tijdvaklengten.
 8. Conclusies.
 9. Naschrift.
- APPENDIX I.
- APPENDIX II.

Over het interpoleren van de neerslaghoogte

door

H.A.R. de Bruin

1. Inleiding

Door de ad-hoc groep B van de Commissie Bestudering Waterhuishouding Gelderland is een mathematisch-hydrologisch model ontwikkeld, dat de waterbeweging in zowel de verzadigde als de onverzadigde zone beschrijft als functie van een aantal tijdsafhankelijke (bijv. de temperatuur, de windsnelheid, de neerslaghoogte en de gewashoogte) en een aantal tijdsonafhankelijke parameters (bijv. bodemkundige gegevens). Met dit model is het in principe mogelijk naar plaats en tijd de grondwaterstand, de vochttoestand van de onverzadigde zone, de verdamping en de oppervlakkige afvoer te simuleren en zodoende de gevolgen van menselijk ingrijpen (zoals grondwateronttrekking) te bepalen.

De neerslaghoogte is één van de tijdsafhankelijke invoergegevens van het model; per tijdstap moet deze grootte op de roosterpunten, waarmee het beschouwde gebied wordt overdekt, bekend zijn.

Het ligt in de bedoeling het model toe te passen op geheel Oost-Gelderland. Het aantal roosterpunten, waarmee voor dit gebied zal worden gewerkt, bedraagt ongeveer 750. Omdat op al deze plaatsen de neerslag niet direkt wordt gemeten, zullen de neerslaggegevens door middel van interpolatie moeten worden bepaald uit de neerslagmetingen, welke door Rijkswaterstaat en het KNMI zijn en worden verricht op een 35-tal meetposten, die op kaart I zijn aangegeven.

Dit rapport handelt over de interpolatietechniek, die zal worden toegepast en geeft tevens een beschouwing over de fout, die hierbij gemiddeld zal worden gemaakt.

De ontwikkelde interpolatiemethode maakt het mogelijk in het algemeen uitspraken te doen betreffende de relatie tussen de dichtheid

van een regenmeternetwerk en de nauwkeurigheid waarmee de neerslaghoogte hieruit kan worden berekend; nagegaan zal worden wat de gevolgen zijn op de interpolatiefout van het verdichten c.q. het uitdunnen van een netwerk van neerslagstations. Bij deze beschouwing zal tevens de meetnauwkeurigheid worden betrokken.

2. Beschrijving van de interpolatiemethode

2.1 De interpolatieformule

Omdat fysisch en meteorologisch over de neerslagverdeling naar plaats en tijd kwantitatief erg weinig bekend is, is men voor het interpoleren van de neerslaghoogte aangewezen op zuiver statistische methoden. Hierbij kan men betrekkelijk willekeurig te werk gaan; de methode, waaraan hier de voorkeur wordt gegeven, is die van Eliassen [1] en van Drozdov en Sepelevskij [2], die de interpolatietechniek onafhankelijk van elkaar ontwikkelden.

Zij gaan ervan uit, dat de neerslaghoogte (of een andere meteorologische grootheid) op een willekeurige plaats, waar in het algemeen niet wordt gemeten, is te schrijven als een lineaire combinatie van de op de meetposten bepaalde neerslaghoogten (c.q. andere meteorologische parameters). Of, mathematisch geformuleerd: als $h_1, h_2 \dots$ en h_n de neerslaghoogten zijn, gemeten op de plaatsen $P_1, P_2 \dots$ en P_n en als we h_0 de op plaats P_0 te interpoleren neerslaghoogte noemen, wordt h_0 geschat met:

$$\hat{h}_0 = \sum_{i=1}^n \alpha_i h_i \quad (1)$$

De gewichtsfactoren α_i worden hierbij geacht onafhankelijk van de tijd te zijn; ze hangen echter wel af van zowel de plaatsen P_1 t/m P_n van de regenmeters als van de plaats P_0 , waarnaar moet worden geïnterpoleerd. In appendix I wordt aangetoond, dat de door (1) beschreven interpolatieformule qua vorm overeenkomt met een in de praktijk veel toegepaste techniek, waarbij de te schatten neerslaghoogte h_0 wordt geschreven als een lineaire functie van polynomen van de plaatscoördinaten van P_0 .

Van het gebied Oost-Gelderland wordt aangenomen, dat het klimatologisch homogeen en isotroop is; er wordt verondersteld, dat - hoewel de neerslaghoogte van dag tot dag en van plaats tot plaats (zelfs aanzienlijk) kan verschillen - gemiddeld over een lange tijd op elke plaats in het gebied evenveel neerslag zal vallen. Verder volgt uit de aanname van klimatologische homogeniteit, dat de frekwentieverdeling van de neerslaghoogte onafhankelijk is van de plaats in het gebied.

De veronderstelde isotropie houdt in, dat de correlatiecoëfficiënt tussen de neerslagreeksen van twee meetposten op afstand \vec{r} van elkaar verwijderd alleen van de absolute waarde van \vec{r} afhangt en niet van de richting van \vec{r} .

Samengevat komen de aannamen op het volgende neer:

$$\bar{h}_0 = \bar{h}_1 = \dots = \bar{h}_n \quad (\equiv \bar{h}) \quad (2)$$

$$\sigma_0 = \sigma_1 = \dots = \sigma_n \quad (\equiv \sigma) \quad (3)$$

$$\rho_{ij}(\vec{r}) = \rho(|\vec{r}|) = \rho(r) \quad (4)$$

Hierbij is σ_i de standaarddeviatie van een lange meetreeks bepaald op plaats P_i ($i = 1, 2, 3, \dots, n$). ρ_{ij} is de correlatiecoëfficiënt tussen de neerslagreeksen van plaatsen P_i en P_j ($i, j = 1, 2, 3, \dots, n$) en r de afstand tussen P_i en P_j . Het symbool '—' betekent 'gemiddeld over een (oneindig) lange tijd'.

2.2 Bepaling van de gewichtsfactoren

In het algemeen zal de werkelijk opgetreden neerslaghoogte h_0 verschillen van de via de hierboven beschreven interpolatietechniek verkregen waarde. De fout, die gemiddeld bij toepassing van (1) zal worden gemaakt, is te schrijven als:

$$F = \sqrt{(\hat{h}_0 - h_0)^2} \quad (5)$$

of wel
$$F^2 = \left(\sum_{i=1}^n \alpha_i h_i - h_0 \right)^2 \quad (6)$$

Uit (1) en (2) volgt, dat de verwachtingswaarde van \hat{h}_0 gelijk moet zijn aan \bar{h} , en dat dus bovendien

$$\sum_{i=1}^n \alpha_i = 1 \quad (7)$$

moet zijn. Uit (6) en (7) volgt:

$$F^2 = \left\{ \sum_{i=1}^n \alpha_i (h_i - h_0) \right\}^2 \quad (8)$$

Daar per definitie geldt dat:

$$\sigma_i^2 = \overline{(h_i - \bar{h}_i)^2} \quad (9)$$

en

$$\rho_{ij} = \frac{\overline{(h_i - \bar{h}_i)(h_j - \bar{h}_j)}}{\sigma_i \sigma_j} \quad (10)$$

kan met behulp van (2) en (3) vergelijking (8) worden geschreven als:

$$F^2 = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \sigma^2 \left[\rho_{ij} - \rho_{oi} - \rho_{oj} + 1 \right] \quad (11)$$

De methode van Eliassen en van Drozdov en Sepelevskij komt er nu op neer, dat de gewichtsfactoren α_i zo worden gekozen, dat F^2 minimaal is, of wel zodanig dat

$$\frac{\partial F^2}{\partial \alpha_k} = 0 \quad (12)$$

voor $k = 1, 2, 3, \dots, n$. *)

Betrekking (12) levert ons in principe n vergelijkingen in $\alpha_1, \alpha_2, \dots$ en α_n . We moeten ons echter bedenken, dat we verondersteld hebben dat Oost-Gelderland een klimatologisch homogeen gebied is, welke

*) Theoretisch dient ook nog te worden bewezen, dat de determinant van de matrix van de tweede afgeleiden (de Hessiaanse matrix) positief definitief is, zodat de oplossing van (12) een minimum is. Omdat F^2 een kwadratische functie is van α_k , bij $\alpha_{i \neq k} = \text{constant}$, wordt hier altijd aan voldaan.

aanname leidde tot (7). Deze extra voorwaarde, die aan de gewichts-
factoren wordt opgelegd, vermindert het aantal vrijheidsgraden in
feite met één. Dit is te ondervangen door in (11)

$$\alpha_n = 1 - \sum_{i=1}^{n-1} \alpha_i \quad \text{te substitueren en daarna de eis}$$

$$\frac{\partial F^2}{\partial \alpha_k} = 0 \quad \text{voor } k = 1, 2, \dots, n-1$$

op te leggen. We verkrijgen zo een stelsel van n-1 lineaire ver-
gelijkingen in $\alpha_1, \alpha_2, \dots, \alpha_{n-1}$, dat er als volgt uitziet:

$$\left. \begin{aligned} \sum_{i=1}^{n-1} \alpha_i \left[\rho_{ik} - \rho_{in} - \rho_{kn} + 1 \right] &= \rho_{ok} - \rho_{on} - \rho_{kn} + 1 \\ k &= 1, 2, 3, \dots, n-1 \\ \alpha_n &\text{ volgt uiteraard uit:} \\ \alpha_n &= 1 - \sum_{i=1}^{n-1} \alpha_i \end{aligned} \right\} \quad (13)$$

Opgemerkt moet worden, dat bij de afleiding van (13) er per definitie
van is uitgegaan dat $\rho_{ii} = 1$ voor $i = 1, 2, \dots, n$.

3. De correlatiefunctie

3.1 Inleiding

Met behulp van het stelsel lineaire vergelijkingen, dat wordt
beschreven door (13), is het mogelijk de neerslaghoogte in elk punt
van het beschouwde gebied door middel van interpolatieformule (1)
te bepalen, op voorwaarde echter, dat de correlatiecoëfficiënt
als functie van de afstand bekend is.

Is deze relatie eenmaal vastgelegd, dan stelt bovendien formule (11)
ons in staat de interpolatiefout F te bepalen. Tevens is het dan
mogelijk de dichtheid van een regenmeternetwerk als functie van F
te berekenen en zijn wij dus in het algemeen in staat het probleem
van de optimale dichtheid van een regenmeternetwerk aan te vatten.

Het verband tussen de correlatiecoëfficiënt en de afstand werd langs empirische weg verkregen door de decadesommen van de 35 neerslagstations, die in Oost-Gelderland aanwezig zijn, te bewerken voor de periode juni 1966 tot en met maart 1972; de meetreeks van elk station werd gecorreleerd met die van de 34 overige. Op deze wijze konden 595 correlatiecoëfficiënten worden berekend. Deze werden grafisch uitgezet tegen de afstand tussen de bijbehorende meetposten, welke afstand varieerde van ca. 700 meter tot ongeveer 41 kilometer. In de figuren 1, 2 en 3 zijn de resultaten te vinden voor drie tijdvakken in het jaar. Alvorens deze resultaten te bespreken, behandelen we eerst een aantal zaken, die voor de hier aan de orde gestelde problematiek van belang zijn, te weten:

1. Er is een jaarlijkse gang aanwezig in de mate waarmee de correlatiecoëfficiënt met de afstand afneemt.
2. Meetreeksen met een eindige lengte geven aanleiding tot een toevallige fout in de berekende correlatiecoëfficiënten.
3. De meetnauwkeurigheid beïnvloedt de berekende correlatiecoëfficiënten.

Het is zinvol deze punten hier één voor één te behandelen.

3.2 Het seizoenseffect

Het blijkt uit bijvoorbeeld het werk van Colenbrander en Stol [5], dat de mate waarin de correlatiecoëfficiënt met de afstand afneemt, een jaarlijkse gang ondergaat. In de zomer is deze afname groter dan in de winter. De verklaring voor dit verschijnsel wordt gevonden in het feit, dat in de zomer regen uit zowel buien als fronten valt, terwijl de neerslag in de winter voornamelijk een frontaal karakter heeft. Buien hebben afmetingen van de orde van grootte van 2-20 kilometer, terwijl fronten 200 tot 2000 kilometer lang zijn. Wegens dit seizoenseffect werd het waarnemingsmateriaal in drie groepen opgesplitst:

- a) november t/m maart;
- b) september, oktober, april en mei;
- c) juni t/m augustus.

De aantallen decaden, die in deze groepen voorkwamen, bedroegen resp. 90, 66 en 54.

De indeling werd gemaakt mede op grond van de resultaten van het werk van Stol [6], die het gedrag van correlatiecoëfficiënten tussen reeksen van dagsommen bestudeerde.

Een indeling per maand kon niet worden gemaakt, omdat dan het aantal waarnemingen in een groep te klein zou worden, en daardoor de onbetrouwbaarheid van de berekende correlatiecoëfficiënten te groot.

3.3 Invloed van een eindige meetreeks

Een correlatiecoëfficiënt, die berekend wordt uit N paren van waarnemingen van bijvoorbeeld twee neerslagstations, is op te vatten als een steekproef met omvang 1 uit een populatie van een groot aantal correlatiecoëfficiënten, die worden verkregen door herhaald sets van N paren van waarnemingen met elkaar te correleren. Door toevallige factoren zullen deze correlatiecoëfficiënten $\rho_1, \rho_2, \rho_3, \dots$ niet aan elkaar gelijk zijn. Men kan aantonen (zie bijv. [3]) dat, indien de twee variabelen, die met elkaar worden gecorreleerd, beide normaal verdeeld zijn, de verdelingsfunctie van de correlatiecoëfficiënten $\rho_1, \rho_2, \rho_3, \dots$ alleen afhangt van N - het aantal paren van waarnemingen - en van ρ^* - de 'werkelijke correlatiecoëfficiënt'. Verder kan men bewijzen, dat voor $N \geq 11$ de variabele z, die gedefinieerd wordt door $z_i = \frac{1}{2} \ln \left(\frac{1+\rho_i}{1-\rho_i} \right)$ normaal verdeeld is met verwachtingswaarde

$$\mathcal{E}(z) = \frac{1}{2} \ln \frac{1+\rho^*}{1-\rho^*} \quad (14)$$

en standaarddeviatie

$$\sigma_z = \frac{1}{\sqrt{N-3}} \quad (15)$$

Zoals eveneens in [3] wordt vermeld, blijkt het dat (14) en (15) ook redelijk geldig zijn voor het geval dat de basis-variabelen niet normaal verdeeld zijn. Dit is bijvoorbeeld met de decadesommen van de neerslag het geval.

Met (14) en (15) kunnen we nu voor verschillende waarden van ρ^* en N de grenzen uitrekenen waarbinnen 95% van alle z_1, z_2, z_3, \dots vallen en dus de grenzen waarbinnen 95% van de correlatiecoëfficiënten $\rho_1, \rho_2, \rho_3, \dots$ zijn gelegen. De grenswaarden van z (z_- en z_+) zijn resp.:

$$z_- = \frac{1}{2} \ln \frac{1 + \rho^*}{1 - \rho^*} - \frac{2}{\sqrt{N - 3}} ; \quad z_+ = \frac{1}{2} \ln \frac{1 + \rho^*}{1 - \rho^*} + \frac{2}{\sqrt{N - 3}}$$

en die van de correlatiecoëfficiënten (ρ_- en ρ_+)

$$\rho_- = \tanh z_- \quad \text{en} \quad \rho_+ = \tanh z_+$$

In tabel I worden voor een aantal waarden van ρ^* en N de 95%-grenzen ρ_- en ρ_+ gegeven.

Tabel I

Grenzen waarbinnen 95% van de 'gemeten' correlatiecoëfficiënten is gelegen voor verschillende waarden van N en ρ^* .

N	$\rho^* = 0.98$		$\rho^* = 0.95$		$\rho^* = 0.90$		$\rho^* = 0.80$	
	ρ_-	ρ_+	ρ_-	ρ_+	ρ_-	ρ_+	ρ_-	ρ_+
54	0.965	0.989	0.914	0.971	0.831	0.942	0.675	0.881
66	0.967	0.988	0.919	0.970	0.840	0.938	0.689	0.874
90	0.970	0.987	0.924	0.967	0.851	0.934	0.709	0.869
210	0.972	0.986	0.928	0.965	0.858	0.930	0.723	0.858

Uit deze tabel blijkt ten eerste dat de correlatiecoëfficiënten scheef zijn verdeeld en ten tweede dat hoe meer de werkelijke correlatiecoëfficiënt van 1 afwijkt, hoe groter de spreiding is.

Uit deze beschouwing mogen we concluderen, dat een deel van de spreiding in de correlatiecoëfficiënten, welke in de figuren 1, 2 en 3 valt te bespeuren, moet worden toegeschreven aan het feit, dat de meetreeksen waaruit ze berekend werden een eindige lengte bezitten.

3.4 Invloed van de meetnauwkeurigheid

Bij de afleiding van (13) hebben we in eerste instantie de meet- en waarnemingsfouten buiten beschouwing gelaten. De vergelijkingen (13) moeten dan ook zo worden gelezen, dat de grootheden, die erin voorkomen, zijn afgeleid van de gemeten neerslagdecadesommen en niet van die, welke werkelijk zijn opgetreden.

Om de invloed van de meet- en waarnemingsfouten te onderzoeken, voeren we in:

$$h_i = \beta_i h_i^* + \epsilon_i \quad (16)$$

Hierin is h_i de gemeten en h_i^* de werkelijk opgetreden neerslaghoogte op het i^e station en karakteriseren β_i en ϵ_i de toevallige en de systematische meet- en waarnemingsfouten.

Een voorbeeld van een systematische meetfout is die, welke het gevolg is van het zgn. 'windeffect'. Dit wordt veroorzaakt door het feit dat een regenmeter, die boven maaiveld is opgesteld, door zijn stoffelijke aanwezigheid het windveld in zijn directe omgeving dusdanig verstoort, dat er systematisch te weinig regendruppels door de opvangtrechter worden opgevangen. Deze fout kan op ca. -6% [5] worden geschat (gemiddeld over het jaar). Ook kunnen door een onjuiste opstelling van de regenmeters systematische meetfouten worden geïntroduceerd.

De toevallige meet- en waarnemingsfouten ontstaan bijvoorbeeld tijdens het aftappen, bij het aflezen van het maatglas of bij het noteren en dupliceren van de meetgegevens.

Opgemerkt moet worden, dat ook systematische fouten een 'toevallig' karakter kunnen hebben. Een voorbeeld is de hierboven genoemde systematische windfout. Deze hangt namelijk - naast de vorm en de opstelhoogte van de regenmeter - af van de windsnelheid en de regendruppelgrootte-verdeling. Deze laatste twee grootheden zijn tijdsafhankelijk. Dit betekent, dat door het windeffect weliswaar steeds te weinig regendruppels worden opgevangen, maar dat het van de meteorologische omstandigheden afhangt 'hoeveel te weinig' neerslag door de regenmeter zal worden opgevangen.

Om een indruk te verkrijgen in hoeverre β_i en ϵ_i de uitkomsten van vergelijkingen (13) beïnvloeden, doen we de volgende veronderstellingen:

$$\left. \begin{aligned} \beta_1 = \beta_2 = \dots = \beta_n \equiv \beta \\ \overline{\varepsilon_i} = 0 \end{aligned} \right\} \quad (17)$$

Bovendien wordt aangenomen dat ε_i onafhankelijk is van ε_j (voor alle $i \neq j$) en van h_k^* (voor alle k), terwijl tevens $\overline{\varepsilon_i^2}$ voor alle stations gelijk wordt verondersteld, of wel:

$$\overline{\varepsilon_k h_k} = 0 \quad (18)$$

$$\overline{\varepsilon_k \varepsilon_k} = 0 \quad (19)$$

$$\overline{\varepsilon_k \varepsilon_k} = \Delta^2$$

Met behulp van (16), (17), (18) en (19) kunnen we afleiden, onder de aanname dat $\sigma_1^* = \sigma_2^* = \dots = \sigma_n^* \equiv \sigma^*$ (waarbij σ_i^* de 'werkelijke' standaarddeviatie is van de meetpost i), dat:

$$\rho_{ij} = \rho_{ij}^* \frac{(\beta \sigma^*)^2}{(\beta \sigma^*)^2 + \Delta^2} \quad (20)$$

Hierin is ρ_{ij}^* de correlatiecoëfficiënt, die volgt uit de op de stations i en j werkelijk opgetreden neerslaghoogten h_i^* en h_j^* .

We kunnen constateren dat meet- en waarnemingsfouten altijd aanleiding geven tot een verlaging van de berekende correlatiecoëfficiënten en daardoor in feite - zoals uit paragraaf 3.3 blijkt - tot vergroting van de spreiding in deze grootheid.

We mogen verwachten, dat ρ_{ij}^* tot 1 nadert, indien de afstand tussen de twee meetposten naar 0 gaat. Uit (20) volgt, dat de 'gemeten' correlatiecoëfficiënt ρ_{ij} dan de limietwaarde

$$\frac{(\beta \sigma^*)^2}{(\beta \sigma^*)^2 + \Delta^2}$$

heeft. Definieren we deze waarde als ρ_0 en noemen we $\lambda = \frac{\Delta}{\beta \sigma^*}$, dan geldt dat

$$\rho_0 = \frac{1}{1 + \lambda^2} \quad (21)$$

In tabel II is ρ_0 gegeven voor een aantal waarden van λ .
Zij $\lambda' = \frac{\Delta}{\sigma^*}$ (λ' zouden we de 'toevallige procentuele meetfout' kunnen noemen), dan is: $\lambda' = \beta\lambda$.

In de praktijk zal de windfout één van de belangrijkste systematische meetfouten zijn. Dit betekent dat β van de orde van grootte van 0.94 is. Uit tabel II volgt dan dat de invloed van β op ρ_0 klein is in vergelijking met λ' op ρ_0 . Bij de correlatiecoëfficiënt - en daardoor bij vergelijkingen (13) - spelen blijkbaar voornamelijk toevallige meetfouten een rol.

Deze conclusie houdt in, dat systematische meetfouten bij het interpolatieprobleem geen belangrijke rol spelen. Dit betekent uiteraard niet dat ze voor het hydrologische model niet van belang zouden zijn. Integendeel, omdat de neerslag vaak de grootste post op de waterbalans is, hebben systematische meetfouten belangrijke consequenties voor de andere hydrologische grootheden, die door het model worden beschreven.

Tabel II

λ (%)	ρ_0
1	~ 1 (0.9999)
5	~ 1 (0.9975)
10	0.99
14	0.98
20	0.96
25	0.94
30	0.92
40	0.86
50	0.80

ρ_0 kan langs empirische weg worden verkregen door de gemeten correlatiecoëfficiënten tegen de afstand uit te zetten en vervolgens te extrapoleren naar afstand gelijk aan 0. Zoals we straks zullen zien, worden op deze wijze waarden van ρ_0 verkregen, die liggen tussen 0.95 en 0.99. Dit betekent dat λ ongeveer ligt tussen 10 en 20%.

Experimenten in De Bilt uitgevoerd hebben aangetoond, dat voor dag-sommen de toevallige instrumentele fouten van de orde van grootte

van enkele procenten zijn [4]. Voor decadesommen zal dat nog minder zijn. Dit wijst erop dat de λ -waarden van 10 tot 20%, welke voor het waarnemingsmateriaal van Oost-Gelderland werden gevonden, voornamelijk moeten worden toegeschreven aan waarnemingsfouten.

We hebben nogal wat aannamen moeten doen om tot deze conclusie te komen en zolang deze aannamen niet op hun geldigheid zijn getoetst, bestaat er geen zekerheid omtrent de juistheid van de conclusie. Er is echter nog een andere aanwijzing dat waarnemingsfouten inderdaad verantwoordelijk zijn voor de betrekkelijk hoge λ -waarden, die experimenteel werden gevonden. Door de Provinciale Waterstaat Gelderland werd op het gehele waarnemingsmateriaal (de dagsommen van juni 1966 tot begin 1974) een kwaliteitscontrole uitgeoefend. Bijgehouden werd per station het aantal onbetrouwbare waarnemingen. De meetposten met minder dan 20 (een willekeurig gekozen grens) onbetrouwbare waarnemingen werden apart bestudeerd. In de figuren 4, 5 en 6 zijn de correlatiecoëfficiënten van deze stations tegen de afstand uitgezet (wederom voor decadesommen) voor dezelfde tijdvakken van het jaar als die, welke bij de figuren 1, 2 en 3 werden beschouwd.

Vergelijken we figuur 1 met figuur 4, figuur 2 met figuur 5 en figuur 3 met figuur 6, dan kunnen we het volgende constateren:

1. De correlatiecoëfficiënten van de geselecteerde stations zijn systematisch hoger dan die van de overige.
2. Dit effect is het duidelijkst in de 'wintermaanden' november tot en met maart, is iets minder evident, doch wel aanwezig voor de 'voor- en najaarsmaanden' april, mei, september en oktober, terwijl het niet meer significant aanwezig is in de 'zomermaanden' juni, juli en augustus.
3. Voor de 'wintermaanden' wordt voor de geselecteerde stations een waarde van 0.99 gevonden, evenals trouwens voor de andere twee perioden. Dit komt overeen met een waarde van λ van 10%. Voor het gehele materiaal wordt voor resp. de winter-, de voor- en najaars- en de zomerperiode een λ -waarde van 22, 14 en 10 gevonden.

Dit is dus inderdaad een aanwijzing, dat waarnemingsfouten de hoge λ -waarden veroorzaken. Dat het effect in de 'zomermaanden' niet duidelijk aantoonbaar is, kan worden verklaard door het feit dat in de zomer de correlatiecoëfficiënten lager zijn door het 'seizoens-effect'. Bovendien is het aantal beschikbare waarnemingen voor de zomermaanden het laagst (54). Beide factoren maken dat de spreiding van de berekende correlatiecoëfficiënten in de zomermaanden het grootst is, wat tot gevolg kan hebben dat het effect, dat meet- en waarnemingsfouten op correlatiecoëfficiënten hebben, als het ware 'in de ruis' verdwijnt.

Het bleek dat tot de 14 stations, die werden geselecteerd, alle de in Oost-Gelderland aanwezige KNMI-stations behoorden. Dit kan erop wijzen, dat de kwaliteitscontrole, die het KNMI - in het kader van de controle van verzamelde neerslaggegevens met het landelijk meet-net - continu uitvoert op deze stations, redelijk doeltreffend is. (Bij deze controle wordt in het geval van onregelmatigheden een vrij intensief contact met de waarnemers onderhouden).

Een belangrijke conclusie, die uit het voorgaande kan worden getrokken, is dat er aanwijzingen zijn, dat er bij het verzamelen van neerslaggegevens in Oost-Gelderland vrij aanzienlijke waarnemingsfouten zijn gemaakt, welke het in dit rapport beschreven neerslagonderzoek ernstig hebben bemoeilijkt.

Aangezien men voor het verkrijgen van een duidelijk inzicht in het statistisch gedrag van de neerslag over betrouwbare neerslaggegevens zal moeten beschikken, is derhalve een tweede belangrijke conclusie dat het waarnemingsmateriaal, in Oost-Gelderland verzameld, ons dit duidelijke inzicht niet kan verschaffen.

Wil men de statistische structuur van de neerslag beter leren kennen dan nu mogelijk is, dan zal men een nieuw regenmeternetwerk moeten inrichten. Om menselijke fouten zoveel mogelijk te elimineren, zal dit netwerk uit automatische, goed opgestelde, grondregometers moeten zijn opgebouwd. De afstand tussen de apparaten zal moeten variëren van 100 meter tot ca. 10 kilometer. Het netwerk zal tenminste 5 jaar in stand moeten worden gehouden.

3.5 Benaderingsformule

Om met het stelsel lineaire vergelijkingen, dat door (13) wordt beschreven, in de praktijk te kunnen werken, is het prettig over een analytische functie te beschikken, die het verband tussen de correlatiecoëfficiënten met de afstand weergeeft. Het is gebruikelijk de correlatiecoëfficiënt te schrijven als:

$$\rho(r) = \rho_0 e^{-r/r_0^*} \quad (22)$$

Dit wordt bijvoorbeeld gedaan door Stol [6], Rusin [7] en Kagan [8]. In (22) is r weer de afstand en zijn ρ_0 en r_0^* constanten. Oost-Gelderland heeft afmetingen van ongeveer 30 bij 40 kilometer. De door Stol [6] gevonden waarden r_0^* voor dagsommen van 0.5 mm en groter, zijn dusdanig dat bij goede benadering voor (22)

$$\rho(r) = \rho_0 - \rho_0 \frac{r}{r_0^*} \quad (23)$$

geschreven kan worden voor afstanden tot 40 kilometer. Relatie (23) wordt verkregen door de e^{-r/r_0^*} in een Taylor-reeks te ontwikkelen en deze na de lineaire term af te breken. Op grond van deze beschouwing en op grond van de berekende correlatiecoëfficiënten, die in de figuren 1 t/m 6 zijn weergegeven, kan worden geconcludeerd, dat

$$\rho(r) = \rho_0 - \frac{r}{r_0} \quad \text{met} \quad r_0 = \frac{r_0^*}{\rho_0} \quad (24)$$

een redelijke beschrijving is van de correlatiecoëfficiënt als functie van de afstand.

Voor Oost-Gelderland werden de constanten ρ_0 en r_0 uit de figuren 1, 2 en 3 verkregen; 'op het oog' werd door elk van de drie puntenwolken een rechte lijn geconstrueerd. Op deze wijze werd gevonden:

voor periode I (nov. t/m maart) : $\rho_0 = 0.95$; $r_0 \geq 1000$ km;
voor periode II (sept., okt., apr., mei): $\rho_0 = 0.98$; $r_0 = 270$ km;
voor periode III (juni t/m aug.) : $\rho_0 = 0.99$; $r_0 = 190$ km.

Wanneer werd afgezien van het seizoenseffect, bleken $\rho_0 = 0.96$ en $r_0 = 350$ km te zijn.

Uit figuur 1 is te zien dat voor de wintermaanden er slechts een geringe afname van de correlatiecoëfficiënt met de afstand is te bespeuren. Dus r_0 is zeer groot.

Beschouwen we het geval dat $r_0 \gg r$, dan gaat (24) over in:

$$\rho(r) = \rho_0$$

Substitueren we dit in (13), dan verkrijgen we (ter herinnering zij vermeld dat per definitie $\rho_{ii} = 1$):

$$\sum_{i \neq k}^{n-1} \alpha_i (1 - \rho_0) + 2 \alpha_k (1 - \rho_0) = 1 - \rho_0 \quad (k = 1, 2, \dots, n-1)$$

Hieruit volgt dat:

$$\sum_{i=1}^{n-1} \alpha_i + \alpha_k = 1$$

hetgeen betekent, dat:

$$\alpha_k = 1 - \sum_{i=1}^{n-1} \alpha_i = \alpha_n$$

Aangezien deze uitkomst voor $k = 1, 2, \dots, n$, heeft dit tot gevolg dat:

$$\alpha_1 = \alpha_2 = \dots = \alpha_n = \frac{1}{n}$$

Met andere woorden: men kan in de winter de neerslaghoogte in de roosterpunten van het rekenmodel het beste schatten met een rekenkundig gemiddelde van de op de neerslagstations gemeten neerslaghoogten.

Door de in 3.1 en 3.2 genoemde oorzaken is de betrouwbaarheid van de hier gevonden waarden van ρ_0 en r_0 niet erg groot. Ze moeten meer als richtgetallen worden opgevat.

Vermeld moet nog worden, dat het aantal geheugenplaatsen van de computer beperkt was. Voor de perioden I en II (dus voor de maanden

april t/m oktober) zijn bij het interpoleren daarom niet alle neerslagstations betrokken, doch alleen de vier meetposten, die het dichtst bij het beschouwde roosterpunt zijn gelegen.

(Voor periode I, de 'wintermaanden', werden bij het interpoleren wel alle stations betrokken, omdat zoals reeds vermeld, de neerslaghoogte op alle roosterpunten dan met het gemiddelde van de gemeten waarden werd geschat).

4. Controle van het interpolatiemodel

Op zichzelf is een controle van de hier beschreven interpolatiemethode eenvoudig uit te voeren. We kunnen namelijk interpoleren naar een plaats waarop een regenmeter is geïnstalleerd, waarna wij vervolgens de gemeten en de geschatte neerslaghoogten met elkaar kunnen vergelijken.

Oplösungen van vergelijkingen (13) worden bepaald door resp.: de afstanden tussen het interpolatiepunt tot de regenmeters, de afstanden tussen de regenstations onderling (de geometrie van het regenmeternet), ρ_0 en r_0 . Betrekken we bij het 'interpoleren' slechts één meetstation, dan treden alleen de afstand tussen dit meetstation en het interpolatiepunt en de constanten ρ_0 en r_0 als bepalende grootheden op.

Dit eenvoudige geval werd gebruikt voor de controle van het interpolatiemodel, omdat dan de - moeilijk te beschrijven - 'geometrie van het regenmeternet' niet meer als variabele optreedt.

Schattingsformule (1) gaat voor dit geval over in:

$$\hat{h}_0 = \alpha_1 h_1 \quad (25)$$

Met (7) geeft dit $\alpha_1 = 1$, waardoor $\hat{h}_0 = h_1$.

Vergelijkingen (11) en (24) gaan dan over in:

$$F^2 = 2 \sigma^2 \left(1 - \rho_0 + \frac{r}{r_0} \right) \quad (26)$$

Deze theoretische interpolatiefout F werd vergeleken met de experimenteel gevonden waarden. Deze laatste werden verkregen door voor alle mogelijke combinaties van i en j en alle neerslaggegevens van de 210 decaden de grootheid

$$F_{ij}^2 = \frac{1}{210} \sum_{k=1}^{210} (h_i^k - h_j^k)^2 \quad (27)$$

te berekenen. (h_i^k is de neerslaghoogte opgetreden in de k^e decade op het i^e meetstation).

In figuur 7 zijn de zo gevonden waarden van F_{ij} tegen de afstand tussen het i^e en het j^e neerslagstation uitgezet.

Bovendien werd in figuur 7 de theoretische kromme getekend, die tot stand kwam door in (26) de voor alle decaden gevonden waarden van $\rho_0 = 0.96$ en $r_0 = 350$ km - de waarden, die werden gevonden voor deze constanten, als we van het seizoenseffect afzagen, te substitueren. Uit figuur 7 blijkt dat, hoewel de spreiding rond de theoretische lijn betrekkelijk groot is, wat natuurlijk weer het gevolg is van het feit dat de correlatiecoëfficiënt zo spreidt en het statistisch interpolatiemodel een benadering is, de aanpassing toch redelijk mag worden genoemd.

Opmerking : De hierboven beschreven controle is uitgevoerd op hetzelfde waarnemingsmateriaal als waarmee de constanten ρ_0 en r_0 werden bepaald. De controle 'bewijst' daarom alleen dat de interpolatiemethode redelijke resultaten oplevert voor de periode juni 1966 - maart 1972. Dit is op zichzelf niet erg, omdat het hydrologische model op ongeveer dezelfde periode zal worden toegepast.

Wil men nagaan of de interpolatiemethode ook voor andere perioden geldig is, dan zal men het waarnemingsmateriaal in twee delen moeten opsplitsen; met het eerste gedeelte bepaalt men dan ρ_0 en r_0 , terwijl men met het tweede gedeelte de controle uitvoert. Het beschikbare waarnemingsmateriaal was echter te klein van omvang om deze procedure te kunnen volgen.

Ter illustratie wordt hier tenslotte nog een voorbeeld gegeven van een toepassing van de interpolatiemethode, waarbij vier neerslagstations werden betrokken. In figuur 8 zijn de gemeten neerslaghoogten (decadesommen) van het station Mauribaan tegen de geïnterpoleerde neerslaghoogten uitgezet, welke laatste werden verkregen uit de neerslaghoogten van resp. Lichtenvoorde, de Wiersse, te Stroete en Nijman. Deze stations zijn op afstanden van 6.7 tot 9,2 km van Mauribaan gelegen. Het bleek dat de interpolatiefout F 4.1 mm bedroeg. Dit is ongeveer 21 procent van de (middelbare) standaarddeviatie.

Langs theoretische weg werd op grond van $\rho_0 = 0.96$ en $r_0 = 350$ km met behulp van (11) een waarde van F/σ verkregen van ca. 25%. Dit is wederom een aanwijzing, dat de interpolatiemethode voor praktische doeleinden redelijk bruikbaar is.

Figuur 8 wekt de indruk, dat de interpolatiefout in feite nog een functie is van de neerslaghoogte zelf: de spreiding rond de 45° -lijn neemt iets toe met de neerslaghoogte. Dit heeft uiteraard consequenties voor de juiste interpretatie van de grootheid F , waarmee hier steeds is gewerkt. In Appendix II zal op dit probleem nader worden ingegaan en een mogelijke oplossing worden gegeven.

5. Enkele voorbeelden van berekende gewichtsfactoren

Om enig inzicht te verkrijgen in het gedrag van de gewichtsfactoren, welke met de in dit rapport beschreven interpolatiemethode worden verkregen, gaan we van de situatie uit waarbij de regenmeters in de hoekpunten van een vierkant, met ribbe van 20 km, zijn geplaatst. Voor een vijftal punten, die resp. A, B, C, D en E worden genoemd en die in figuur 9 zijn aangegeven, werden met behulp van vergelijkingen (13) en (17) voor $\rho_0 = 1$ en 0.95 en $r_0 = 100$ en 500 km de gewichtsfactoren α_1 , α_2 , α_3 en α_4 berekend. Tevens werd met behulp van vergelijking (11) de relatieve interpolatiefout F/σ bepaald. De resultaten zijn in tabel III weergegeven.

In de laatste kolom van tabel III is tevens de waarde van de relatieve interpolatiefout F^*/σ gegeven, welke wordt verkregen, indien men de gewichtsfactoren α_1 , α_2 , α_3 en α_4 eenvoudig 0.25 kiest.

Uit tabel III kunnen we het volgende concluderen:

1. De methode ter bepaling van de gewichtsfactoren voldoet aan de eisen, die de wetten der symmetrie ons opleggen; neerslagstations op gelijke afstand van het interpolatiepunt gelegen, hebben gelijke gewichtsfactoren. (Zie bijv. punten A, B, C en E).
2. De dichtstbijgelegen neerslagstations verkrijgen de hoogste gewichtsfactoren. De gewichtsfactoren van verderweg gelegen meetposten nemen echter toe, indien ρ_0 af- en r_0 (bij $\rho_0 \neq 1$) toeneemt. Dit is ook in overeenstemming met wat men moet

Tabel III

	α_1	α_2	α_3	α_4	$\bar{F}/\sigma \cdot 100$	ρ_0	r_0 (km)	$(F^*/\sigma, 100)$
A	0.250	0.250	0.250	0.250	34	1	100	34
A	0.250	0.250	0.250	0.250	15	1	500	15
A	0.250	0.250	0.250	0.250	42	0.95	100	42
A	0.250	0.250	0.250	0.250	29	0.95	500	29
B	0.469	0.469	0.031	0.031	31	1	100	39
B	0.469	0.469	0.031	0.031	14	1	500	18
B	0.436	0.436	0.064	0.064	41	0.95	100	46
B	0.366	0.366	0.134	0.134	30	0.95	500	31
C	0.571	0.179	0.071	0.179	30	1	100	36
C	0.571	0.179	0.071	0.179	14	1	500	16
C	0.512	0.200	0.081	0.200	40	0.95	100	44
C	0.405	0.227	0.140	0.227	29	0.95	500	30
D	0.723	0.229	0.016	0.031	27	1	100	40
D	0.723	0.229	0.016	0.031	12	1	500	18
D	0.634	0.250	0.034	0.082	39	0.95	100	47
D	0.476	0.264	0.101	0.159	29	0.95	500	31
E	0.908	0.042	0.008	0.042	16	1	100	40
E	0.908	0.042	0.008	0.042	7	1	500	18
E	0.778	0.104	0.014	0.104	33	0.95	100	47
E	0.555	0.183	0.078	0.183	29	0.95	500	31

verwachten. Als het namelijk inderdaad zo is, dat het feit dat ρ_0 kleiner is dan 1, moet worden toegeschreven aan toevallige waarnemingsfouten, zullen zoveel mogelijk meetposten bij de interpolatie moeten worden betrokken om deze waarnemingsfouten zoveel mogelijk uit te middelen. Dit houdt in, dat bij afnemende ρ_0 (dus een toename van de waarnemingsfout) verderweg gelegen stations een hoger gewicht zullen moeten krijgen, hetgeen inderdaad het geval is.

Verder houdt een toename van r_0 in, dat verderweg gelegen stations ons meer informatie zullen verschaffen, waardoor zij derhalve ook een hoger gewicht zullen moeten krijgen. Ook dit is het geval.

3. Een vergelijking tussen de vijfde en de laatste kolom van tabel III leert ons, dat in het geval van $\rho_0 = 0.95$ en $r_0 = 500$ km de relatieve interpolatiefout F/σ en F^*/σ niet veel van elkaar verschillen. Hieruit kan worden geconcludeerd dat bij een lage ρ_0 -waarde en een grote r_0 de gewichtsfactoren even goed 0.25 kunnen worden gekozen. (Voor het geval dat r_0 zeer groot is, hadden we dat reeds ontdekt in paragraaf 3.4).

6. De optimale dichtheid van een regenmeternetwerk m.b.t. interpolatie van decadesommen

Met de hierboven verkregen resultaten is het nu mogelijk ons te buigen over het probleem van de optimale dichtheid van een regenmeternetwerk. Om de beschouwingen niet te gecompliceerd te maken, gaan we, evenals in hoofdstuk 5, uit van de situatie waarbij de regenmeters in de hoekpunten van een vierkant (met ribbe-lengte 1) zijn opgesteld. Bovendien zullen we ons beperken tot het geval, dat we alleen naar het middelpunt van dit vierkant interpoleren. We houden dan als bepalende grootheden over: ρ_0 en r_0 en de lengte van de ribbe van het vierkant, 1. De gewichtsfactoren zullen voor de vier regenstations gelijk zijn en wel de waarde 0.25 bezitten.

Met behulp van (11) kunnen we dan de relatieve interpolatiefout bepalen, waarvoor in dit geval geldt:

$$\frac{F}{\sigma} = \sqrt{(1 - \rho_0)1.25 + 0.56 \frac{1}{r_0}} \quad (28)$$

In figuur 10 is voor $\rho_0 = 0.96$ en $r_0 = 350$ km deze grootheid als functie van 1 getekend (kromme I). Uit deze krommen zijn een aantal interessante conclusies te trekken:

a. De waarde van F/σ is ook bij zeer hoge netwerkdichtheden nog van de orde van grootte van 20-25%. Hierbij moet nog worden vermeld, dat voor decadesommen in Nederland σ ongeveer gelijk is aan de gemiddelde neerslaghoogte.

Het is dus bijvoorbeeld niet mogelijk een hydrologisch model, waarvoor neerslagdecadesommen met een nauwkeurigheid beter

dan 20% worden vereist, op het materiaal dat in Oost-Gelderland is verzameld toe te passen.

- b. Kromme I verloopt nogal vlak; dit betekent dat het verdichten van het regenmeternetwerk relatief weinig winst t.a.v. de interpolatiefout oplevert. Ter illustratie geven we het volgende voorbeeld: Stel, we verkleinen l van 40 km tot 20 km, wat neerkomt op een verviervoudiging van het aantal regenmeters per km^2 ; F/σ neemt dan af van ongeveer 34% tot ca. 29%. Economisch bezien is dit erg onvoordelig.

Uit paragraaf 3.3 kwam naar voren, dat er duidelijke aanwijzingen zijn dat de waarnemingsfouten verantwoordelijk zijn voor de lage waarde van ρ_0 . Hoewel hieromtrent geen definitief uitsluitsel valt te verkrijgen, is het interessant te onderzoeken wat de invloed zou zijn op F/σ , indien we er inderdaad, door verbetering van de kwaliteit van de waarnemingen, in zouden slagen de waarde van ρ_0 te verhogen.

Stel, het lukt ons op deze wijze λ van 20% (welke waarde overeenkomt met een ρ_0 van 0.96) te verlagen tot 14%, wat volgens tabel II een waarde van $\rho_0 = 0.98$ oplevert.

Voor het gemak veronderstellen we dat r_0 niet zal veranderen. (In feite verandert deze ook iets - zie bijv. (20), maar dit effect is klein).

We kunnen dan wederom met (28) F/σ als functie van l berekenen. Het resultaat is weergegeven in figuur 10, kromme II.

Tenaanzien van de krommen I en II kunnen we het volgende opmerken:

1. Ook kromme II verloopt betrekkelijk vlak. Bij $l = 5$ km bedraagt F/σ nog ca. 18%.
2. Indien men naar een verlaging van F/σ streeft, kan het voordeliger zijn de kwaliteit van de waarnemingen te verbeteren dan dat men het netwerk verdicht. (Wat in de praktijk dikwijls betekent dat de kwaliteit van de waarnemingen gemiddeld achteruit gaat).

Dit wordt zeer duidelijk geïllustreerd met het volgende voorbeeld: Stel, we gaan uit van kromme I bij $l = 20$ km, dan heeft een verkleining van l tot ca. 5 km - wat overeenkomt met meer

dan een verzeestienvoudiging van het aantal regenmeters per km^2 - hetzelfde effect als wanneer we bij $l = 20$ van kromme I naar kromme II kunnen gaan, omdat we erin geslaagd zijn λ van 20% op 14% te brengen door verbetering van de kwaliteit van de waarnemingen. Het kan dus economisch veel voordeliger zijn, in plaats van een bestaand netwerk te verdichten, de kwaliteit van de waarnemingen te verhogen.

In figuur 10 hebben we tevens F/σ als functie van l getekend voor $\lambda = 5\%$ ($\rho_0 = 0.9975$) en $r_0 = 350$ km (kromme III). Een vergelijking tussen de krommen I en III leert ons, dat het zelfs voordeliger kan zijn een regenmeternet uit te dunnen, mits we er tegelijkertijd maar voor zorgen dat de kwaliteit van de overblijvende stations wordt verbeterd.

Stel namelijk, dat we een (dicht) regenmeternetwerk hebben met $\lambda = 20\%$ ($\rho_0 = 0.96$) en $l = 5$ km (1 regenmeter per 25 km^2). Uit kromme I blijkt, dat dan de relatieve interpolatiefout ca. 24% bedraagt. Ditzelfde resultaat wordt verkregen met een regenmeternet met $\lambda = 5\%$ en $l = 35$ km (1 regenmeter per 1225 km^2).

Er zij nog opgemerkt, dat de geometrie van het regenmeternet invloed op de interpolatiefout zal hebben; in het algemeen zal deze toenemen als het netwerk niet regelmatig is opgebouwd. De invloed van dit effect is hier niet onderzocht, omdat het begrip 'onregelmatigheid' zo moeilijk in algemene termen is te beschrijven, maar in ieder geval zal het zo zijn, dat ernaar moet worden gestreefd een netwerk zo regelmatig mogelijk in te richten. Dit blijkt speciaal van belang te zijn, indien gebiedsgemiddelde neerslag met het regenmeternet moet worden bepaald.

7. Uitbreiding naar gebiedsgemiddelde neerslag en naar andere tijdvaklengten

Tot slot van dit rapport willen we een aantal punten naar voren brengen, die nog niet zijn behandeld, doch die een belangrijke rol spelen bij het probleem van de optimale dichtheid van een regenmeternetwerk:

1. Dit rapport handelt in eerste instantie over het interpoleren van de neerslaghoogte. Dit komt, omdat, in verband met het hydrologisch model van de Commissie Bestudering Waterhuishouding Gelderland, dit probleem moest worden aangevat.

Vaak zal echter de neerslaghoogte gemiddeld over een zeker gebied worden verlangd. Het is in principe mogelijk de, in dit rapport beschreven, interpolatietechniek uit te breiden tot een schattingsmethode van gebiedsgemiddelde neerslag, waarbij het tevens mogelijk is een schatting te geven van de fout, die men gemiddeld maakt bij het toepassen van deze methode.

Deze uitbreiding is door de schrijver van dit rapport nog niet geheel uitgewerkt; wel is het mogelijk ter illustratie enkele voorlopige resultaten te geven.

Stel, er is een gebied G met oppervlak S gegeven, dat is opgebouwd uit N congruente vierkanten. In het middelpunt van deze vierkanten is een regenmeter opgesteld. We schatten de gebiedsgemiddelde neerslag met het rekenkundig gemiddelde van de N gemeten neerslaghoogten. Hierbij maken we gemiddeld een fout F_S . F_S zal een functie zijn van N , van S en, evenals dit bij het interpolatieprobleem het geval was, van de wijze waarop de correlatiecoëfficiënt van de afstand afhangt, ofwel van de constanten ρ_0 en r_0 .

In tabel III.a is F_S (wederom uitgedrukt als een percentage van de standaarddeviatie σ) gegeven voor $\rho_0 = 0.96$, $r_0 = 350$ km en $S = 1000$ km². In tabel III.b is hetzelfde gedaan voor $S = 100$ km². Deze tabellen gelden weer voor decadesommen. Omdat ze eigenlijk buiten het bestek van dit rapport vallen, worden deze voorlopige resultaten hier zonder afleiding en zonder commentaar gegeven.

Tabel III

a. $S = 1000$ km ²		b. $S = 100$ km ²	
<u>N</u>	<u>$(F_S/\sigma) \cdot 100$</u>	<u>N</u>	<u>$(F_S/\sigma) \cdot 100$</u>
1	25	1	21
10	11	10	8
100	5	30	5
1000	0.8	100	3

2. Het hydrologisch model van de Commissie Bestudering Waterhuishouding Gelderland werkt met een tijdstap van 10 dagen. Dit is de reden dat dit rapport in eerste instante over decadesommen handelt.

Het is uiteraard mogelijk de in dit rapport beschreven interpolatiemethode toe te passen op andere tijdvaklengten.

Omdat blijkt, dat opeenvolgende decadesommen statistisch onafhankelijk van elkaar zijn, is dit voor tijdvakken met een lengte, welke een geheel aantal maal een decade is, eenvoudig. Stel namelijk, dat we bij decadesommen een relatieve fout F/σ van gemiddeld $p\%$ maken, dan zal deze voor maandsommen $p/\sqrt{3}\%$ en voor jaarsommen ongeveer $p/\sqrt{36}\%$ bedragen.

Aangezien geldt, dat de standaarddeviatie van maandsommen $\sqrt{3}$ maal groter is dan de standaarddeviatie van decadesommen en die van jaarsommen ongeveer $\sqrt{36}$ maal, dan volgt hieruit dat de absolute fout F , die bij het interpoleren wordt gemaakt, voor resp. decade-, maand- en jaarsommen ongeveer gelijk is, en wel ongeveer 4-6 mm. Dit is resp. ca. 25, 10 en 1% van de gemiddelde neerslaghoogte.

Opeenvolgende dagsommen zijn statistisch niet meer van elkaar onafhankelijk. Voor perioden korter dan 10 dagen zullen we daarom de hierboven beschreven procedure moeten herhalen, wat er in feite op neerkomt dat voor bijv. 1, 2, 3, 5 en 7-daagse perioden de relatie tussen correlatiecoëfficiënt en de afstand afzonderlijk zal moeten worden vastgesteld.

Voor decadesommen geldt dat de standaarddeviatie σ en de gemiddelde neerslaghoogte \bar{h} ongeveer aan elkaar gelijk zijn. Voor dagsommen geldt echter dat $\sigma \approx 3\bar{h}$. Aangezien de interpolatiefout recht evenredig is met σ , volgt hieruit dat het interpoleren van dagsommen slechts met een beperkte mate van nauwkeurigheid mogelijk zal zijn. Dit wordt nog versterkt door het feit, dat de ruimtelijke variabiliteit van dagsommen groter zal zijn dan die van decadesommen.

8. Conclusies

1. Neerslag is van nature sterk variabel naar plaats en tijd. Men is daarom bij het interpoleren van de neerslaghoogte aangewezen op zuiver statistische methoden. Voor het verkrijgen van een duidelijk inzicht in het statistisch gedrag van de neerslag, moet men over betrouwbare waarnemingen beschikken, anders zijn natuurlijke fluctuaties in het neerslagpatroon niet meer te onderscheiden van fluctuaties, die door meet- en waarnemingsfouten worden geïntroduceerd. Uit dit rapport blijkt dat de neerslaggegevens, die sinds 1966 in Oost-Gelderland zijn verzameld, niet betrouwbaar zijn; ze zijn waarschijnlijk behept met een gemiddelde waarnemingsfout van 10 tot 20%. Het is dus niet mogelijk, op grond van het in Oost-Gelderland verzamelde waarnemingsmateriaal, een duidelijk inzicht in het statistisch gedrag van de neerslag te verkrijgen. Wil men alsnog dit inzicht verkrijgen, dan zal een dicht netwerk van automatische (grond)regenmeters moeten worden ingericht. Hierbij zal de afstand tussen de apparaten moeten variëren van 100 m tot ca. 10 km. Het net zal tenminste 5 jaar in stand moeten worden gehouden.
2. Bij toepassing van de in dit rapport beschreven interpolatiemethode op het waarnemingsmateriaal van Oost-Gelderland maakt men bij decadesommen een absolute fout van 4-6 mm. Procentueel is dit ongeveer 20 tot 25% van de standaarddeviatie en omdat deze ongeveer gelijk is aan het gemiddelde voor decadesommen, ook ongeveer 20 tot 25% van het gemiddelde. De absolute interpolatiefout voor maand- en jaarsommen bedraagt eveneens ongeveer 4-6 mm. Dit is ca. 10 en 1% van resp. de gemiddelde maand- en de gemiddelde jaarsom. De vraag of de hier gevonden interpolatiefouten acceptabel zijn of niet kan door de schrijver van dit rapport niet worden beantwoord, omdat dit antwoord afhangt van het doel waarvoor de neerslaggegevens worden gebruikt en van de nauwkeurigheidseisen, die eraan worden gesteld.

Eén van de stappen, die in het kader van de werkzaamheden van de ad-hoc Groep B van de Commissie Bestudering Waterhuishouding Gelderland zal moeten worden gedaan, is het uitvoeren van een gevoeligheidsanalyse op het hydrologisch model om de nauwkeurigheidseisen, die door het model aan de neerslaggegevens worden gesteld, te bepalen.

3. Blijkt het, dat de interpolatiefout, die wordt gemaakt, te groot is voor het doel waarvoor de neerslaggegevens worden gebruikt, dan kan men in het algemeen twee stappen ondernemen:
 - a. het bestaande regenmeternet verdichten;
 - b. de kwaliteit van het regenmeternet verbeteren, (D.w.z. de meet- en waarnemingsfouten van de puntwaarnemingen verkleinen).

Uit dit rapport blijkt, dat t.a.v. het regenmeternet in Oost-Gelderland het laatste de voorkeur verdient.

9. Naschrift

De schrijver dankt Dr. P.J. Rijkoort van het KNMI, Ir. Ph.Th. Stol van het Instituut voor Cultuurtechniek en Waterhuishouding en Ir. H.J. Colenbrander en zijn medewerkers van de Provinciale Waterstaat Gelderland voor het grondig doorlezen van het manuscript en voor het plaatsen van kritische kanttekeningen.

Voorts wordt dank gebracht aan Drs. S. Kruizinga en de Heer G.J. Yperlaan voor de vele adviezen op statistisch gebied, die de schrijver van hen heeft ontvangen en voor de vele leerzame discussies, welke hij met hen over het behandelde onderwerp heeft gevoerd.

Appendix I

Interpoleren met polynomen

Een in de praktijk veel toegepaste interpolatietechniek is die, waarbij - om het algemeen te stellen - grootheid z , welke een functie is van de plaatscoördinaten x , y en van de tijd t , op de willekeurige plaats (x_0, y_0) wordt geschat met:

$$\hat{z}(x_0, y_0, t) = \sum_{i=0}^N a_i(t) P_i(x_0, y_0) \quad A(1)$$

waarbij $P_i(x_0, y_0)$ een machtsfunctie is van x_0 en y_0 .

We meten $z(x, y, t)$ (gelijktijdig op tijdstip t) op m plaatsen met de coördinaten (x_1, y_1) , (x_2, y_2) (x_m, y_m) .

We noemen deze gemeten waarden resp. $z_1(t)$, $z_2(t)$ $z_m(t)$.

Het is nu mogelijk, indien $m \geq N$, de coëfficiënten $a_i(t)$ te bepalen met behulp van de methode der kleinste kwadraten; de coëfficiënten worden zo gekozen, dat de grootheid

$$U = \sum_{j=1}^m \{ z_j(t) - \hat{z}(x_j, y_j, t) \}^2$$

minimaal wordt.

Dit houdt in, dat we eisen dat:

$$\frac{\partial U}{\partial a_i(t)} = 0, \quad i = 0, 1, 2, \dots, N \quad A(2)$$

Dit levert ons het volgend stelsel van $N+1$ lineaire vergelijkingen op in $a_0(t)$ t/m $a_N(t)$:

$$\sum_{j=1}^m z_j(t) P_k(x_j, y_j) = \sum_{i=0}^N a_i(t) \sum_{j=1}^m P_i(x_j, y_j) P_k(x_j, y_j) \quad A(3)$$

$k = 0, 1, 2, \dots, N$

We voeren in:

$$f_{ik} = \sum_{j=1}^m P_i(x_j, y_j) P_k(x_j, y_j)$$

$$i, k = 0, 1, 2, \dots, N$$

f_{ik} is alleen een functie van de plaatscoördinaten $(x_1, y_1) \dots (x_m, y_m)$

Vergelijkingen A(3) kunnen nu worden geschreven als:

$$\begin{array}{ccccccc} a_0(t) f_{00} + a_1(t) f_{10} + \dots + a_N(t) f_{NO} & = & \sum_{j=1}^m z_j(t) P_0(x_j, y_j) & & & & \\ \vdots & & \vdots & & \vdots & & \\ a_0(t) f_{0N} + a_1(t) f_{1N} + \dots + a_N(t) f_{NN} & = & \sum_{j=1}^m z_j(t) P_N(x_j, y_j) & & & & \end{array} \quad A(4)$$

Zij Δ de determinant:

$$\Delta = \begin{vmatrix} f_{00} & f_{10} & \dots & f_{NO} \\ f_{01} & f_{11} & \dots & f_{N1} \\ \vdots & \vdots & & \vdots \\ f_{0N} & f_{1N} & \dots & f_{NN} \end{vmatrix}$$

en Δ_k de determinant, die ontstaat als we de k^e kolom van Δ vervangen door:

$$\begin{array}{c} Q_0 \\ \vdots \\ Q_N \end{array}$$

waarbij

$$Q_k = \sum_{j=1}^m z_j(t) P_k(x_j, y_j)$$

dan is volgens de regel van Cramer:

$$a_k(t) = \frac{\Delta_k}{\Delta} \quad A(5)$$

Opgemerkt moet worden, dat Δ alleen een functie is van (x_1, y_1) t/m (x_m, y_m) en dat we Δ_k kunnen schrijven als:

$$\Delta_k = \sum_{l=0}^N Q_l (-1)^{k+l} \Lambda_{kl} \quad A(6)$$

waarbij Λ_{kl} de determinant is, die ontstaat door in Δ_k de k^e kolom en de l^e rij weg te laten, hetgeen betekent dat ook Λ_{kl} alleen afhankelijk is van de plaatscoördinaten (x_1, y_1) t/m (x_m, y_m) en onafhankelijk van $z_1(t)$ t/m $z_m(t)$.

Verder merken we op, dat Q_k voor $k = 0, 1, \dots, N$ een lineaire combinatie is van $z_1(t)$ t/m $z_m(t)$.

Uit A(6) volgt dan dat ook Δ_k dit is, en dús $a_k(t)$ ook, omdat Δ alleen van de plaatscoördinaten van de meetposten afhangt.

Uit A(5) volgt vervolgens, dat $a_k(t)$ in het algemeen als volgt kan worden geschreven:

$$a_k(t) = \sum_{j=1}^m c_{kj} z_j(t) \quad A(7)$$

waarbij c_{kj} alleen een functie kan zijn van de plaatscoördinaten van de meetposten, (x_1, y_1) t/m (x_m, y_m) .

Substitueren we A(7) in A(1), dan verkrijgen we:

$$\hat{z}(x_0, y_0, t) = \sum_{i=0}^N \left(\sum_{j=1}^m c_{ij} z_j(t) \right) P_i(x_0, y_0) \quad A(8)$$

wat na verwisseling van sommatie over i en j leidt tot:

$$\hat{z}(x_0, y_0, t) = \sum_{j=1}^m \left(\sum_{i=0}^N c_{ji} P_i(x_0, y_0) \right) z_j(t) \quad A(9)$$

Definiëren we $\alpha_j = \sum_{i=0}^N c_{ji} P_i(x_0, y_0)$ dan gaat dit over in:

$$\hat{z}(x_0, y_0, t) = \sum_{j=1}^m \alpha_j z_j(t)$$

waarbij α_j alleen een functie is van de coördinaten van de plaats waarnaar we interpoleren en van de plaats van de meetposten.

Hiermee is bewezen, dat de interpolatieformule (1), waarmee in dit rapport wordt gewerkt, wat betreft zijn wiskundige vorm, volkomen identiek is met formule A(1).

-o-o-o-

Appendix II

De interpretatie van de grootheid F

Uit figuur 8 kwam naar voren, dat de interpolatiefout in feite een functie is van de neerslaghoogte zelf. Omdat deze eigenschap de interpretatie van de grootheid F, welke een centrale rol speelt bij de in dit rapport besproken interpolatiemethode, bemoeilijkt, geven we in deze appendix een mogelijkheid aan om de interpolatiefout als functie van de neerslaghoogte h_0 en de grootheid F te schrijven.

Volgens (5) is F gedefinieerd als

$$F^2 = \overline{(h_0 - \hat{h}_0)^2}$$

Het symbool '—' betekent 'gemiddeld over een (oneindig) lange reeks'. Als $f(h_0, \hat{h}_0)$ de verdelingsfunctie is van h_0 en \hat{h}_0 , mogen we F^2 ook schrijven als:

$$F^2 = \int_0^\infty \int_0^\infty (h_0 - \hat{h}_0)^2 f(h_0, \hat{h}_0) dh_0 d\hat{h}_0$$

Kruizinga [9] vond langs empirische weg dat

$$\mathcal{E}^2(h_0) = \int_0^\infty (h - \hat{h}_0)^2 f(\hat{h}_0 | h_0) d\hat{h}_0 = c h_0$$

waarbij c een constante is en $f(\hat{h}_0 | h_0)$ de voorwaardelijke verdelingsfunctie van \hat{h}_0 bij h_0 is constant.

$\mathcal{E}(h_0)$ is de fout, die men gemiddeld maakt bij de toepassing van de interpolatiemethode, indien de werkelijke neerslaghoogte h_0 is.

Er geldt, dat

$$f(\hat{h}_0 | h_0) = \frac{f(h_0, \hat{h}_0)}{f(h_0)}$$

waarbij

$$f(h_0) = \int_0^\infty f(h_0, \hat{h}_0) d\hat{h}_0$$

Dus

$$\int_0^{\infty} (h_0 - \hat{h}_0)^2 \frac{f(h_0, \hat{h}_0)}{f(h_0)} d\hat{h}_0 = c h_0$$

waaruit volgt dat

$$F^2 \equiv \int_0^{\infty} c h_0 f(h_0) dh_0 = c \bar{h}_0$$

Hieruit volgt dat

$$c = \frac{F^2}{\bar{h}_0}$$

waardoor

$$\mathcal{E}^2(h_0) = F^2 \frac{h_0}{\bar{h}_0}$$

-o-o-o-

Literatuur

- [1] Eliassen, A. : "Provisional report on calculation of spatial covariance and autocorrelation of the pressure field".
Inst. Weather and Climate Res., Norw. Acad. of Science and Letters, Rep. No. 5, 1954.
zie tevens : WMO No. 94, T.P.38, Techn. Note No. 29, 1960.
- [2] Drozdov, O.A, and Sepelevskij, A.A. : "The Theory of the Interpolation of Meteorological Elements in a Stochastic Field and their Application for Questions of Weather Maps and Network Rationalization".
Trudy NIV/GUGMS, Series I, No. 13, 1946.
zie tevens
Gandin, L.S. : "On the Planning of Meteorological Station Networks".
WMO No. 265, T.P.149, Techn. Note No.111, 1970.
- [3] Kendall, M.G. : "The Advanced Theory of Statistics".
Vol. I, Hafner Publ. Comp. New York, 1952.
- [4] Denkema, A. : "Representativiteit van een puntmeting".
KNMI W.R. 70-6, 1970.
- [5] Colenbrander, H.J. : "Neerslag en neerslagverdeling naar plaats en en Stol, Ph.Th. tijd".
Comm. Waterhuishouding Gelderse Landbouwgronden, 5e deelrapport v.h. 2e interimrapport, 1970.
- [6] Stol, Ph.Th. : "Dagelijkse neerslaghoeveelheden in de Gelderse Achterhoek in verband gebracht met de onderlinge afstand van de plaats van meting".
ICW nota 426, 1967.
- [7] Rusin, N.P. : "Estimation of Areal Precipitation".
Symp. on Distribution of Precipitation in Mountain Areas, Geilo, Norway, 1972, WMO No. 526, Vol. I, 1973.
- [8] Kagan, R.L. : "Precipitation - Statistical Principles".
Casebook on Hydr. Netw. Desing Prac., I-1.1, WMO No. 324, 1972.
- [9] Kruizinga, S. : Mondelinge mededeling.

KAART 1

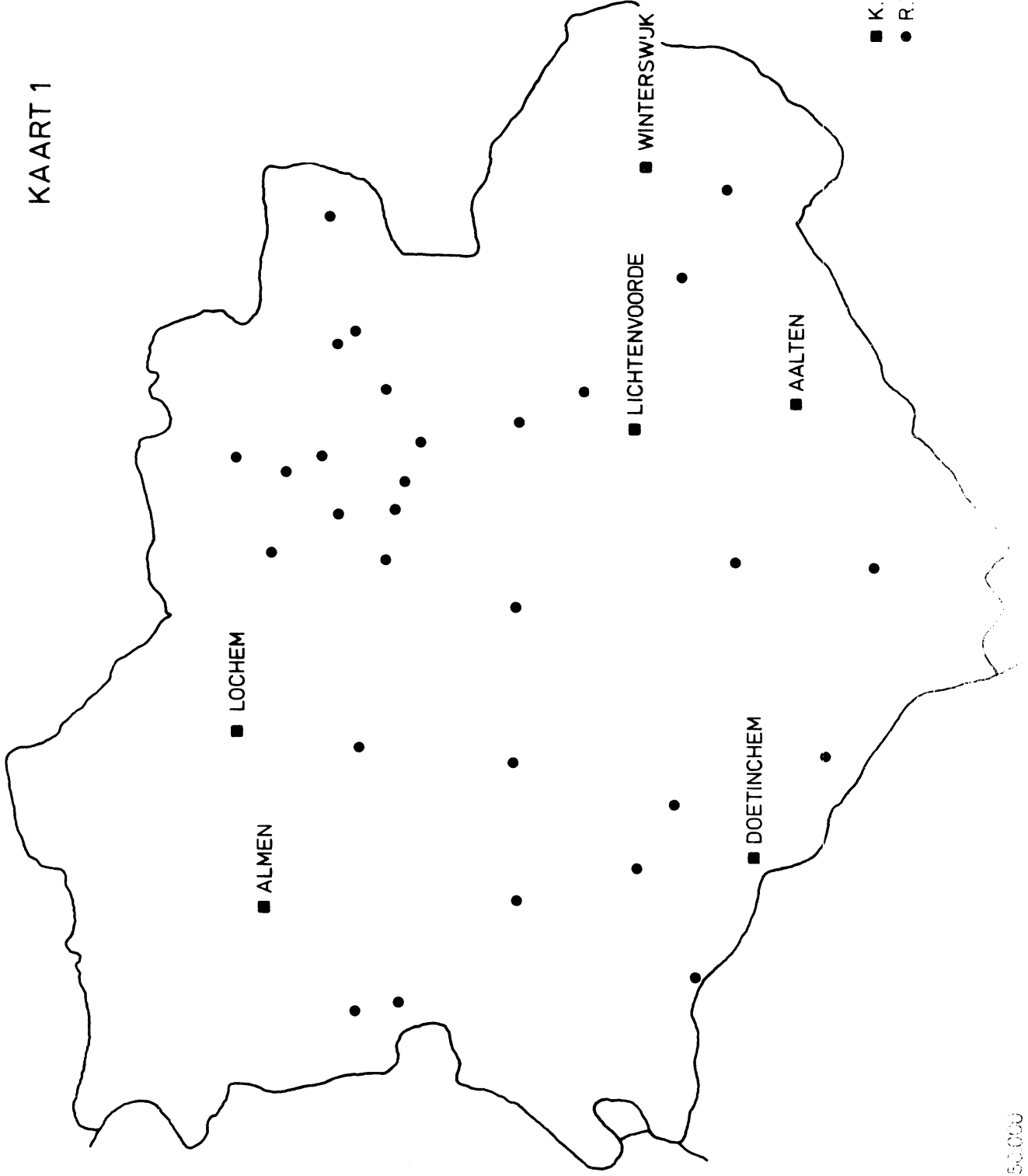
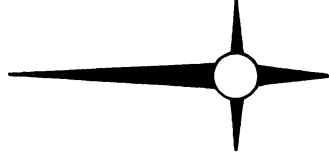


Fig.1 nov. t/m mrt.

$\rho_0 \approx 0.95$ $r_0 \approx 1000$ km

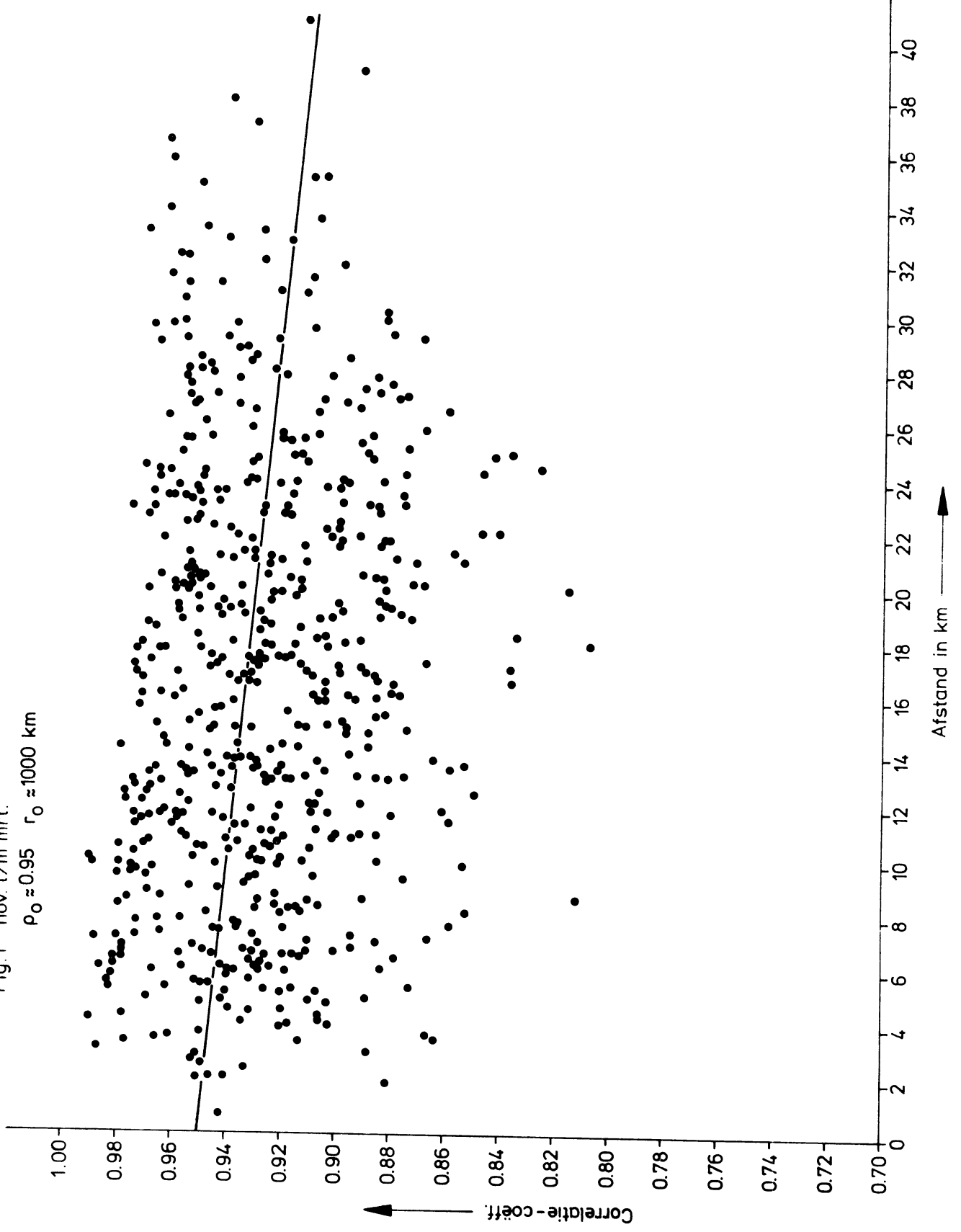


Fig.2 juni t/m aug.
 $\rho_0 \approx 0.99$ $r_0 \approx 190$ km

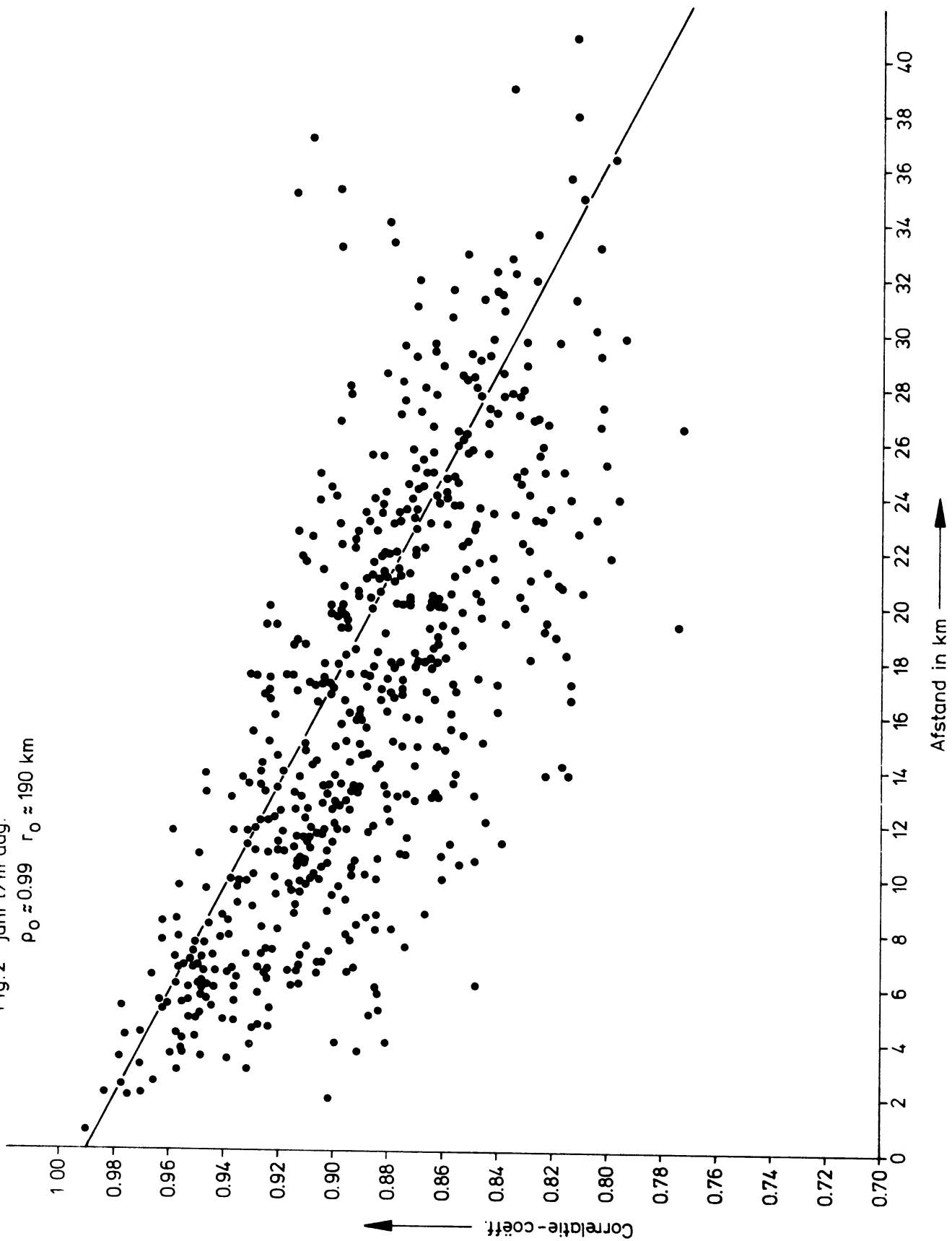


Fig.3 apr., mei, sept. en okt.
 $\rho_0 \approx 0.98$ $r_0 \approx 270$ km

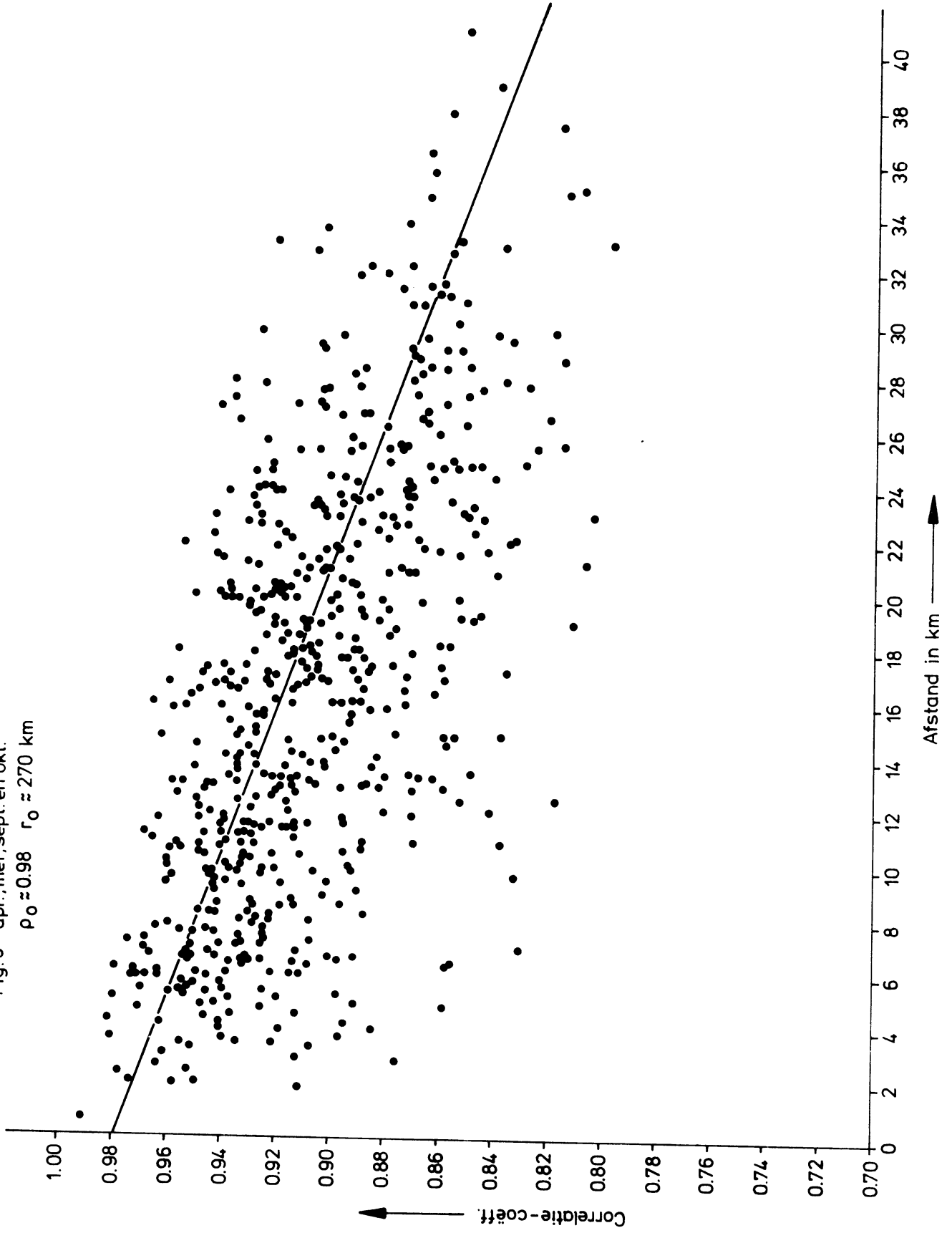


Fig 4 Winter maanden
15 Geselecteerde stations

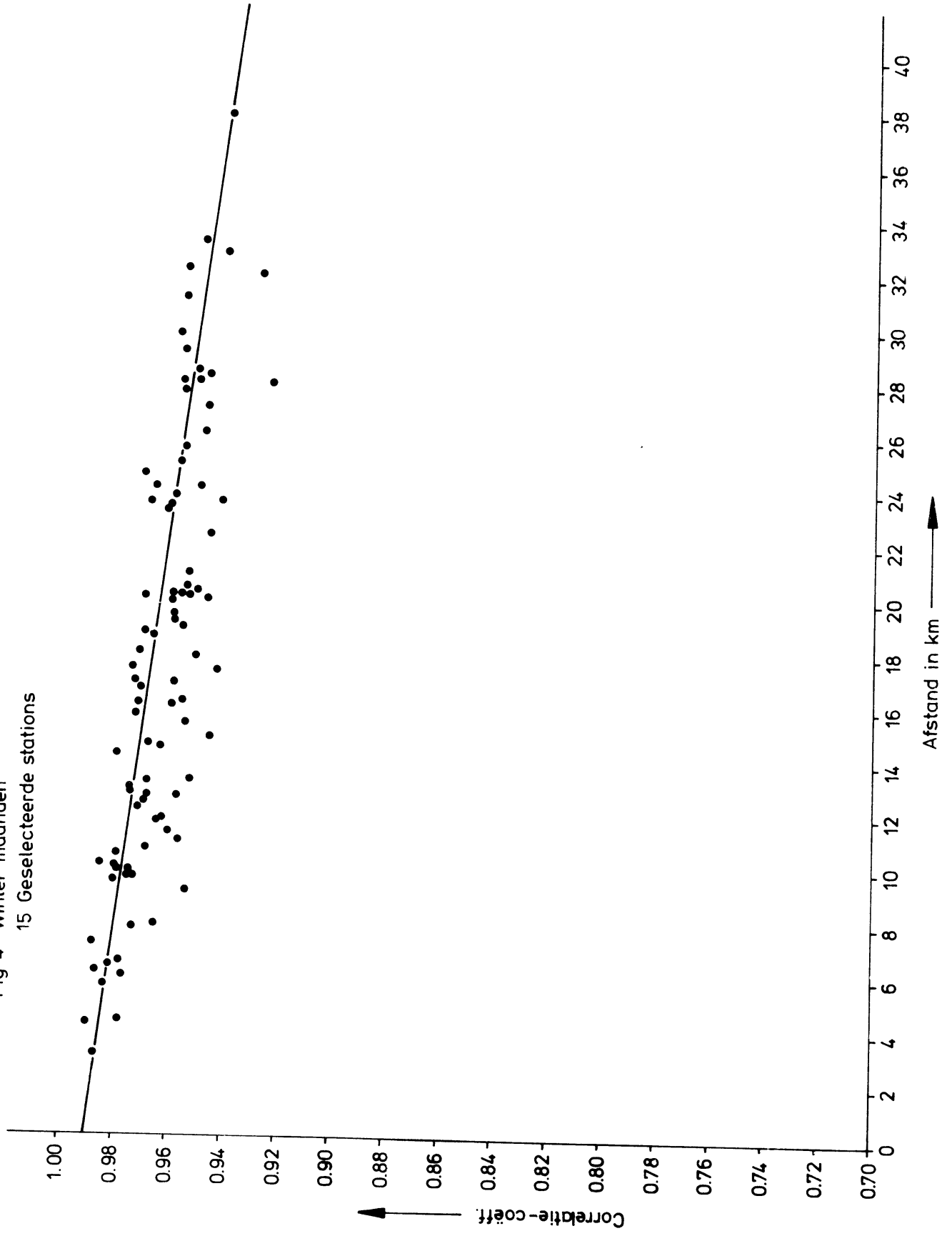


Fig. 5 Zomer maanden juni t/m aug.
Geselecteerde stations

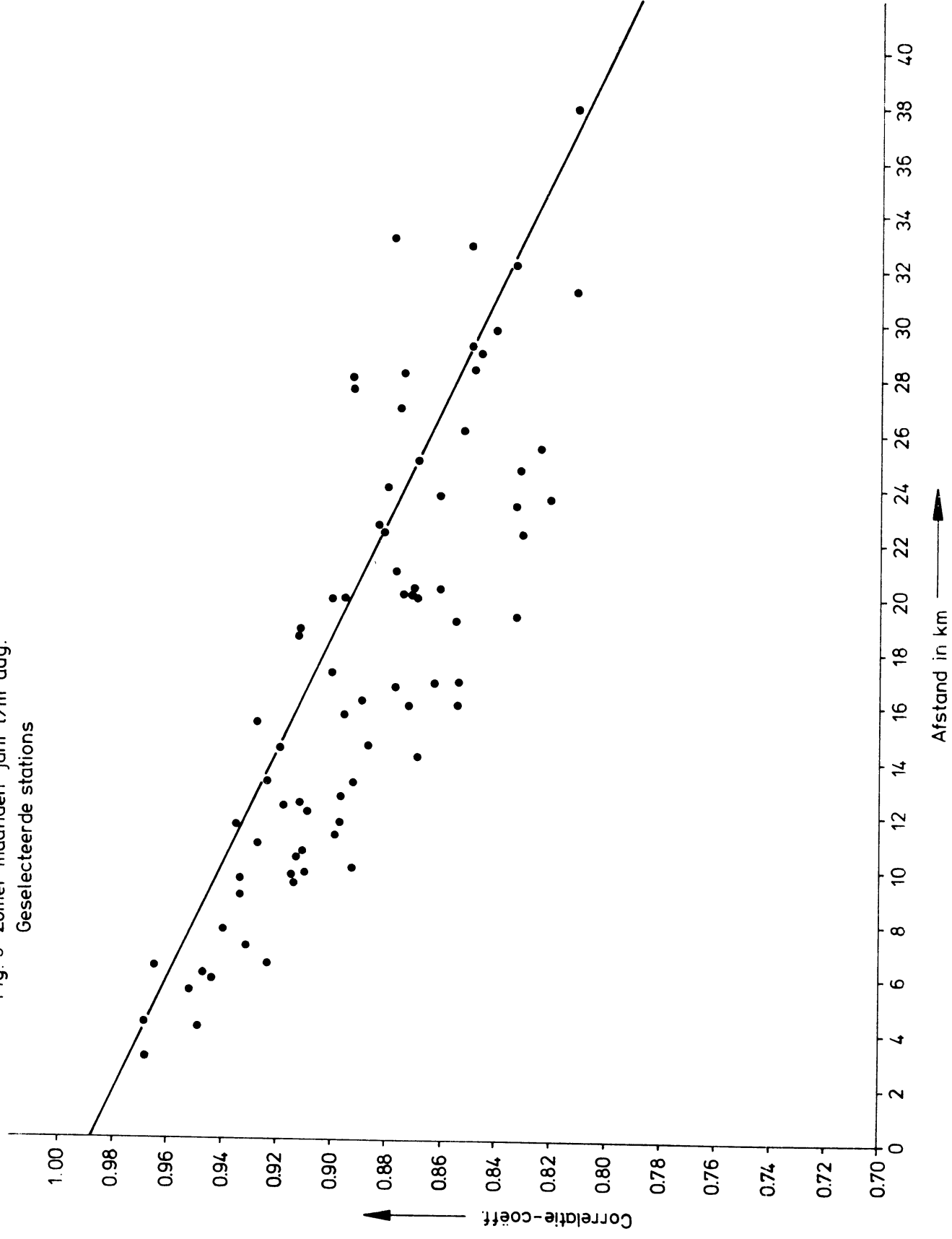


Fig.6 Voor- en najaar
Geselecteerde stations

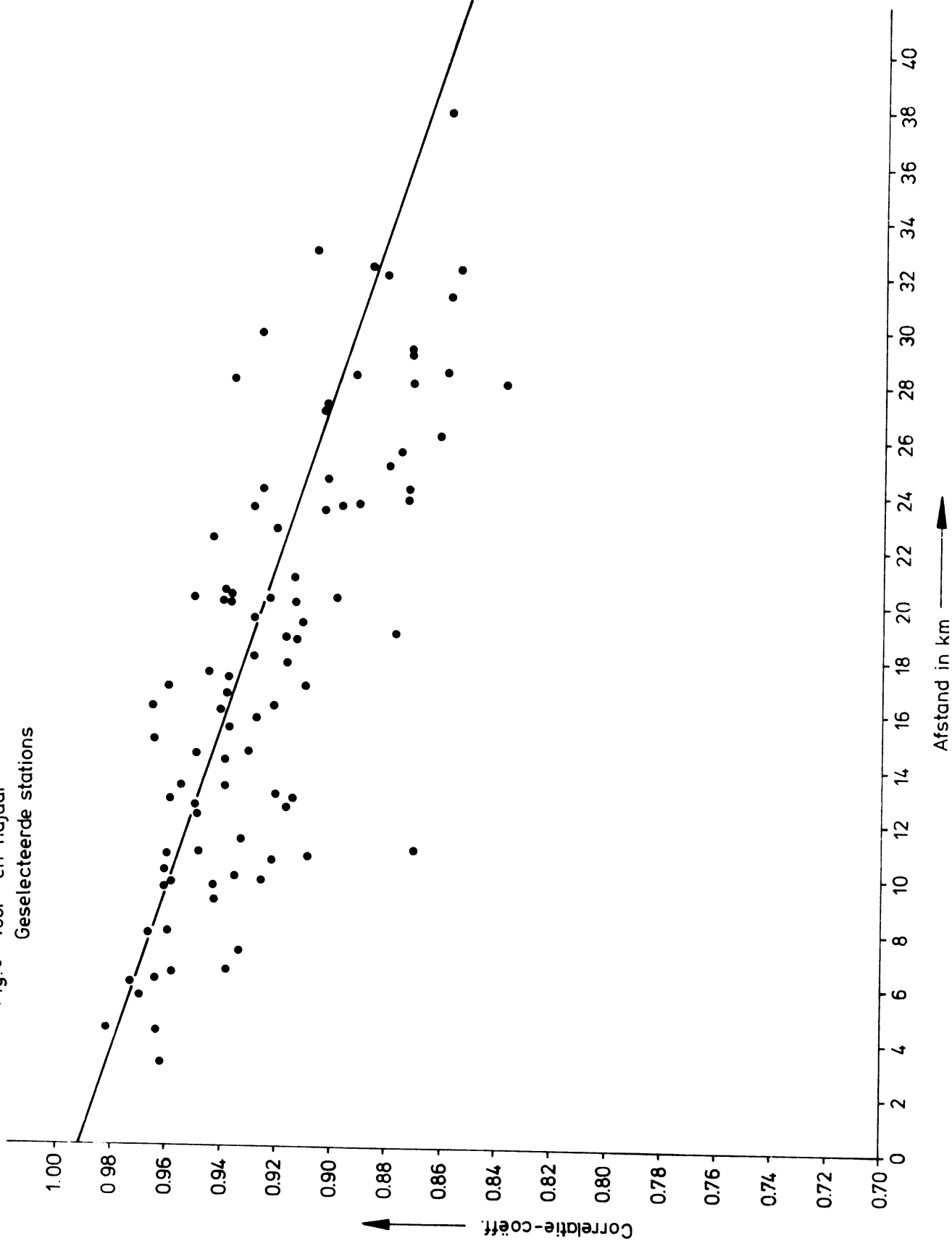


Fig. 7

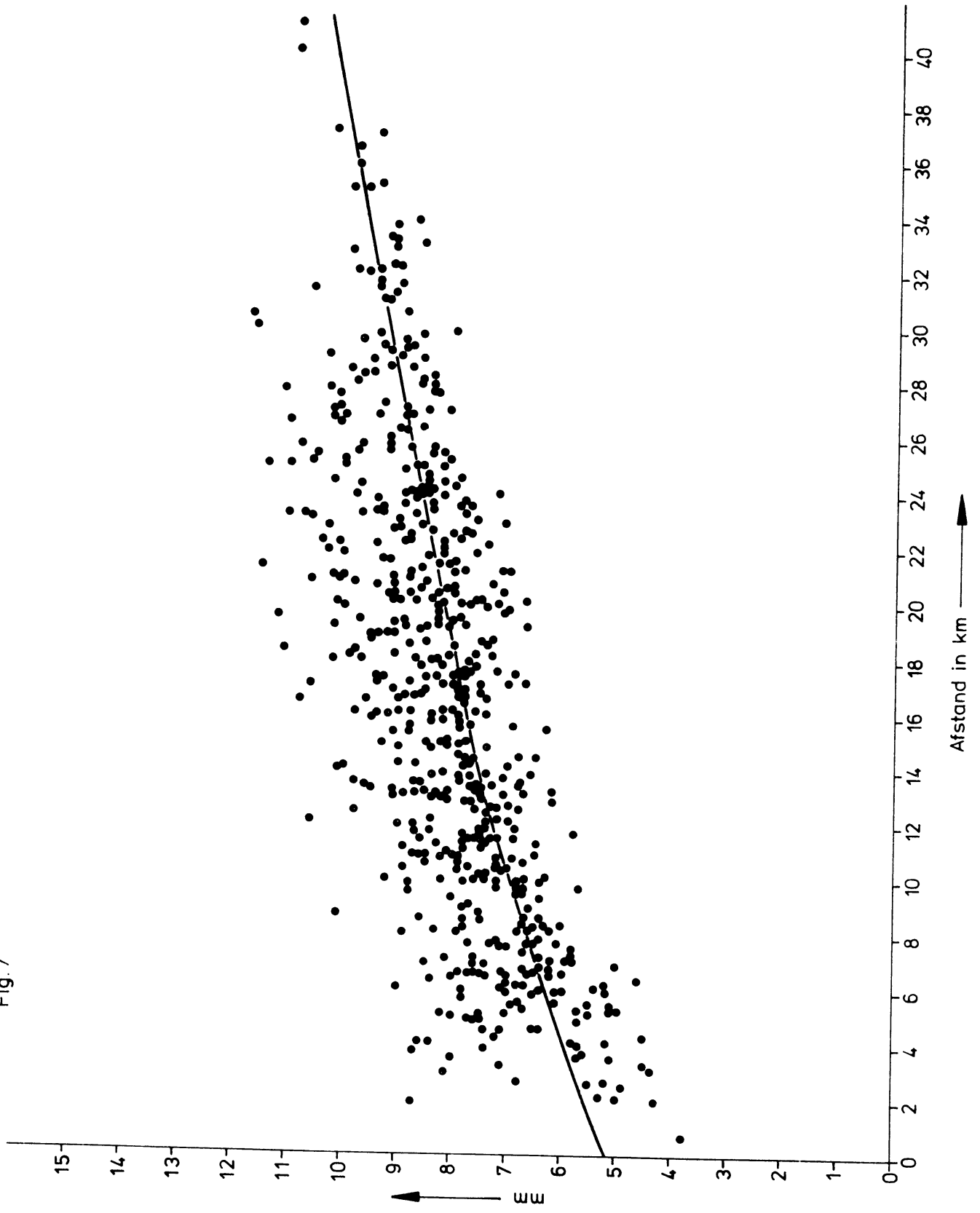
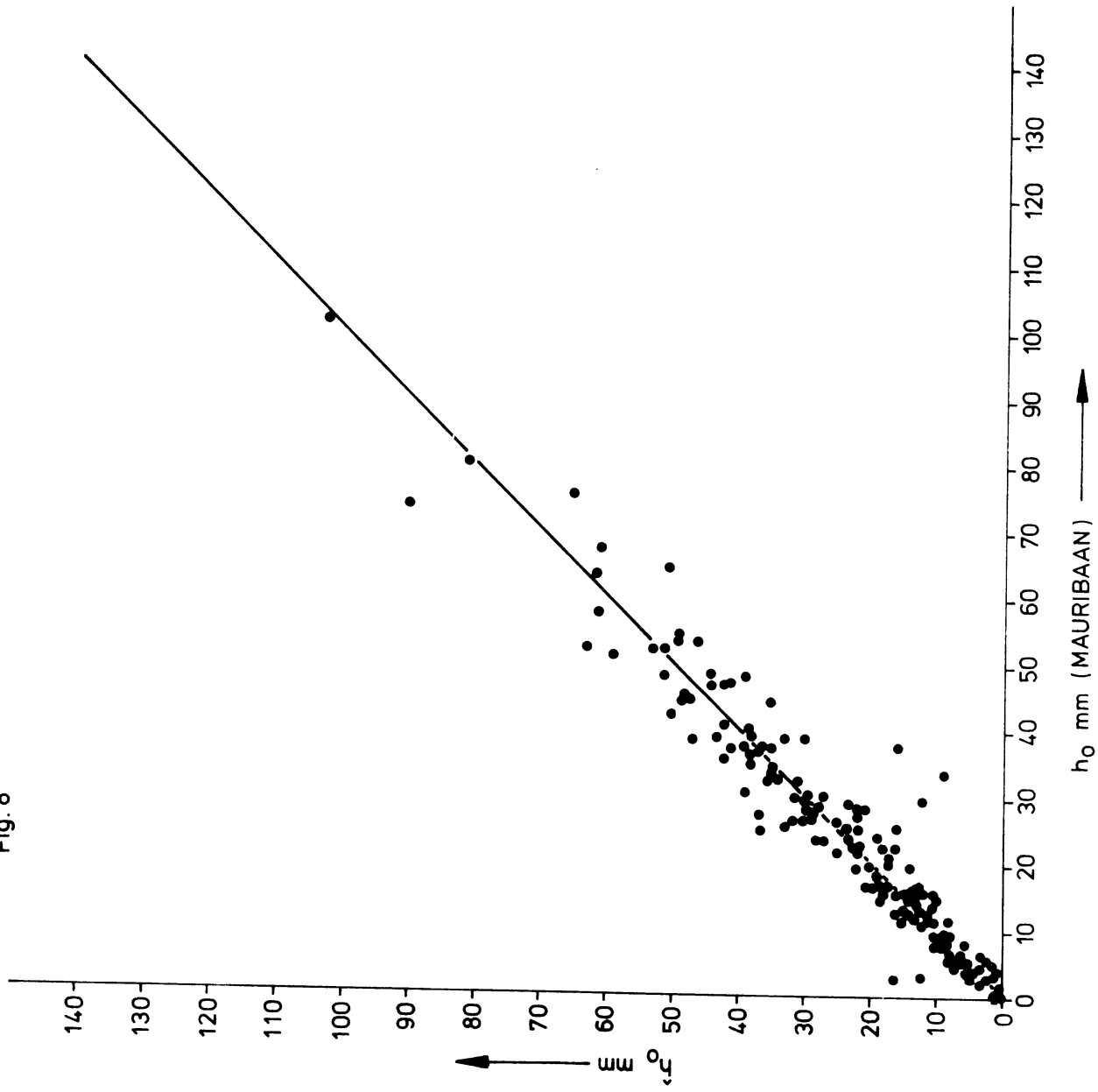


Fig. 8



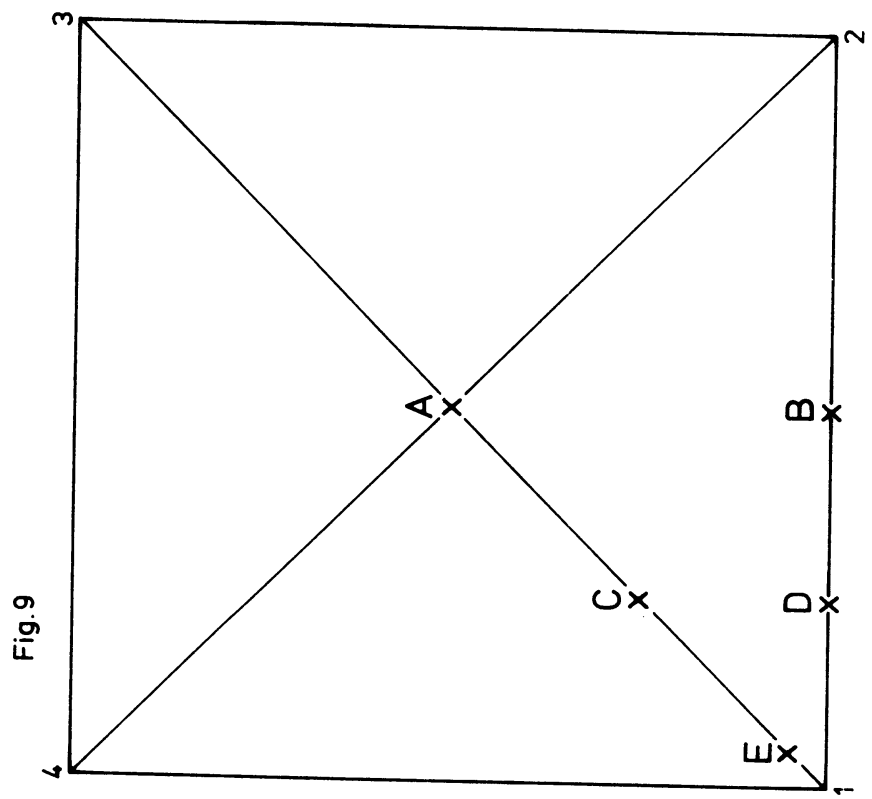


Fig. 9

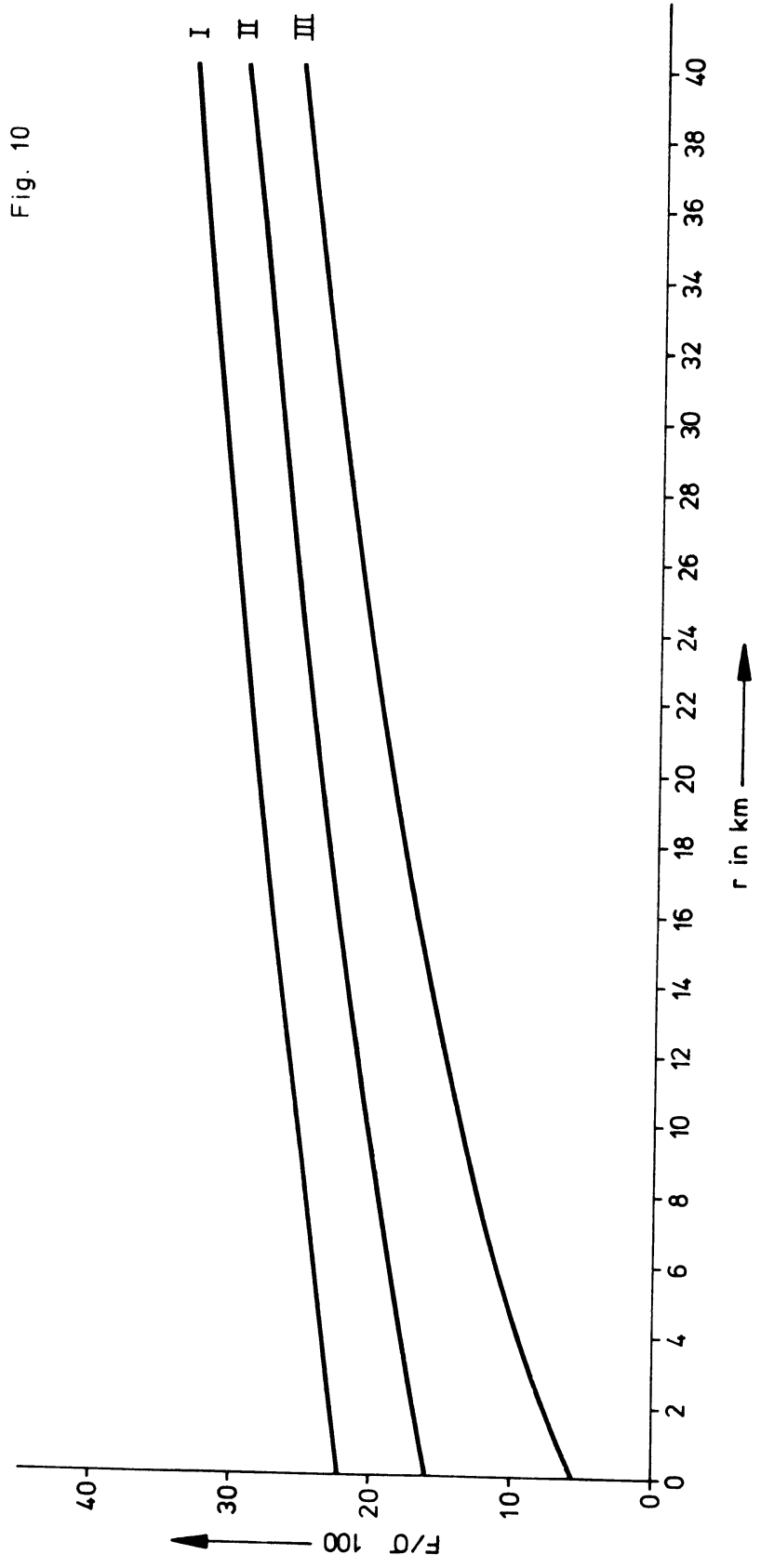


Fig. 10