HOMOGENITEITSONDERZOEK MET BETREKKING TOT FREQUENTIES VAN ZWARE DAGNEERSLAGSOMMEN

NOTA 55

J.V. Witter

LABORATORIUM VOOR HYDRAULICA EN
AFVOERHYDROLOGIE
LANDBOUWHOGESCHOOL
Februari 1982

162.168

INHOUD

	blz.
1. Inleiding	1
2. Trends in neerslagreeksen	3
2.1 Trends in gemiddeldenreeksen	4
2.2 Trends in jaarsgewijs gereduceerde stationsreeksen	6
2.3 Verdeling van trends in gereduceerde stationsreeksen over Nederland.	7
3. Neerslagverschillen tussen gebieden in Nederland.	12
4. De mogelijkheid van indeling van Nederland in gebieden.	15
5. Samenvatting en conclusies.	19
6. Literatuur.	20

FIGUREN

APPENDICES:

APPENDIX I: Lijst van beschouwde neerslagstations

APPENDIX II: M.A.J. van Montfort: "Enige statistische opmerkingen met betrekking tot homogeniteitsonderzoek"

1. Inleiding

Bij bestudering van afvoerprocessen is het onderkennen van trends in de te gebruiken neerslaggegevens van groot belang. Aangezien in eerdere onderzoekingen (KRAIJENHOFF VAN DE LEUR, PRAK; 1979) een toename van het jaarlijks aantal zomerdagen met minstens 30 mm neerslag werd geconstateerd voor een vijftal gebieden in Nederland, alsmede een afname voor enkele andere gebieden, werd nader onderzoek verricht naar het voorkomen van trends in de neerslag.

Dit nadere onderzoek, waarvan hier verslag wordt gedaan, richtte zich op de vraag:

is er sprake van trends in neerslagkenmerken gemeten op een aantal stations en, zo ja, zijn er verschillen in trend?

Aangezien bij een beoordeling van het practisch belang van trends onder andere het niveau waarop de trend zich voordoet, bekend moet zijn, richtte het onderzoek zich eveneens op de vraag:

zijn er verschillen in niveau van de kenmerken?

Tenslotte werd onderzocht:

leiden de geconstateerde verschillen tot een zinvolle indeling?

Voor het onderzoek werd beschikt over de gegevens van 140 K.N.M.I. neerslagstations, waarvan de dagneerslagen voor de periode 1951 t/m 1979 op magneetband (de "REGEN"-band) staan. Voorts waren van 18 stations de dagneerslagen voor de periode 1893 t/m 1979 op magneetband beschikbaar (voor de periode t/m 1953 op de "CODE2X"-band). Deze databestanden zijn in het hiernavolgende aangeduid met D140, resp. D18. Alhoewel voor beide periodes van meer stations de dagneerslagen beschikbaar waren, zijn de 140 stations geselecteerd op basis van een gering aantal ontbrekende waarnemingen en een gering aantal aanvankelijk meerdaagse sommen. Bij de selectie van de 18 stations gold eveneens het criterium van de ontbrekende waarnemingen, alsmede een indruk van de kwaliteit van de reeksen (BUISHAND, VELDS; 1980, pp. 137 t/m 144). De ligging van de stations voor beide databestanden D140 en D18 is aangegeven in figuur 1. In de appendix 1 staat van ieder station: de plaatsnaam, het K.N.M.I. codenummer en een aanduiding met betrekking tot ontbrekende waarnemingen.

Bestudeerd werden een aantal kenmerken met betrekking tot de hoeveelheid en de frequentie (aantal dagen in een jaargetijde met neerslagsom in een bepaald interval) van de neerslag. Als jaargetijden werden een zomerperiode (z: mei t/m september) en een winterperiode (w: oktober t/m april) beschouwd. De bestudeerde kenmerken zijn vermeld in tabel 1; hierbij duidt de hoofdletter

aan of het kenmerk betrekking heeft op een frequentie (F), dan wel op een hoeveelheid (N). De benedenindex duidt op het interval van de neerslagsom dat beschouwd werd. De bovenindex duidt op het beschouwde jaargetijde: z, w, j (met j: jaar).

Kenmerk	Beschouwd interval (mm)
F^z_{01}, F^w_{01}	$[1,\infty)$
F^z_{15}, F^w_{15}	$[15,\infty)$
F^z_{25}, F^w_{25}	$[25,\infty)$
N^j	n.v.t.

Tabel 1. Lijst der beschouwde neerslagkenmerken met hun notatie.

In de hierna volgende hoofdstukken 2, 3 en 4 zullen de hierboven genoemde vraagstellingen behandeld worden.

2. Trends in neerslagreeksen

Voor alle beschouwde kenmerken werden de uit beide databestanden ter beschikking staande reeksen op trend onderzocht. Daartoe werden de elementen van deze reeksen indien nodig getransformeerd en jaarsgewijs gereduceerd.

Transformatie was wenselijk bij de reeksen neerslagkenmerken met betrekking tot frequentie van de neerslag. Hierbij werden aantallen overschrijdingen geteld. Op deze aantallen overschrijdingen werd een variantie-stabiliserende transformatie, met als neveneffect bevordering van de normaliteit, toegepast, namelijk:

$$\tilde{n} = \sqrt{n} + \sqrt{n+1} \qquad (1)$$

met: $\tilde{n}$ = getransformeerd aantal

n = niet-getransformeerd aantal

Vervolgens werd voor elk jaar het gemiddelde bepaald van de getransformeerde aantallen voor alle stations. Ook voor de jaartotalen N^j werd een dergelijk gemiddelde bepaald. Er resulteerden 14 "gemiddeldenreeksen": voor beide databestanden D140 en D18 voor ieder van de in tabel 1 genoemde kenmerken één. Deze gemiddeldenreeksen werden gebruikt bij de jaarsgewijze reductie van de reeksen voor de individuele stations ("stationsreeksen").

Het trendonderzoek werd uitgevoerd met behulp van de toets op helling van de regressielijn, waarbij de nulhypothese van "geen trend" werd getoetst met behulp van de grootheid $\underline{T}$, met:

$$\underline{T} = \frac{\underline{r}\sqrt{m-2}}{\sqrt{1-\underline{r}^2}} \qquad (2)$$

met: m = lengte reeks

r = steekproefcorrelatiecoëfficiënt tussen de getallenparen (x_j, j)

met $j = 1, 2, \ldots, m$

en x_j = de waarde van het beschouwde kenmerk in het jaar j.

De steekproefcorrelatiecoëfficiënt r werd als volgt berekend:

$$r = \left[\sum_{j=1}^{m} (x_j-\bar{x}) * (j-\bar{j})\right] / \left[\sqrt{\Sigma\,(x_j-\bar{x})^2} * \sqrt{\Sigma\,(j-\bar{j})^2}\right] \qquad (3)$$

met: $\bar{x} = \frac{1}{m} \sum_{j=1}^{m} x_j$, $\bar{j} = \frac{1}{m} \sum_{j=1}^{m} j = (m+1)/2$

Onder de nulhypothese ("geen trend") geldt bij benadering:

$$\underline{T} \simeq \underline{t}_{m-2}$$

met: $\underline{t}_{m-2}$ = Student's variabele met m-2 vrijheidsgraden.

2.1. Trends in gemiddeldenreeksen

Een indruk van de mate waarin trend in de gemiddeldenreeksen aanwezig was, in het vervolg "landelijk opgetreden trend" genoemd, verschaft tabel 2.

Kenmerk	T voor gemiddeldenreeksen voor databestand			
	D140		D18	
	z	w	z	w
F_{01}	-1.29	0.84	0.75	2.12
F_{15}	-1.93	0.68	0.91	2.45
F_{25}	-2.03	-0.33	1.08	1.58
N	-0.55		2.74	

Tabel 2. Waarden van de toetsingsgrootheid $\underline{T}$ voor de toets op aanwezigheid van trend voor gemiddeldenreeksen met betrekking tot een aantal kenmerken (waarden van $\underline{T}$ die bij tweezijdige toetsing met $\alpha = 0.10$ in het kritieke gebied liggen, zijn onderstreept).

Geconcludeerd kan worden dat de lange reeksen een positieve, en de korte reeksen overwegend een negatieve trend vertonen. Daarnaast is er een verschil in landelijk opgetreden trends voor zomerreeksen en voor winterreeksen.

De positieve trend voor reeksen uit het databestand D18 is in de eerste plaats het gevolg van veranderingen in de neerslagmeting. Zo werden in de periode 1946 t/m 1950 de regenmeters vervangen door een type dat in plaats van 150 cm maar 40 cm boven het maaiveld opgesteld is. Na 1962 zijn deze weer vervangen door een type met dezelfde opstellingshoogte, maar met een opening van 200 cm^2 in plaats van 400 cm^2 (DEIJ; 1968). BUISHAND (1977) concludeert voor de periode na de eerste wijziging een toename van de gemeten neerslag in de orde van 2% voor ver van de kust verwijderde stations, tot iets meer dan 10% voor kuststations.

Dat de positieve trend bij de lange reeksen vooral aanwezig is bij in de winter gemeten neerslagkenmerken is aannemelijk aangezien trend in deze reeksen vooral veroorzaakt zal zijn door de veranderingen in opstellingshoogte van de regenmeters, waardoor de windfout in de metingen, dit is de verminderde neerslagvangst als gevolg van aerodynamische effecten, afneemt. Aangezien vooral in de winter regenmetingen behept zijn met een windfout, resulteren verschillen in trend voor in de zomer en in de winter gemeten neerslagkenmerken. Voor de seizoensverschillen in trend die optraden bij korte reeksen kon geen verklaring gevonden worden.

Landelijk opgetreden trends worden ook beïnvloed door de algemene circulatie in het beschouwde tijdvak. De frequentie van voorkomen van de verschillende circulatietypen is namelijk aan wisselingen onderhevig en er kunnen episoden onderscheiden worden, die gekenmerkt zijn door het overheersen van bepaalde typen van circulatiepatronen (BARRY, PERRY; 1973). De verschillende circulatietypen resulteren in weertypen met verschillende neerslagkansen. Zo leiden zonale circulatietypen in West Europa tot vochtige, milde winters en tot vochtige en koele zomers.

De opgetreden circulatietypen boven Nederland gedurende de periode 1881 t/m 1978 zijn bekend (HESS; 1977), evenals de neerslagkansen gegeven het opgetreden circulatietype voor een viertal K.N.M.I. stations (BIJVOET, SCHMIDT; 1960); daarmee kon de invloed van het circulatiepatroon op trends in de neerslag nagegaan worden door voor een bepaalde periode het op basis van de conditionele neerslagkansen berekende verwachte jaarlijks aantal dagen met een dagsom groter of gelijk aan een bepaalde drempelwaarde te vergelijken met het werkelijk opgetreden jaarlijks aantal dergelijke dagen. Als periode werd beschouwd het tijdvak 1956 t/m 1978, daar de genoemde neerslagkansen bepaald waren over het tijdvak 1881 t/m 1955. Als drempelwaarde werd 5 mm gekozen, in overeenstemming met de hierboven genoemde gepubliceerde conditionele neerslagkansen. De verwachte en werkelijk opgetreden jaarlijkse aantallen dagen met minstens 5 mm neerslag, zijn voor het station Den Helder/De Kooij in figuur 2$^{(a)}$ (voor de zomer) en in figuur 2$^{(b)}$ (voor de winter) afgebeeld.

Na een variantie-stabiliserende transformatie volgens vergelijking (1) op de verwachte en werkelijk opgetreden aantallen werd de steekproefcorrelatie-coefficiënt r tussen beide reeksen getransformeerde aantallen berekend volgens vergelijking (3), hetgeen resulteerde in:

$r_{zomer} = 0.64$

$r_{winter} = 0.55$

Beide waarden zijn significant bij toetsing met behulp van Student's t met een

eenzijdig alternatief en $\alpha = 0.05$, zodat er enige aanwijzing bestaat voor beïnvloeding van landelijk opgetreden trends door de algemene circulatie in het beschouwde tijdperk.

2.2. Trends in jaarsgewijs gereduceerde stationsreeksen

De stationsreeksen uit het databestand D140 voor kenmerken met betrekking tot aantallen overschrijdingen, getransformeerd volgens vergelijking (1), werden jaarsgewijs gereduceerd met behulp van de gemiddeldenreeksen. Ook de stationsreeksen met betrekking tot het kenmerk N^j werden jaarsgewijs gereduceerd, evenwel zonder voorafgaande transformatie. De zo verkregen stationsreeksen werden op de reeds beschreven wijze onderzocht op trend.

De jaarsgewijze reductie heeft een gunstige invloed op het onderscheidingsvermogen van de toets, aangezien de landelijke tussen-jaren variatie geelimineerd wordt. In plaats van dat beschouwd worden (eventueel eerst getransformeerde) getallen $x_{i,j}$ worden nu beschouwd getallen $z_{i,j}$ met:

$$z_{i,j} = x_{i,j} - x_{.j} \qquad (4)$$

met: $x_{i,j}$ = element van de beschouwde neerslagreeks voor station i in jaar j

$x_{.j} = \frac{1}{k} \sum_{i=1}^{k} x_{i,j}$

k = aantal beschouwde stations.

Vanwege de sterke, positieve gecorreleerdheid van $x_{i,j}$ en $x_{.j}$ zal de variantie van gereduceerde reekselementen $z_{i,j}$ kleiner zijn dan de variantie van niet-gereduceerde reekselementen $x_{i,j}$, waardoor bij trendonderzoek beschouwing van gereduceerde reeksen te verkiezen is boven beschouwing van niet-gereduceerde (zie eveneens BUISHAND; 1981).

In appendix 1 is voor beide databestanden voor alle beschouwde stations aangegeven, of trend in één van de beschouwde reeksen aangetoond werd (bij tweezijdige toetsing met $\alpha = 0.10$). Het blijkt dat gegeven het station en het databestand, trends in overeenkomstige reeksen, bijvoorbeeld zomerreeksen of winterreeksen, redelijk samenhangen.

De aantallen stations waarvoor bij een reeks met betrekking tot een bepaald neerslagkenmerk de nulhypothese verworpen moest worden, zijn aangegeven in tabel 3. Deze aantallen blijken groter te zijn dan de verwachte aantallen bij afwezigheid van correlatie, respectievelijk 14 en 1.8 voor de korte en de

lange periode.

Het relatief hoge aantal reeksen met trend beïnvloedde de steekproefvarianties van de berekende toetsingsgrootheden, waardoor ze afweken van de bij ongecorreleerde stationswaarnemingen te verwachten waarden: var($\underline{t}_{27}$) = 1.08 voor databestand D140, en var($\underline{t}_{85}$) = 1.02 voor D18.

De gevonden steekproefvarianties zijn opgenomen in tabel 3.

Kenmerk	D140				D18			
	n		s^2		n		s^2	
	z	w	z	w	z	w	z	w
F_{01}	21	32	1.59	2.39	4	11	3.65	8.56
F_{15}	23	24	1.78	1.49	5	6	1.68	3.36
F_{25}	20	19	1.29	1.28	1	3	0.97	1.13
N^j	48		2.74		11		5.45	

Tabel 3. Aantallen reeksen n waarvoor de nulhypothese ("geen trend") verworpen moest worden, alsmede steekproefvarianties s^2 van de berekende toetsingsgrootheden voor reeksen met betrekking tot 7 neerslagkenmerken.

2.3. Verdeling van trends in gereduceerde stationsreeksen over Nederland

Aangezien het opsporen van eventuele verschillen in trend tussen (groepen van) stations een deel van de onderzoeksvraag vormde, en voor de periode 1951 t/m 1979 voor 140 stations verspreid over Nederland trends ten opzichte van de landelijk gemiddelde trend met betrekking tot één of ander kenmerk berekend waren, werden deze op de stations berekende trends met behulp van een interpolatiemethode vereffend over Nederland. Hiertoe werd als indicatie voor de trend de in de stations berekende waarden van de toetsingsgrootheid $\underline{T}$ gekozen.

De toetsingsgrootheid $\underline{T}$ werd hierbij opgevat als een ruimtelijke variabele $\underline{T}(u)$ met:

u = plaatscoördinaat in het (x,y) vlak.

De waarde van $\underline{T}(u_o)$ met:

u_o = plaatscoördinaat van het punt waarnaar geïnterpoleerd wordt

kan zuiver en met minimale schattingsvariantie geschat worden met behulp van "kriging". Voor een gedetailleerde beschrijving van deze interpolatiemethode zij verwezen naar JOURNEL, HUIJBREGTS (1978) en naar GUARASCIO et al., ed. (1976).

Aan de schattingen met behulp van kriging liggen zwakke veronderstellingen met betrekking tot het gedrag van de ruimtelijke variabele ten grondslag, de zogenaamde "intrinsieke hypothese":

$$E\,\{\underline{T}\,(u+h) - \underline{T}\,(u)\} = 0 \qquad (5a)$$

$$\mathrm{Var}\,\{\underline{T}\,(u+h) - \underline{T}\,(u)\} = 2\gamma(h) \qquad (5b)$$

met: h = afstand

$\gamma(h)$ = het zogenaamde "semi-variogram", een functie van slechts de onderlinge afstand.

De eisen van zuiverheid en minimale schattingsvariantie leiden tot het volgende stelsel vergelijkingen:

$$\sum_{j=1}^{k} \lambda_j \gamma_{i,j} + \mu = \gamma_{i,o} \qquad (i = 1, \ldots, k) \qquad (6)$$

$$\sum_{j=1}^{k} \lambda_j = 1 \qquad (7)$$

met: k = aantal punten waarop de ruimtelijke variabele $\underline{T}(u)$ gedefinieerd is (k=140)

$\gamma_{i,j}$ = waarde van het semi-variogram voor de afstand tussen de punten i en j

$\gamma_{i,o}$ = waarde van het semi-variogram voor de afstand tussen de punten i en o

μ = Lagrange multiplicator

λ_j = gewicht in de interpolatiefunctie (8)

Nadat uit het stelsel vergelijkingen (6), (7) de gewichten λ_j bepaald zijn, volgt de waarde $\underline{T}(u_o)$ van de ruimtelijke variabele in het punt met plaatscoördinaat u_o uit de lineaire interpolatiefunctie:

$$\underline{T}(u_o) = \sum_{j=1}^{k} \lambda_j\, \underline{T}(u_j) \qquad (8)$$

Opgemerkt dient te worden dat het stelsel vergelijkingen (6), (7) slechts geldig is in het geval dat de ruimtelijke variabele geen "drift" vertoont.

Indien een drift aanwezig is geldt:

$$E\{\underline{T}(u)\} = f(u) \tag{9}$$

en worden aan het stelsel (6), (7) een aantal randvoorwaarden en Lagrange multiplicatoren toegevoegd. Tot aanwezigheid van een drift kan geconcludeerd worden uit het verloop van het semi-variogram $\gamma(h)$ met de afstand, indien dit meer dan kwadratisch toeneemt met de afstand.

Het semi-variogram $\gamma(h)$ is een grootheid, welke geschat wordt met behulp van het steekproef-variogram:

$$\hat{\gamma}(h^*) = \sum_{h^*-\delta<h_{ij}<h^*+\delta} (\underline{T}(u_i) - \underline{T}(u_j))^2/2n_h \tag{10}$$

met: h_{ij} = afstand van station i en station j

h^* = klassemidden van de te beschouwen afstandsklassen, waarover $\hat{\gamma}(h^*)$ bepaald wordt. In dit onderzoek werd beschouwd: $h^* = 5, 15, 25, \ldots$ (km)

δ = halve intervalbreedte van de te beschouwen afstandsklassen; hier gold: $\delta = 5$ (km)

n_h = aantal paren stations met onderlinge afstand tussen $h^*-\delta$ en $h^*+\delta$ (km).

Aan de met behulp van (10) berekende eerste 16 punten, dat wil zeggen voor onderlinge afstanden tussen punten kleiner dan 160 km, van het steekproefvariogram $\hat{\gamma}(h^*)$ werd vervolgens een semi-variogrammodel $\gamma(h)$ aangepast van de vorm:

$$\gamma(h) = \omega\ \{1 - \exp\ [(-1/a)h^b]\} \tag{11}$$

met ω, a, b: parameters (> 0).

Naast het hier toegepaste exponentiële model voor het semi-variogram volgens (11) bestaan er andere modellen voor het semi-variogram, bijvoorbeeld een lineair model in h en een logaritmisch model. De waarden van de parameters ω, a en b werden geschat met behulp van het kleinste kwadraten criterium (LS-schatters: Least Squares). Beginschattingen werden daarbij uit het steekproef-variogram afgelezen, aangezien bij een waarde van b gelijk 1 geldt, dat 3a ongeveer gelijk is aan het bereik van het semi-variogram, dit is de afstand h, waarna

het semi-variogram horizontaal blijft lopen. Bovendien geldt in dat geval dat ω gelijk is aan de drempel van het semi-variogram: de maximale waarde ervan.

De steekproef-variogrammen en de schattingen van de variogrammen volgens het exponentiële model zijn afgebeeld in de figuren $3^{(a)}$ t/m $3^{(g)}$. Enkele semi-variogrammen hebben een voortdurend stijgend verloop met de afstand, zij het minder dan kwadratisch. Voor deze kenmerken is de veronderstelling van tweede orde stationariteit klaarblijkelijk niet meer toegestaan, maar geldt nog wel de intrinsieke hypothese die voldoende is voor toepassing van de kriging procedure in deze vorm. In een drietal gevallen (figuren $3^{(c)}$, $3^{(d)}$ en $3^{(g)}$ werd voor afstanden $\geq$ 160 km een horizontaal verloop van het semi-variogram aangenomen, op grond van de waarden van het steekproef-variogram op grotere afstanden dan 160 km, waaraan het theoretisch verloop niet aangepast was. De LS-schattingen van ω, a en b voor de beschouwde neerslagkenmerken zijn vermeld in tabel 4. Geïnterpoleerd werd naar een regelmatig puntennet, met punten op afstanden van 7.9 km van elkaar, hetgeen de figuren $4^{(a)}$ t/m $4^{(g)}$ opleverde.

Kenmerk	ω		a		b	
	z	w	z	w	z	w
F_{01}	1.67	2.83	6.54	3.46	0.67	0.40
F_{15}	1.73	4.40	29.86	13.48	1.21	0.36
F_{25}	2.97	9.13	9.06	87.34	0.36	0.54
N^j	5.53		10.63		0.43	

Tabel 4. LS-schattingen van ω, a en b in uitdrukking (11) voor het semi-variogram voor 7 beschouwde neerslagkenmerken.

Voor trends in reeksen neerslagkenmerken met betrekking tot aantallen overschrijdingen is het globale beeld als volgt:

- bij de winterreeksen treden aan de kust positieve trends op, en in Oost-Nederland negatieve;
- bij de zomerreeksen is het beeld gecompliceerder, met positieve trends in de kop van Noord-Holland, in een strook Noord-Zuid over midden Nederland en in de Noord-Oost Polder, en met negatieve trends in Oost Drente, Oost Overijssel en Oost Gelderland, in Limburg en in een deel van Noord-Holland.

Naast het bovengenoemde globale beeld trekken de seizoensverschillen tussen de figuren voor de neerslagkenmerken F^z_{01} en F^w_{01}, met een lage drempelwaarde

van de beschouwde neerslagsommen, en de figuren voor de overige kenmerken, met een hoge drempelwaarde, de aandacht. Bij vergelijking van de figuren $4^{(a)}$ en $4^{(d)}$ voor de neerslagkenmerken F^{z}_{01} en F^{w}_{01} blijken seizoensverschillen in trend op te treden in Groningen, Noord-Holland, de Randstad en midden Nederland en Zeeland. Bij vergelijking van de figuren $4^{(b)}$ en $4^{(e)}$ voor de neerslagkenmerken F^{z}_{15} en F^{w}_{15} blijken deze seizoensverschillen in trend in Groningen en Zeeland grotendeels verdwenen te zijn, maar in Noord-Holland en in de Randstad en midden Nederland zijn zij blijven bestaan.

Voorts kunnen beschouwd worden verschillen tussen figuren voor neerslagkenmerken met een lage drempelwaarde van de beschouwde neerslagsom en kenmerken met een hoge drempelwaarde, gegeven het seizoen. Bij vergelijking van de figuren $4^{(a)}$ en $4^{(b)}$ voor de neerslagkenmerken F^{z}_{01} en F^{z}_{15} blijken dan verschillen in trend op te treden in Groningen, Friesland, Noord-Holland, Midden Nederland en Zeeland. Bij vergelijking van de figuren $4^{(d)}$ en $4^{(e)}$ voor de neerslagkenmerken F^{w}_{01} en F^{w}_{15} blijken de verschillen in trend in Groningen en Noord-Holland grotendeels verdwenen.

Interpretatie van de figuren wordt overigens bemoeilijkt door de volgende drie overwegingen:

a. de figuren voor één seizoen zijn niet onafhankelijk van elkaar;
b. de 3 verschillende drempels zijn bedoeld om onderscheid te maken tussen alle neerslagen ($\geq$ 1 mm), zware neerslagen ($\geq$ 15 mm) en zware convectieve neerslagen ($\geq$ 25 mm in het zomerseizoen). Het is duidelijk dat de neerslaghoeveelheid voor het onderscheiden tussen neerslagtypen een grof middel is;
c. locale effecten (land-zee effecten, topografie, differentiële opwarming, stadseffecten) zijn, zeker in kwantitatieve zin, moeilijk te verdisconteren.

3. Neerslagverschillen tussen gebieden in Nederland

Bij een beoordeling van het practisch belang van trends is tevens van belang het niveau waarop de trends zich voordoen. In verband hiermee is voor de 7 kenmerken onderzocht in hoeverre er verschillen tussen gebieden optreden. Hiertoe werden de gemiddelde waarden over de periode 1951 t/m 1979 van aantallen met betrekking tot elk van de 7 beschouwde kenmerken voor alle 140 stations ("stationsgemiddelden") bepaald. Gemiddelden daarvan, alsmede de steekproefvarianties en variatiecoëfficiënten zijn vermeld in tabel 5. Histogrammen van de verdeling van de stationsgemiddelden voor de kenmerken F^z_{01}, F^w_{01}, F^z_{15} en F^w_{15} zijn in de figuren $5^{(a)}$ t/m $5^{(d)}$ opgenomen.

Kenmerk	$\bar{x}$		s^2		vc (%)	
	z	w	z	w	z	w
F_{01}	52.38	80.52	$(2.58)^2$	$(2.57)^2$	4.9	3.2
F_{15}	4.69	3.50	$(0.53)^2$	$(0.58)^2$	11.3	16.4
F_{25}	1.15	0.48	$(0.24)^2$	$(0.14)^2$	21	29
N^j	775.80 (mm)		$(36.4)^2$ (mm^2)		4.7	

Tabel 5. Steekproefgemiddelden $\bar{x}$, steekproefvarianties s^2 en variatiecoëfficiënten vc voor stationsgemiddelden betreffende een zevental kenmerken.

Ten behoeve van een optimale schatting van de neerslagverdeling over Nederland werden de op de 140 stations bepaalde stationsgemiddelden geïnterpoleerd naar een dicht en regelmatig puntennet, geheel volgens de in hoofdstuk 2 beschreven methode. De steekproef-variogrammen en de schattingen van de variogrammen volgens het exponentiële model zijn afgebeeld in de figuren $6^{(a)}$ t/m $6^{(g)}$. Ook nu blijken de semi-variogrammen voor enkele neerslagkenmerken een voortdurend stijgend verloop met de afstand te hebben. LS-schattingen van de parameters ω, a en b voor de beschouwde neerslagkenmerken zijn vermeld in tabel 6. Interpolatie leverde de figuren $7^{(a)}$ t/m $7^{(g)}$ op.

Kenmerk	ω		a		b	
	z	w	z	w	z	w
F_{01}	9.26	25.62	429.92	23.82	1.34	0.42
F_{15}	0.30	0.35	35.72	41.62	1.13	1.18
F_{25}	0.066	0.034	19.68	8.53	1.08	0.38
N^j	1390.80		38.10		1.12	

Tabel 6. LS-schattingen van de parameters ω, a en b in uitdrukking (11) voor het semi-variogram, voor de 7 beschouwde neerslagkenmerken.

Voor het kenmerk N^j, de jaarneerslag, die nu per station over de beschouwde periode gemiddeld werd, resulteerde figuur $7^{(g)}$, welke goed overeenstemt met de in BUISHAND, VELDS (1980) gegeven beschrijving van de verdeling van dit neerslagkenmerk over Nederland:

"De natste plaatsen in Nederland zijn de Veluwe en de omgeving van Vaals. Betrekkelijk nat zijn verder het Drentse plateau en het oosten van Friesland, de heuvels van Overijssel, de Utrechtse heuvelrug en gebieden achter de Hollandse duinen. De droogste delen van Nederland zijn de Groningse waddenkust, het oosten van Brabant en het noorden en midden van Limburg" (p. 78).

De in KRAIJENHOFF VAN DE LEUR, PRAK (1979) geconstateerde Zuid West-Noord Oost georiënteerde strook over Nederland met hoge frequenties van zware zomerse neerslagen, is terug te vinden in de figuren $7^{(b)}$ en $7^{(c)}$.

Voor de verdeling over Nederland naar seizoen van de zwaardere buien is vergelijking van de figuren $7^{(b)}$ en $7^{(e)}$ voor de neerslagkenmerken F^z_{15}, resp. F^w_{15} interessant. Voor de volgende gebieden blijken bij het beschouwde neerslagkenmerk seizoensverschillen op te treden:

- het gebied Rotterdam-Dordrecht
- de Noord-Oost Polder
- de kop van Noord-Holland
- Noord-Oost Groningen.

Bij vergelijking van de figuren $7^{(a)}$ en $7^{(b)}$ voor de neerslagkenmerken F^z_{01}, resp. F^z_{15} blijken verschillen op te treden in de volgende gebieden:

- Randstad en Noord-West Brabant
- het oosten van de provincies Groningen en Drente.

De verschillen tussen de figuren $7^{(a)}$ en $7^{(b)}$ zijn aanmerkelijk groter dan de verschillen tussen de daarmee corresponderende, maar nu voor de winterperiode bepaalde figuren $7^{(d)}$ en $7^{(e)}$.

Een indruk van de mogelijke verschillen in jaarlijkse frequentie van zware neerslagen tussen twee willekeurige plaatsen in Nederland, verschaffen de histogrammen voor de kenmerken F^{z}_{15} en F^{w}_{15} in de figuren $5^{(c)}$ en $5^{(d)}$.

Een indruk van de consequenties van neerslagverschillen tussen 2 plaatsen werd verkregen door ter beschikking staande uurneerslaggegevens over de periode 1957 t/m 1977 van de stations Den Helder/De Kooij en De Bilt als invoer te gebruiken in rioleringsberekeningen voor een woonwijk in Enschede. Hierbij werd gebruik gemaakt van het rioleringsmodel STORM (HEC; 1977) met de voor de Enschedese woonwijk bepaalde parameterwaarden (LEENEN, GROOT; 1980).

Gebruik van de neerslaggegevens van De Bilt en van Den Helder/De Kooij leidde tot aanzienlijke verschillen in het berekende gemiddeld jaarlijks aantal riooloverstortingen $\bar{n}_{ov}$, namelijk voor De Bilt: $\bar{n}_{ov} = 5.90$ (met geschatte standaardafwijking $s = 2.83$) en voor Den Helder/De Kooij: $\bar{n}_{ov} = 4.43$ (met geschatte standaardafwijking $s = 2.82$).

4. De mogelijkheid van indeling van Nederland in gebieden

De in de vorige hoofdstukken geconstateerde plaatselijke verschillen in trend en niveau van een aantal kenmerken zijn getoetst op statistische significantie. Onderzocht werd of sterkere locale clustering optrad dan reeds op grond van de positieve correlatie tussen waarnemingen op verschillende stations verwacht mocht worden, aangezien zelfs toevalstrekkingen bij onderlinge gecorreleerdheid te interpreteren figuren opleveren. Als voorbeeld moge dienen de figuren $8^{(a)}$ t/m $8^{(d)}$ die tot stand kwamen door viermaal aan elk van de 140 stations een trekking uit de standaard-normale verdeling toe te wijzen, waaraan met behulp van de Choleski-procedure (WILKINSON; 1965) een correlatiestructuur opgelegd werd, beschreven door het model:

$$\rho(h_{i,j}) = \exp\,(-\lambda_1 . h_{i,j}^{\lambda_2}) \qquad (12)$$

met $h_{i,j}$ = onderlinge afstand van de stations i en j

ρ = correlatiecoëfficiënt

λ_1, λ_2 = parameters ($\lambda_1 = 0.07$, $\lambda_2 = 0.80$)

Bij de statistische toetsing werd allereerst de algemene vraag gesteld:

a. *zijn alle verwachtingen (van één of ander kenmerk, gemeten op 140 stations) gelijk?*

Vervolgens werden twee meer specifieke vragen gesteld, waarbij indelingen in gebieden getoetst werden:

b. *zijn bij indeling in een aantal gebieden deze gebieden intern homogeen?*

c. *zijn er verschillen tussen intern homogene gebieden?*

Ter beantwoording van de vragen b. en c. werd een viertal gebiedsindelingen opgesteld:

1. een indeling in 15 deelgebieden, corresponderend met de indeling in "deelgebieden" van het K.N.M.I.
2. een indeling in 5 deelgebieden, ruwweg corresponderend met een indeling in kust-land en Noord-Zuid (zie fig. 9).
3. een indeling in een "stedelijk" gebied, bestaand uit een aantal stations in de omgeving van Rotterdam en uit een aantal stations in Zeeuws Vlaanderen en Zuid Limburg, en een "landelijk" gebied, bestaand uit overig Nederland (zie fig. 10). Deze indeling is geïnspireerd op de door KRAIJENHOFF VAN DE LEUR, PRAK (1979) gevonden verstedelijkte

gebieden, gekenmerkt door hoge jaarlijkse frequenties van zware zomerse neerslagen.

4. een indeling in gebieden gekenmerkt door een hoge, een lage en een middelmatige jaarneerslag (gemiddeld over de periode 1951 t/m 1979) (zie fig. 11).

In appendix 1 staat aangegeven in welke deelgebieden de verschillende stations liggen bij verschillende gebiedsindelingen.

Beschouwd werden niveaus voor de kenmerken N^j, F^z_{15}, F^w_{15} en F^z_{25}, die een indicatie geven van de verdeling over Nederland van zware, uit hydrologisch oogpunt interessante neerslagen, alsmede trends voor de kenmerken F^z_{25} en F^z_{15}, omdat verstedelijking en industrialisatie zou resulteren in plaatselijke neerslagverschillen voor zware zomerse buien.

Het beschouwde statistisch model is beschreven in VAN MONTFORT (1981), en is als appendix 2 opgenomen. Bij bovengenoemde vraagstelling a. werd de nulhypothese van "alle verwachtingen zijn onderling gelijk" getoetst met de in appendix 2 vermelde toetsingsgrootheid $\underline{T}_2$. De bij vraagstelling b. behorende nulhypothese van "intern homogene gebieden" werd getoetst met de daar vermelde toetsingsgrootheid $\underline{T}_4$, en de nulhypothese van "geen verschillen tussen gebieden", behorend bij vraagstelling c. werd getoetst met behulp van $\underline{T}_5$.

Vanwege het gebruikte correlatie-afstand model volgens vergelijking (12) en het gebruikte model voor het semi-variogram volgens vergelijking (11) geldt voor de coëfficiënten λ_1 en λ_2 in de correlatie-afstand relatie:

$$\lambda_1 = 1/a \tag{13}$$

$$\lambda_2 = b \tag{14}$$

De resulterende waarden voor λ_1 en λ_2 voor de beschouwde kenmerken zijn vermeld in tabel 7.

Aangezien bij de toetsing beschouwd werd een m-dimensionaal Normaal verdeelde waarnemingsvector $\underline{z}$ (m = 140, het aantal stations) met verwachtingsvector $\xi = 0$, werden op de gegevens nog enige transformaties toegepast.

Bij beschouwing van trends in de kenmerken F^z_{15} en F^z_{25} werden onderzocht de genormeerde $\underline{T}$-waarden voor station i in de toets voor het betreffende kenmerk op helling van de regressielijn:

$$\underline{z}_i = \underline{T}_i/s \tag{15}$$

waarbij de standaardafwijking s werd bepaald uit het semi-variogram.

	Kenmerk	λ_1	λ_2
Trends	F^z_{15}	0.034	1.21
	F^z_{25}	0.070	0.80
Niveaus	F^z_{15}	0.028	1.13
	F^w_{15}	0.024	1.18
	F^z_{25}	0.050	1.08
	N^j	0.026	1.12

Tabel 7. Waarden van de parameters λ_1 en λ_2 in de correlatie-afstand relatie volgens vergelijking (12) voor trends en niveaus voor enige kenmerken.

Bij beschouwing van niveaus voor de kenmerken F^z_{15}, F^w_{15} en F^z_{25} werd onderzocht het volgens vergelijking (1) getransformeerde en vervolgens gestandaardiseerde (door aftrekking van het gemiddelde en deling door de uit het semi-variogram afgelezen standaardafwijking) gemiddelde jaarlijkse aantal neerslagen in de beschouwde neerslagklasse. Deze zelfde procedure (transformeren, gevolgd door standaardiseren) werd toegepast bij beschouwing van jaartotalen neerslag, het kenmerk N^j. De resultaten van de verschillende toetsingen zijn opgenomen in tabel 8.

Opgemerkt kan allereerst worden dat voor het kenmerk N^j, dus de gemiddelde jaarlijkse neerslagsom, meer evidentie gevonden werd voor een heterogene verdeling over Nederland, dan voor kenmerken met betrekking tot jaarlijkse aantallen zware zomerse neerslagen F^z_{15} en F^z_{25}. Voorts kan worden opgemerkt dat de indelingen 1 t/m 3 niet erg bevredigend zijn, daar de aanname van interne homogeniteit van de deelgebieden in meer dan de helft van de gevallen verworpen moet worden, en er slechts in 1 geval evidentie is voor verschillen tussen gebieden (overigens bij intern inhomogene deelgebieden). Indeling 4 die ontworpen is op basis van de ruimtelijke verdeling van de jaarneerslag, werkt ook goed onderscheidend bij de overige hier beschouwde kenmerken met betrekking tot het niveau van de neerslag.

	Kenmerk	$\underline{T}_2$	indeling 1		indeling 2		indeling 3		indeling 4	
			$\underline{T}_4$	$\underline{T}_5$	$\underline{T}_4$	$\underline{T}_5$	$\underline{T}_4$	$\underline{T}_5$	$\underline{T}_4$	$\underline{T}_5$
Trends	F^z_{15}									
	F^z_{25}	0	0		0		0		0	
Niveaus	N^j	*	*		*		*			*
	F^z_{15}	0	0		0					*
	F^w_{15}	0			0		0			*
	F^z_{25}							*		*

Tabel 8. Resultaten van een drietal toetsen betreffende clustering voor 6 verschillende kenmerken bij 4 gebiedsindelingen. De legenda voor het kritieke niveau (P) van de toets: $P \leq 0.05$: *

$0.05 < P \leq 0.10$: 0

5. Samenvatting en conclusies

1. De jaren 1951 t/m 1979 worden gekenmerkt door een lichte, significante afname van het jaarlijks aantal zware, zomerse dagneerslagen (hfdst. 2).
2. Met betrekking tot deze afname voor dagneerslagen zwaarder dan 25 mm mag Nederland niet-homogeen geacht worden (hfdst. 4).
3. Het met behulp van een optimale lineaire interpolatiemethode vereffende patroon van trends in neerslagreeksen met betrekking tot de beschouwde kenmerken is niet eenvoudig fysisch verklaarbaar. Enige voorgestelde gebiedsindelingen bleken niet in overeenstemming te brengen met dit patroon (hfdst. 2 en 4).
4. Met betrekking tot aantallen zware dagneerslagen voor zowel het winter- als het zomerseizoen, als met betrekking tot de gemiddelde jaarneerslag mag Nederland niet-homogeen geacht worden (hfdst. 3 en 4). Eén der voorgestelde gebiedsindelingen bleek overeen te stemmen met de na vereffening resulterende ruimtelijke patronen voor de 4 in dit verband beschouwde kenmerken (hfdst. 4).

6. Literatuur

BARRY, A.G., PERRY, A.H. (1973): "Synoptic climatology: methods and applications", Methuen and Co., London.

BUISHAND, T.A. (1977): "Stochastic modelling of daily rainfall sequences", Landbouwhogeschool, Wageningen.

BUISHAND, T.A. (1981): "The analysis of homogenity of long-term rainfall records in the Netherlands", Wetenschappelijk Rapport K.N.M.I. 81-7, K.N.M.I., De Bilt.

BUISHAND, T.A., VELDS, C.A. (1980): "Neerslag en verdamping", K.N.M.I., De Bilt.

BIJVOET, H.C., SCHMIDT, F.H. (1960): "Het weer in Nederland in afhankelijkheid van circulatietypen, deel II", Wetenschappelijk Rapport K.N.M.I. 60-1, K.N.M.I., De Bilt.

DEIJ, L.J.L. (1968): "Enige opmerkingen over de neerslag en neerslagverdeling naar tijd en plaats in ons land", in: Verslagen en Mededelingen T.N.O. No. 14, T.N.O., Den Haag.

HEC (1977): "Storage, treatment, overflow, runoff model "Storm", Users manual", Hydrologic Engineering Center, U.S. Army Corps of Engineers, Davis, California.

HESS, P. (1977): "Katalog der Grosswetterlagen Europas", Deutsche Wetterdienst, Offenbach am Main.

JOURNEL, A.G., HUIJBREGTS, Ch.J. (1978): "Mining Geostatics", Academic Press, London.

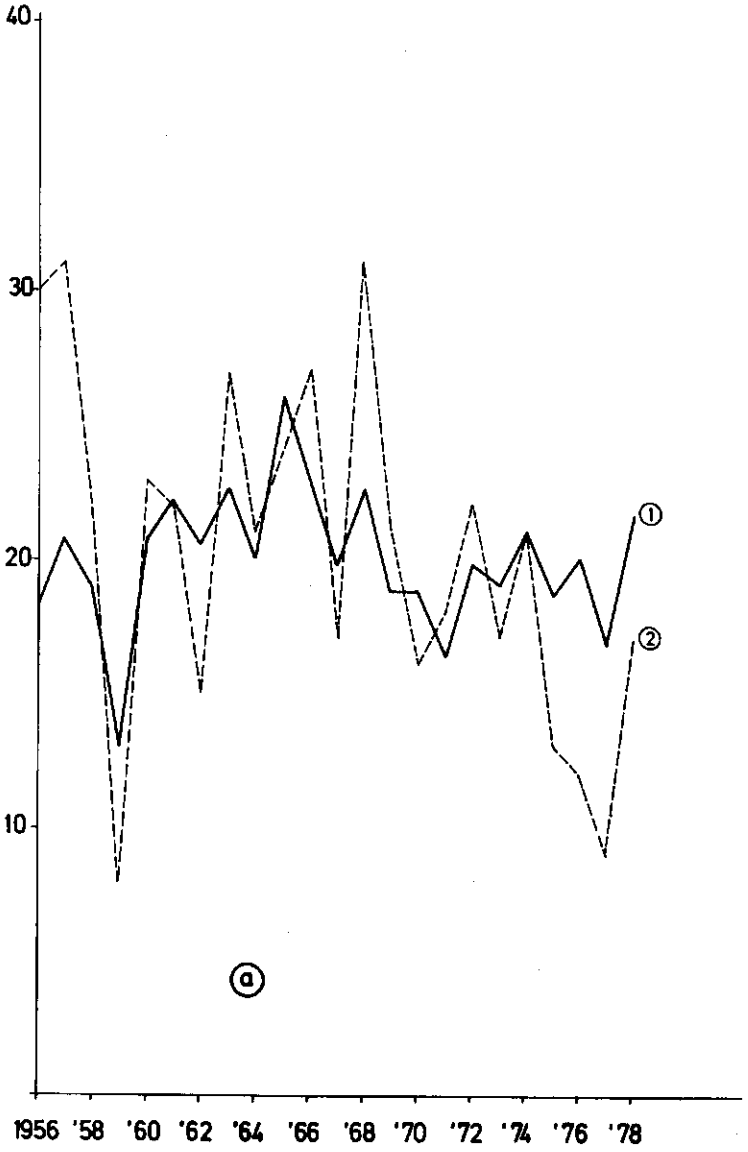
KRAIJENHOFF VAN DE LEUR, D.A., PRAK, H. (1979): "Verstedelijking, industrie en zware zomerregens: een verkennende studie", H_2O 12, pp. 75-81.

LEENEN. J.D., GROOT, S. (1980): "Kwalitatieve en kwantitatieve aspecten van de afvoer van stedelijke gebieden", Nota 45, Vakgroep Hydraulica en Afvoerhydrologie, Landbouwhogeschool, Wageningen.

VAN MONTFORT, M.A.J. (1981): "Enige statistische opmerkingen m.b.t. homogeniteitsonderzoek", Statistisch advies MM-81-06, Vakgroep Wiskunde, Landbouwhogeschool, Wageningen.

WILKINSON, J.H. (1965): "The algebraic eigenvalue problem", Clarendon Press, Oxford.

Fig. 1. Ligging van 140 en 18 geselecteerde neerslagstations voor databestanden D140 en D18, aangegeven met resp. • en ○



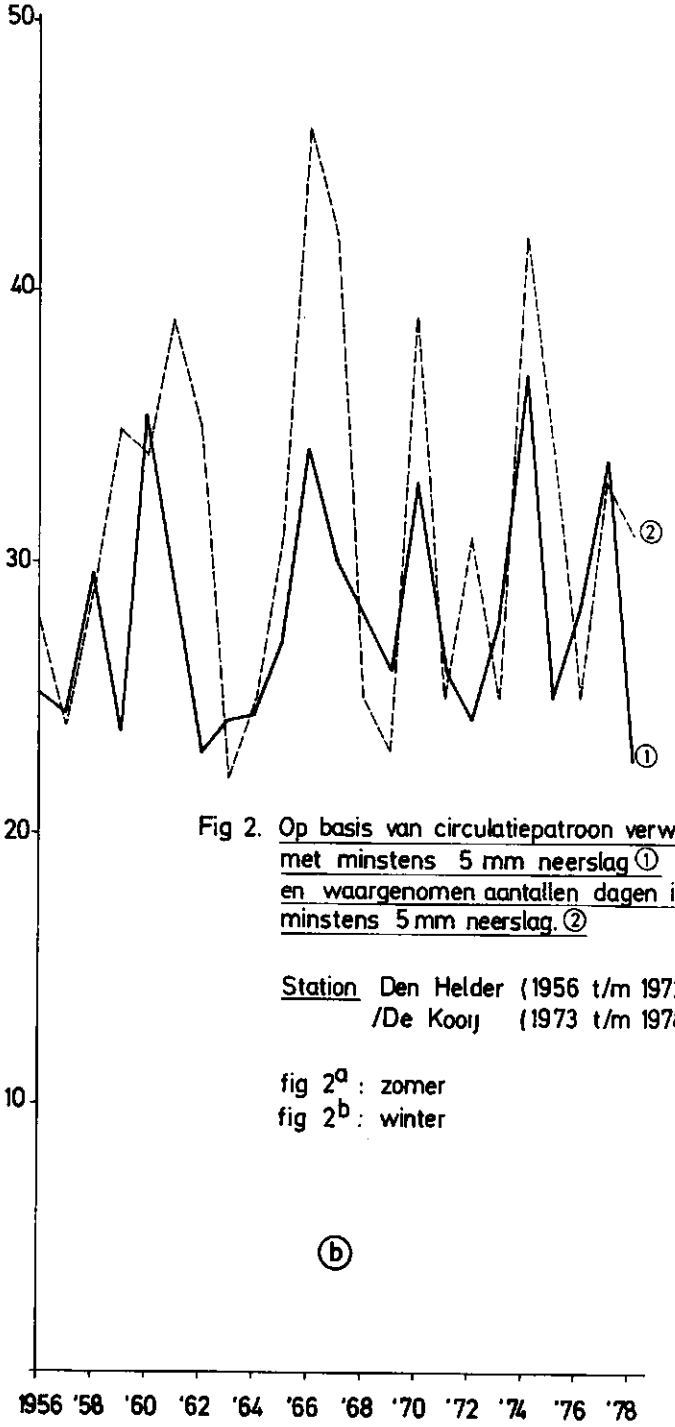
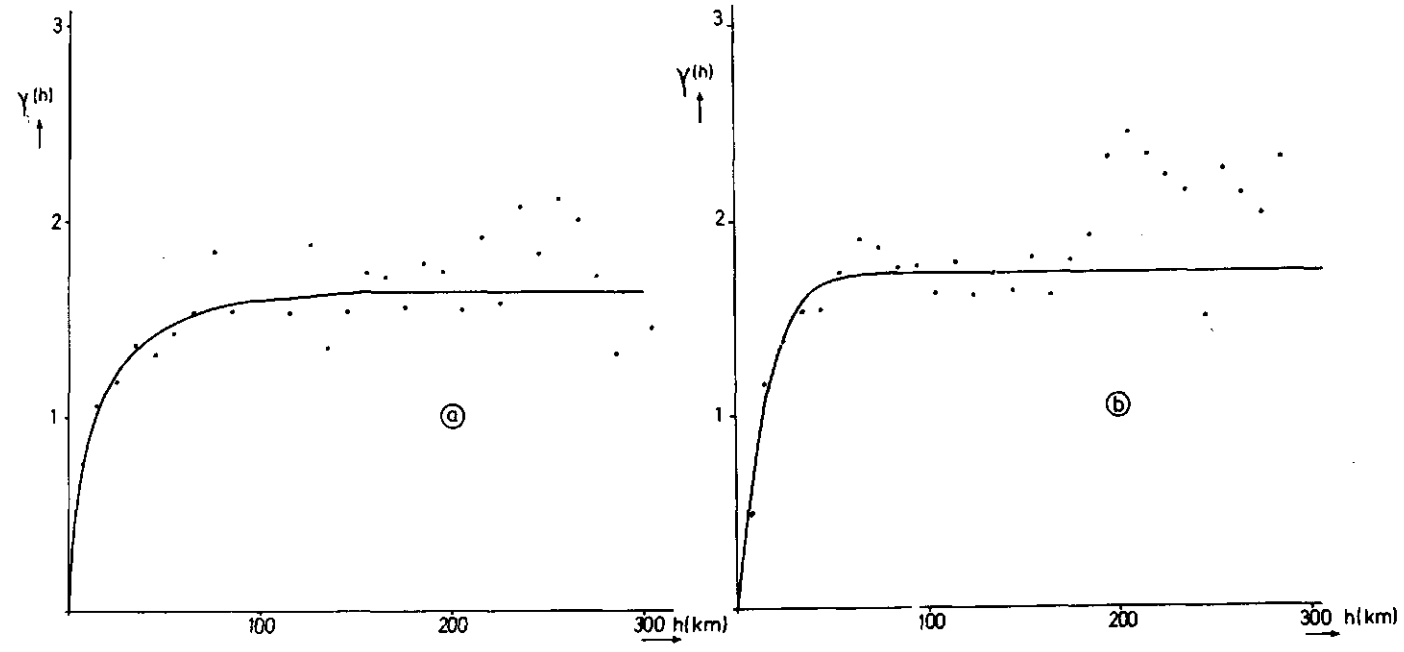


Fig 2. Op basis van circulatiepatroon verwacht aantal dagen/seizoen met minstens 5 mm neerslag ①
en waargenomen aantallen dagen in dat seizoen met minstens 5 mm neerslag. ②

Station Den Helder (1956 t/m 1972)
/De Kooij (1973 t/m 1978)

fig 2^a : zomer
fig 2^b : winter



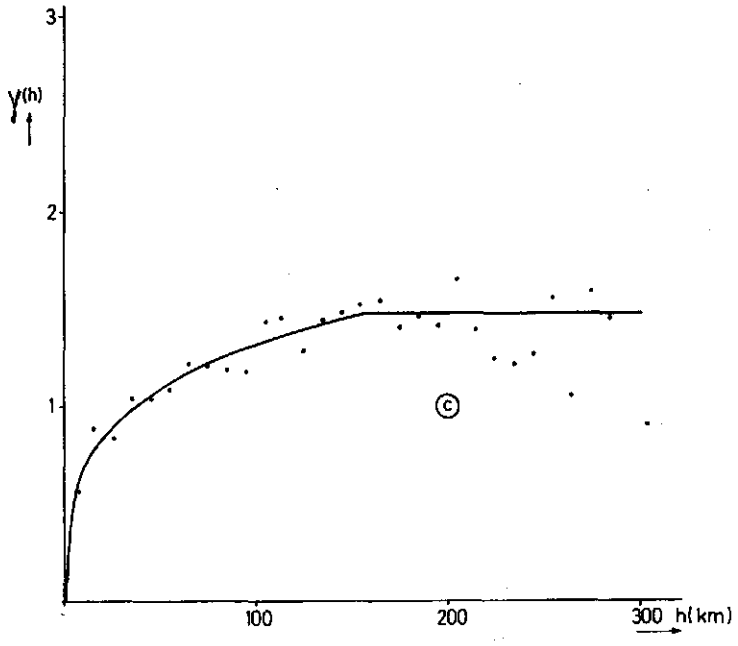


Fig. 3: Steekproef-variogrammen, alsmede de aan de eerste 16 klassengemiddelden aangepaste variogrammen volgens het exponentiële model voor plaatselijke trendverschillen m.b.t. de kenmerken

(a) F^z_{01}

(b) F^z_{15}

(c) F^z_{25}

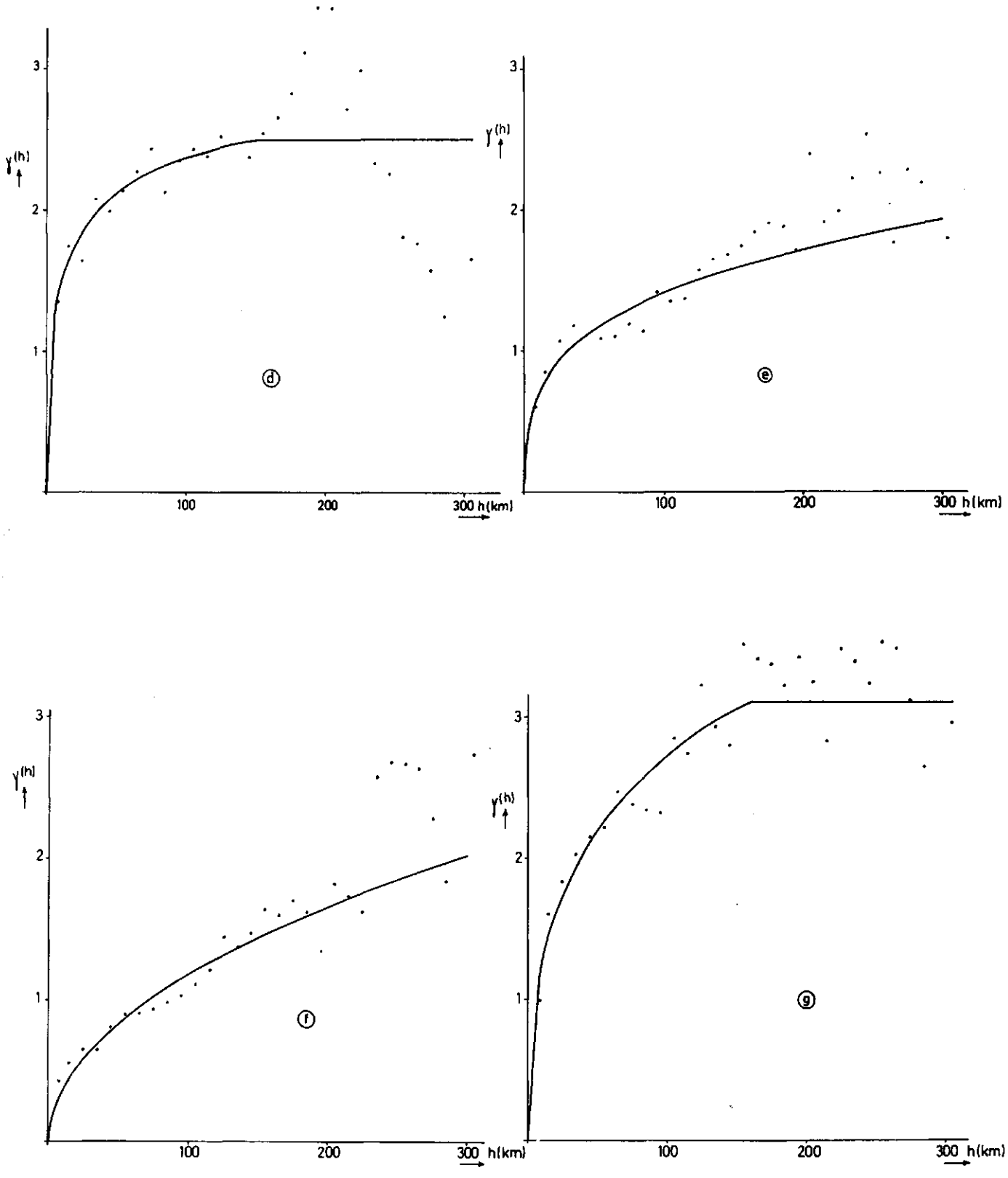


Fig 3. (vervolg) : Steekproef-variogrammen, alsmede de aan de eerste 16 klassengemiddelden aangepaste variogrammen volgens het exponentiële model voor plaatselijke trendverschillen m.b.t. de kenmerken

(d) F_{01}^{w}

(e) F_{15}^{w}

(f) F_{25}^{w}

(g) N^{j}

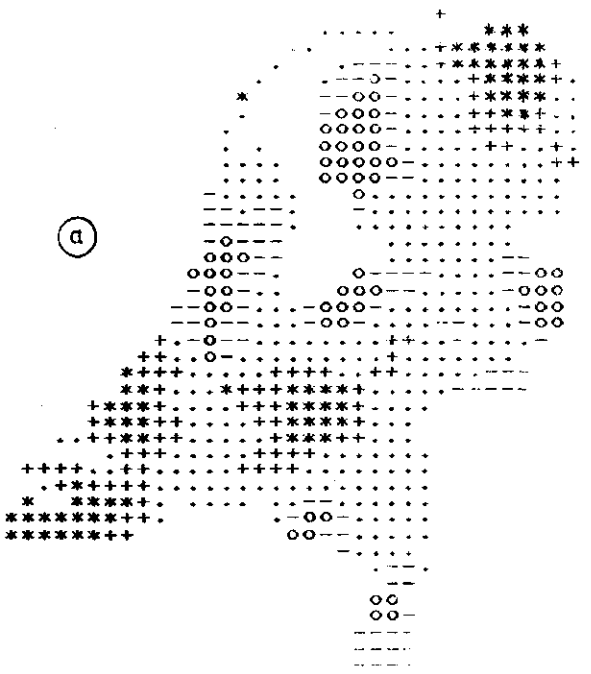


Fig 4. Verdeling over Nederland van geïnterpoleerde $\underline{T}$ waarden, behorende bij de toets op trend, voor de neerslagkenmerken

(a) F^z_{01}

(b) F^z_{15}

(c) F^z_{25}

Legenda bij de figuren $4^{(a)}$ t/m $4^{(c)}$

0 : $\underline{T} \leq -1.00$

- : $-1.00 < \underline{T} \leq -0.50$

. : $-0.50 < \underline{T} < 0.50$

+ : $0.50 \leq \underline{T} < 1.00$

x : $\underline{T} \geq 1.00$

Fig. 4 (vervolg) : Verdeling over Nederland van geïnterpoleerde $\underline{I}$ waarden, behorende bij de toets op trend, voor de neerslagkenmerken

(d) F_{01}^{w}

(e) F_{15}^{w}

(f) F_{25}^{w}

(g) N^{j}

Legenda bij de figuren $4^{(d)}$ t/m $4^{(g)}$

0 : $\underline{I} \leq -1.00$

- : $-1.00 < \underline{I} \leq -0.50$

. : $-0.50 < \underline{I} < 0.50$

+ : $0.50 \leq \underline{I} < 1.00$

x : $\underline{I} \geq 1.00$

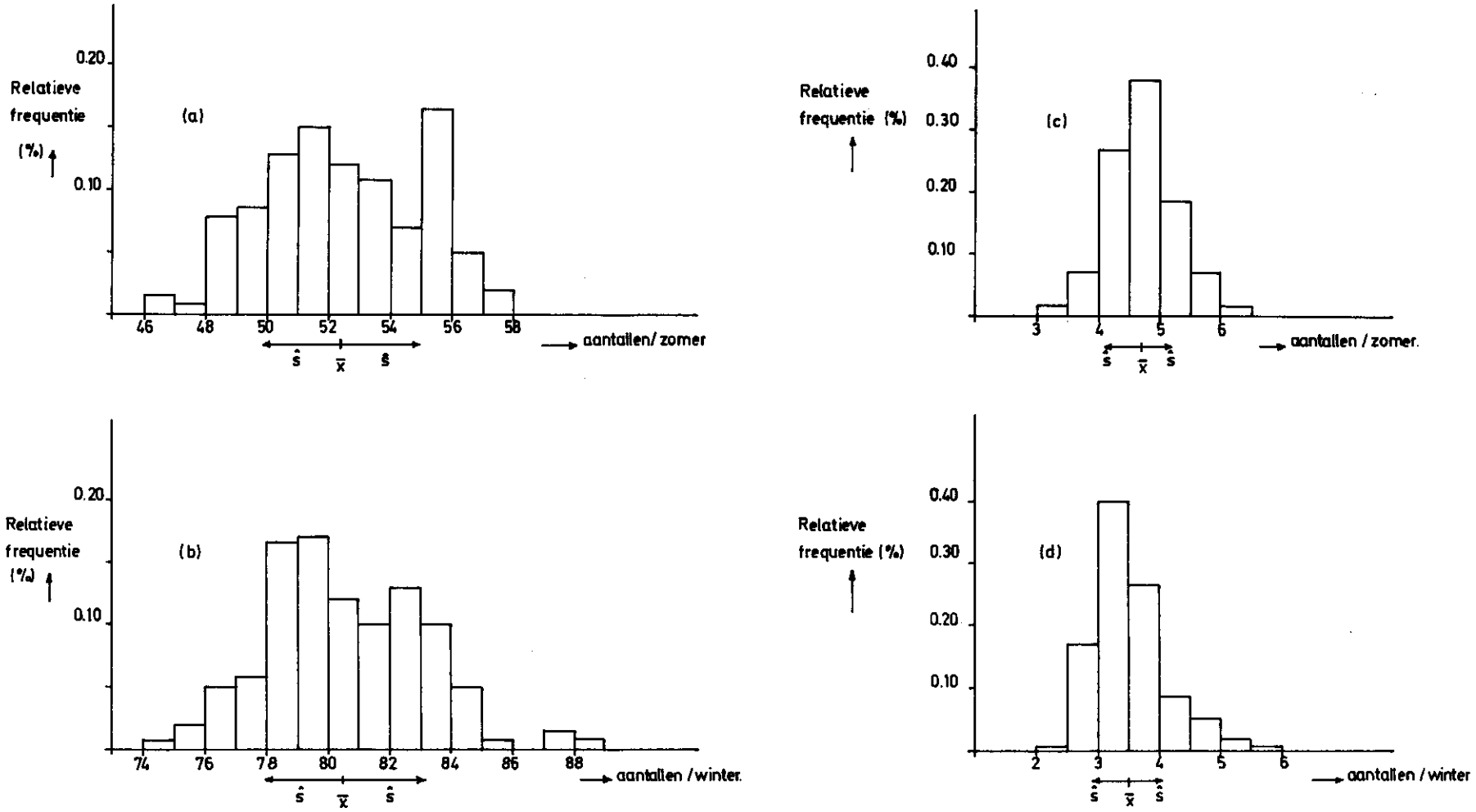
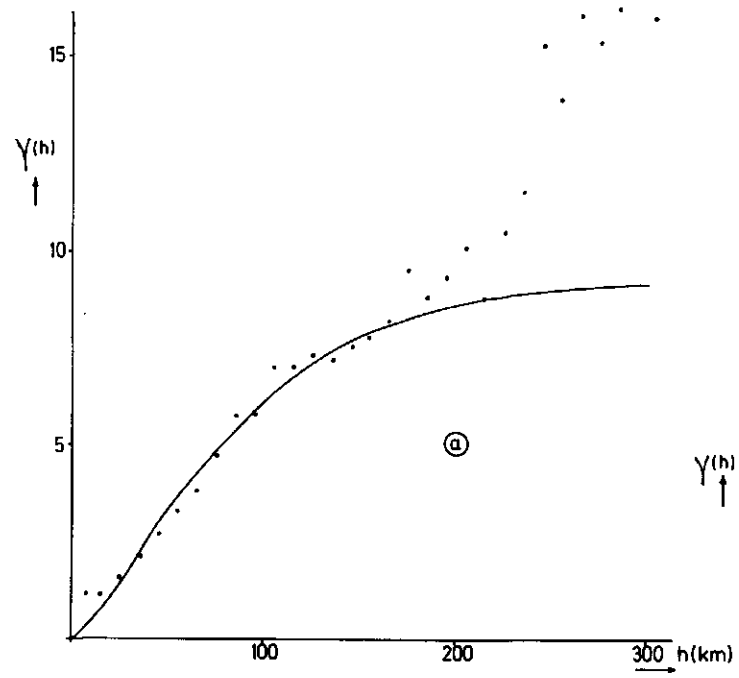
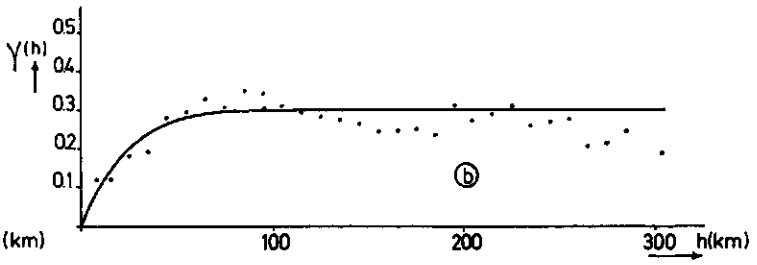


Fig. 5 Relatieve frequentie van stationsgemiddelden m.b.t. de kenmerken

- F_{01}^{z} (fig. a.)
- F_{01}^{w} (fig. b.)
- F_{15}^{z} (fig. c.)
- F_{15}^{w} (fig. d.)





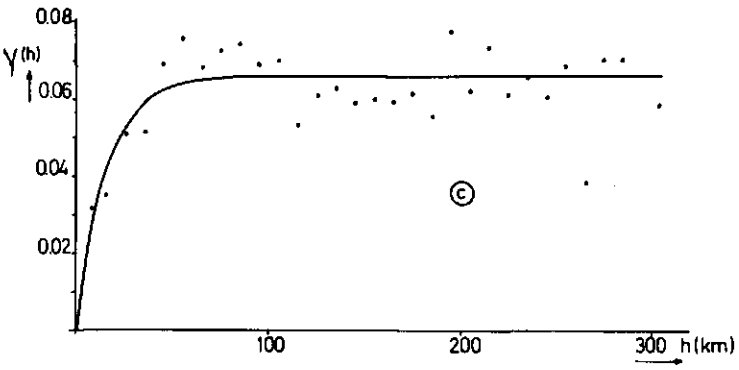


Fig. 6 : Steekproef-variogrammen, alsmede de aan de eerste 16 klassengemiddelden aangepaste variogrammen volgens het exponentiële model voor niveauverschillen m.b.t. de kenmerken

(a) F^z_{01}
(b) F^z_{15}
(c) F^z_{25}

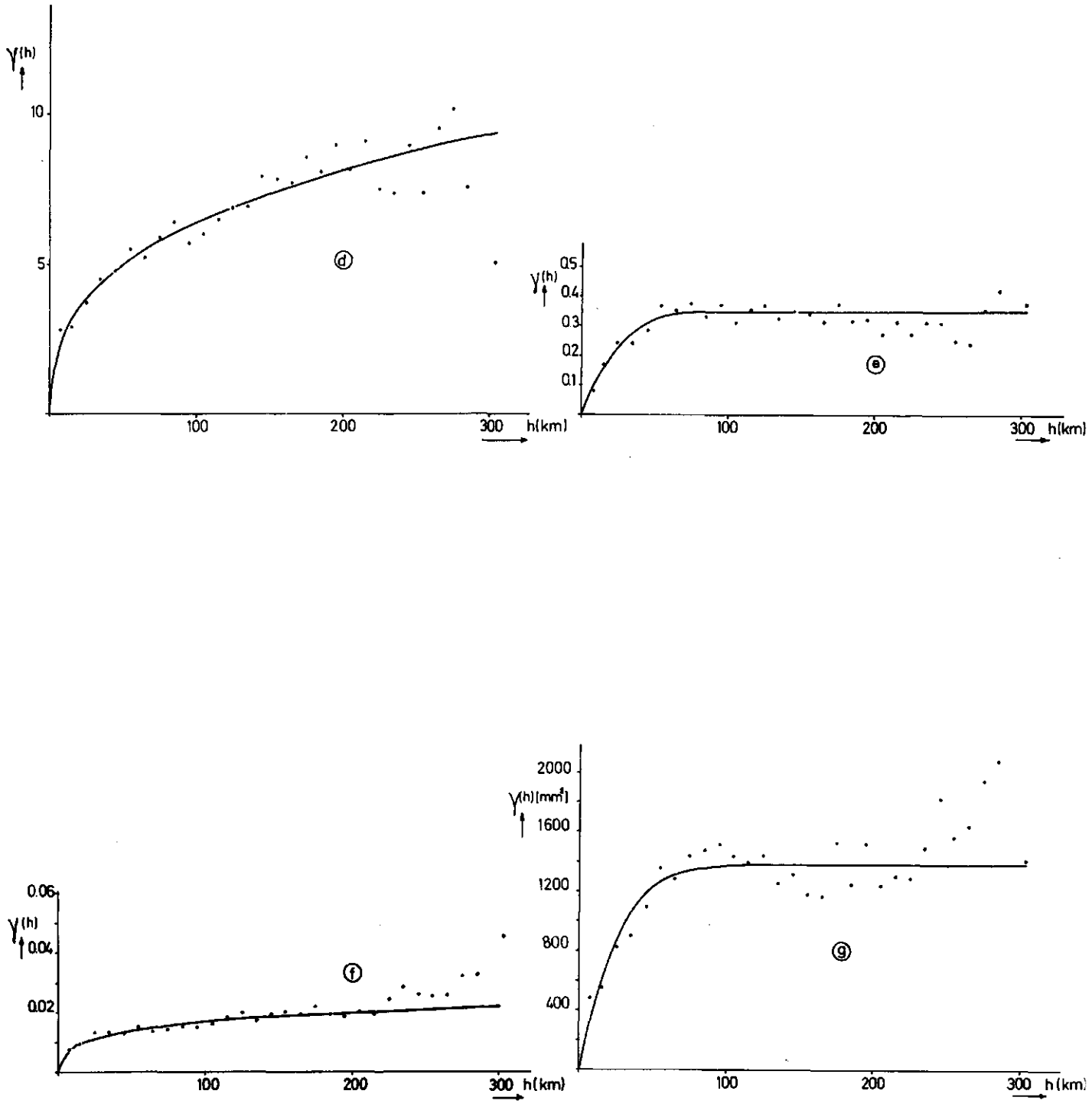


Fig. 6 (vervolg): Semi-variogrammen, alsmede de aan de eerste 16 klassengemiddelden aangepaste variogrammen volgens het exponentiële model voor niveauverschillen m.b.t. de kenmerken

(d) F_{01}^{w}

(e) F_{15}^{w}

(f) F_{25}^{w}

(g) N^{j}

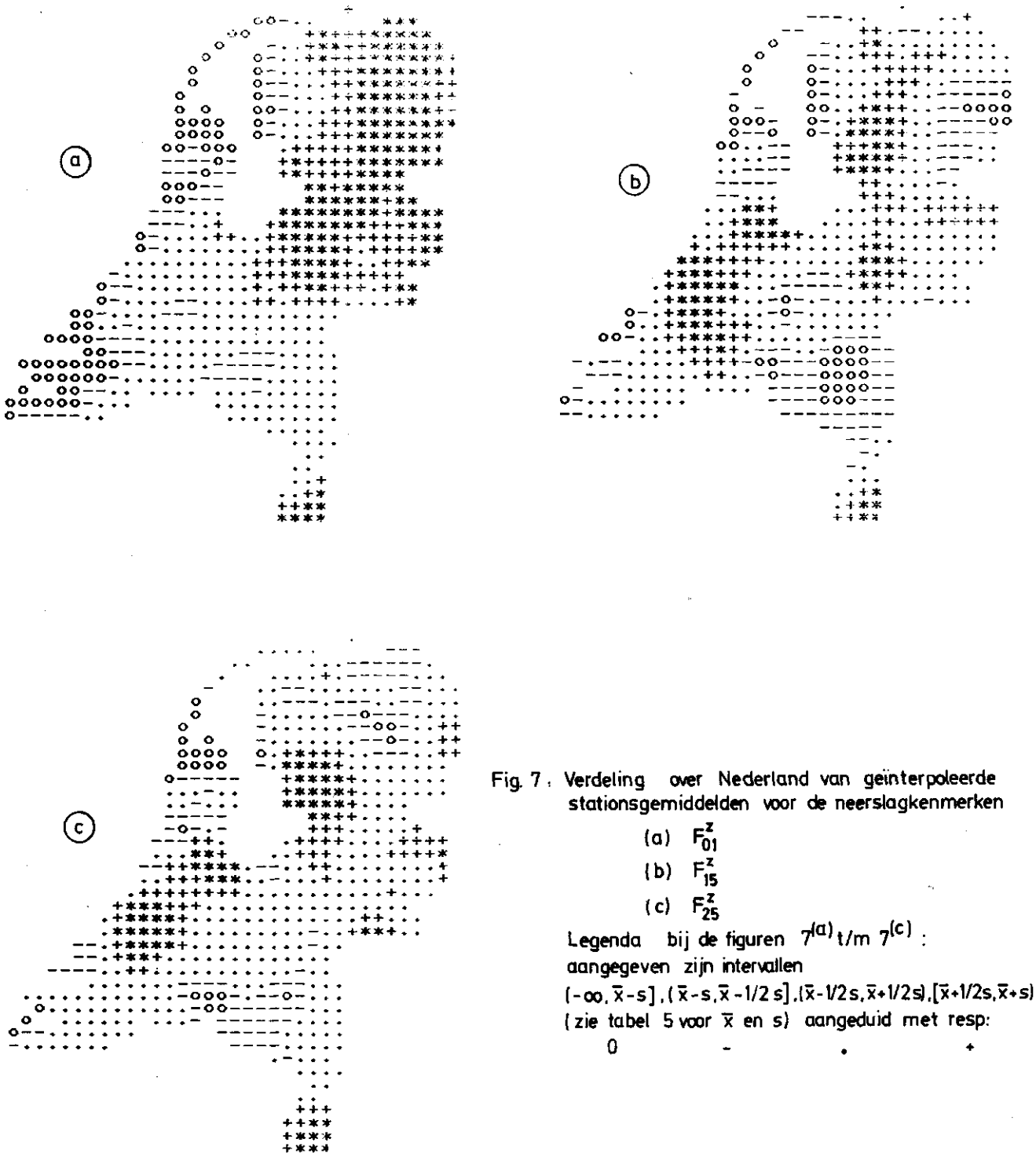


Fig. 7 : Verdeling over Nederland van geïnterpoleerde stationsgemiddelden voor de neerslagkenmerken

(a) F^z_{01}

(b) F^z_{15}

(c) F^z_{25}

Legenda bij de figuren $7^{(a)}$ t/m $7^{(c)}$:
aangegeven zijn intervallen
$(-\infty, \bar{x}-s]$, $(\bar{x}-s, \bar{x}-1/2\,s]$, $(\bar{x}-1/2s, \bar{x}+1/2s)$, $[\bar{x}+1/2s, \bar{x}+s)$, $[\bar{x}+s, \infty)$
(zie tabel 5 voor $\bar{x}$ en s) aangeduid met resp:

0 - . + *

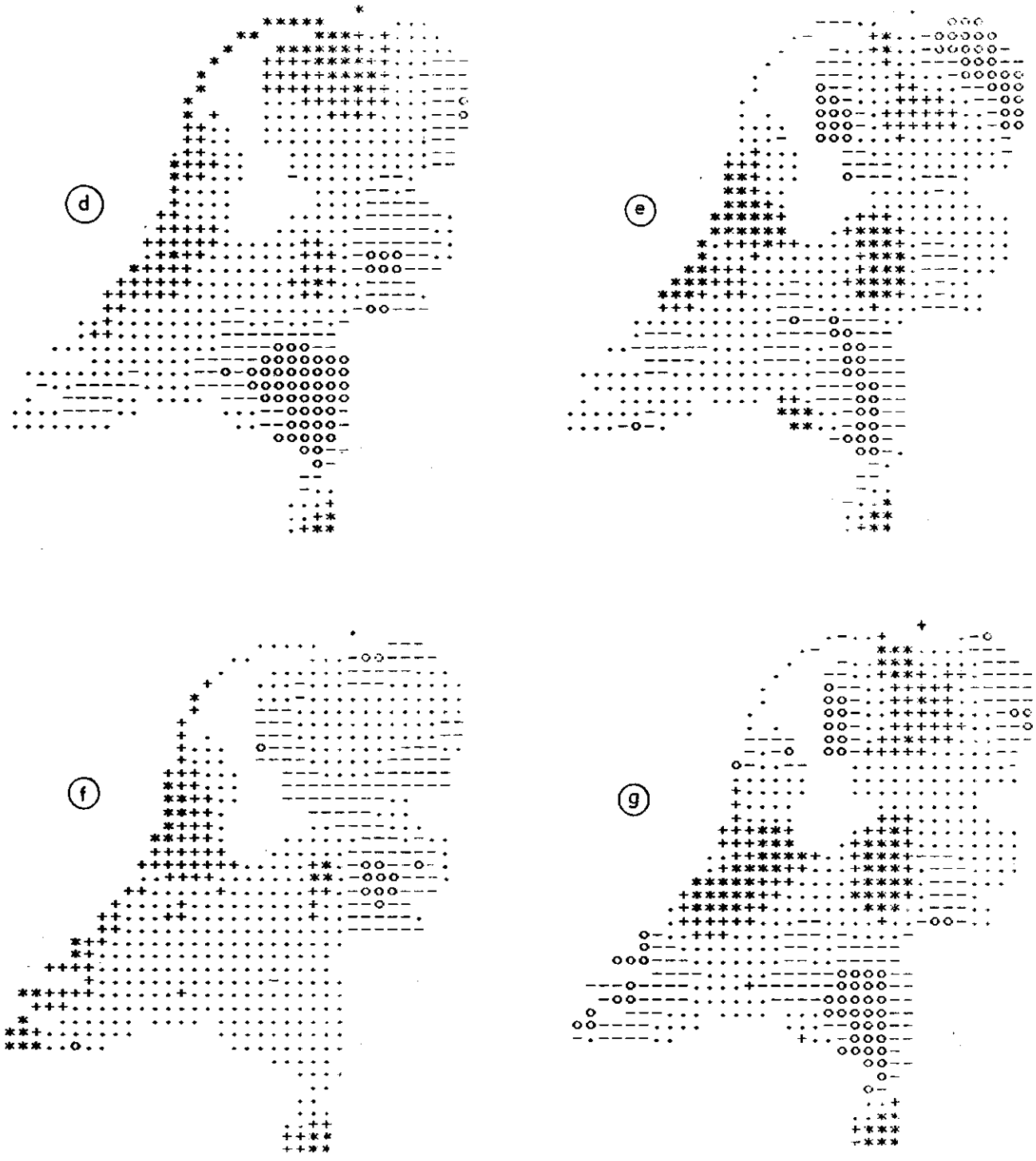


Fig. 7 (vervolg): Verdeling over Nederland van geïnterpoleerde stationsgemiddelden voor de neerslagkenmerken

(d) F_{01}^{w}

(e) F_{15}^{w}

(f) F_{25}^{w}

(g) N^{j}

Legenda bij de figuren $7^{(d)}$ t/m $7^{(g)}$:
aangegeven zijn de intervallen
$(-\infty, \bar{x}-s]$, $(\bar{x}-s, \bar{x}-1/2\,s]$, $(\bar{x}-1/2\,s, \bar{x}+1/2\,s)$, $[\bar{x}+1/2\,s, \bar{x}+s)$, $[\bar{x}+s, \infty)$
(zie tabel 5 voor $\bar{x}$ en s), aangeduid met resp:

0 - • + x

Fig. 8 : Verdeling over Nederland van geïnterpoleerde, gecorreleerde trekkingen $\underline{Z}_i$ uit Normale verdeling

(a) reeks 1

(b) reeks 2

(c) reeks 3

(d) reeks 4

Legenda bij de figuren $g^{(a)}$ t/m $g^{(d)}$

0 : $\underline{Z}_i \leq -1.00$

- : $-1.00 < \underline{Z}_i \leq -0.50$

. : $-0.50 < \underline{Z}_i < 0.50$

+ : $0.50 \leq \underline{Z}_i < 1.00$

x : $\underline{Z}_i \geq 1.00$

Fig. 9: Gebiedsindeling in 5 deelgebieden, corresponderend met een indeling kust - land en Noord - Zuid.

Fig. 10: Gebiedsindeling in 2 deelgebieden, corresponderend met een verstedelijkt (2) en een landelijk deelgebied (overig Nederland.)

Fig. 11: Gebiedsindeling in 3 deelgebieden, corresponderend met een indeling in een nat (1), een droog (2) en een overig deel van Nederland.

APPENDIX I: Lijst van beschouwde stations.

Kolomnummer
1. Plaatsnaam
2. K.N.M.I.-nummer
3. Perioden zonder waarnemingen
4. Neerslagstation dat gebruikt is voor aanvullingen
5. Trend in het kenmerk F^{w}_{01} (bij significante positieve dan wel negatieve trend, aangeduid met '+', resp. '-')
6. Trend in het kenmerk F^{w}_{15}
7. Trend in het kenmerk F^{w}_{25}
8. Trend in het kenmerk F^{z}_{01}
9. Trend in het kenmerk F^{z}_{15}
10. Trend in het kenmerk F^{z}_{25}
11. Trend in het kenmerk N^{j}
12. Nummer van deelgebied bij indeling 1
13. Nummer van deelgebied bij indeling 2
14. Nummer van deelgebied bij indeling 3
15. Nummer van deelgebied bij indeling 4.

Databestand D18:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Den Helder/De Kooij	9/25							-						
W. Terschelling	11				-	-		-		-				
Leeuwarden	85			-						-				
Groningen	139													
Ter Apel	144						+							
Hoorn	222			-			-			+				
Heerde	328									-				
Lemmer	359/44			+	-		+		+					
Hoofddorp	438			-	+	+	-			-				
Scheveningen	440									+				
Putten	542			+				+		+				
Winterswijk	666			-						-				
Vlissingen	733			+	+					+				
Kerkwerve	737/743/462	jan. en feb. '54	Noord-gouwe	-										
Oudenbosch	828			-										
Helmond	896							-		-				
Roermond	961	maart en nov. '71	Beesel	-	-									
Maastricht	976/951	sept. '62	Maas-tricht	+	-	-		+		+				

N.B.: Voor de stations W. Terschelling, Heerde, Kerkwerve, Helmond en Maastricht vertoonden de gegevens voor de periode vanaf 1951 enige lacunes voor de periode 1951 t/m 1953, die met behulp van de gegevens op de CODE2X-band werden aangevuld. Voor de periode na 1953 moesten nóg enige aanvullingen verricht worden; voor zover de hier beschouwde stations niet opgenomen zijn in databestand D140 zijn aanvullingen vermeld in bovenstaande tabel.

Databestand D140:

	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Hollum	10										1	1	1	3
Schiermonnikoog	12							-			3	1	1	3
Petten	16			+							5	1	1	2
Den Burg (Texel)	17										1	1	1	3
Cocksdorp (Texel)	19										1	1	1	3
Dokkum	67										3	2	1	1
Lemmer (Tacozijl)	74			-	-	-	-			-	2	2	1	3
Oldeholtpade	75					-					2	3	1	1
Kornwerderzand	76			-							1	2	1	2
Stavoren	80										1	2	1	2
Gorredijk	82										2	3	1	1
Ezumazijl	84										3	2	1	3
Leeuwarden	85			-	-		-	-		-	2	2	1	3
Groningen	139			-				-			3	3	1	3
Assen	140	aug.'55	Vlagtwedde	-							4	3	1	3
Ter Apel	144					-					4	5	1	2
Zoutkamp	145				+		+			+	3	2	1	3
Sappemeer	148					-	+				3	3	1	2
Roodeschool	151						+			+	3	2	1	2
Winschoten	153			-	-					-	4	3	1	2
Eenrum	154	mrt.'64	Ulrum				+				3	2	1	3
Vlagtwedde	156			-							4	5	1	2
Onnen	158			-	-			-	-		3	3	1	2
Eelde	161							-	-		3	3	1	3
Niekerk	162										3	3	1	3
Marum (Gr.)	166										3	3	1	1
Enkhuizen	221	feb.'51	Hoorn	+		+				+	1	1	1	3
Hoorn (N.H.)	222										6	1	1	3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Overveen	225										5	1	1	3
Schagen	228										5	1	1	3
Zaandijk	230					+		-			6	1	1	3
Bergen (N.H.)	234							-			5	1	1	3
Castricum	235										5	1	1	3
Medemblik	236								+	+	6	1	1	3
De Haukes (N.H.)	238									+	6	1	1	3
Den Oever	239			-							1	1	1	3
Kreileroord	240				+				+	+	1	1	1	2
Marken	246									-	6	1	1	3
Kolhorn	252										6	1	1	3
Wapenveld	329			-	-					-	12	3	1	1
Zwolle	330			+							12	3	1	3
Emmen	333							-		-	6	5	1	3
Rheezerveen	339				-	-				-	13	5	1	3
Zweelo	341								-	-	4	5	1	3
Vroomshoop	345				-	-			-	-	13	5	1	3
Kraggenburg (NOP)	346			-	-						2	2	1	3
Urk	347										1	2	1	3
Emmeloord	348					-		+	+		2	2	1	3
Nagele (NOP)	352							+			2	2	1	3
Blokzijl	353			+							2	2	1	3
Dedemsvaart	354				-	-				-	4	3	1	3
Kuinre	356										2	2	1	3
Lemmer (Buma)	359	dec.'58	Lemmer (Tacoz.)	-			-			-	2	2	1	3
Groot Ammers	434	jan.'53	Oud Alblas	+			+	+		+	7	4	2	3
Sassenheim	436										5	1	1	3
Lijnsden	437						-				6	1	1	1
Hoofddorp	438										6	1	1	1
Oude Wetering	439										5	1	1	1
Scheveningen	440										5	1	1	1
Amsterdam (H.B.)	441										6	1	1	1
Boskoop	442										7	1	2	1
Gouda	443										7	1	2	1
Katwijk a/d Rijn	444									+	5	1	1	3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Rotterdam (W)	445				-				+		7	1	2	1
Delft	449	feb.'51	Naaldwijk								7	1	2	1
Numansdorp	450			+	+					+	9	1	1	3
Bergschenhoek	453			+							7	1	2	1
Mookhoek	455	feb.'53	Numansdorp		+			+		+	10	1	1	3
Oostvoorne	456						+				7	1	1	3
Aalsmeer	458							-			6	1	1	1
Dordrecht	459			+	+			+		+	10	1	2	3
Dirksland	462			+	+		+			+	8	1	1	3
Wassenaar	466			+						+	5	1	1	1
Poortugaal	467	juni,juli '51	IJsselmonde	+						+	7	1	1	1
Leiden	469									+	5	1	1	1
Ouddorp	471	feb.'53	Brielle								8	1	1	2
Nijmegen	539									-	14	5	1	3
Arnhem	541	sept.'55	Arnhem								12	5	1	1
Apeldoorn	543									-	12	4	1	1
Nijkerk	547	11 t/m 28 feb.'55	Lunteren				-				11	4	1	3
De Bilt	550							+			11	4	1	3
Bussum	556										7	4	1	1
Lunteren	558			+				+			11	4	1	1
Tiel	562			+		-	+		+		11	4	1	3
Hulshorst	564	21 t/m 28 feb.'55, dec.'55	Beekbergen	-	-		-			-	12	4	1	3
Harskamp	571			+							11	4	1	1
Beekbergen	573										12	4	1	1
Oosterbeek	578										12	4	1	1
Veenendaal	579										11	4	1	3
Geldermalsen	584								+	+	11	4	1	3
Hilversum	586								+		7	4	1	1
Almelo	664					-			-	-	13	5	1	3
Enschede	665					-					13	5	1	3
Winterswijk	666			-	-						13	5	1	3
Doetinchem	667										13	5	1	3
Hengelo (Ov.)	668					-					13	5	1	3
Borculo	669	14 mei '78	Borculo			-					13	5	1	3
Vliegveld Twente	670				-	-	-			-	13	5	1	3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Gendringen	673			-	-		-		-	-	13	5	1	2
Rekken	674										13	5	1	3
Oldenzaal	676					-	-		-	-	13	5	1	3
Deventer	677	dec.'52	Deventer							-	12	4	1	3
Almen	678										12	5	1	2
Lettele	681	oct.'52	Lettele								12	3	1	3
Vlissingen	733										8	1	1	2
St. Kruis	740						+				9	1	1	3
Terneuzen	742							-	-		9	1	2	3
Axel	745										9	1	2	3
Krabbendijke	747	6 t/m 28 feb.'53	Kapelle						-		9	1	2	3
Vrouwenpolder	751				+						8	1	1	2
Haamstede	752	feb.'53	Oranje-zon							+	8	1	1	2
Middelburg	756					+					8	1	1	2
St. Annaland	759			+	+						8	1	1	3
's-Herenhoek	760										8	1	1	2
Cadzand	763					+		+		+	9	1	1	2
Tilburg	827							+			10	4	1	2
Oudenbosch	828									-	9	4	1	3
Herwijnen	830							+		+	10	4	1	3
Bergen op Zoom	832									-	9	1	1	3
Oosterhout	833										10	4	1	3
Chaam	834										10	4	1	3
Andel	835			+							10	4	1	3
Ginneken	838										10	4	1	3
Hoogerheide	839				+						9	1	1	3
Nieuwendijk (N.B.)	840										10	4	1	3
Vlgb. Gilze Rijen	843										10	4	1	3
Capelle (N.B.)	844	jan,feb. '51	Oosterhout					+	+	+	10	4	1	3
Helmond	896	aug.'51	Gemert								14	5	1	2
Gemert	899									+	5	1	1	2
Nuland	901			+			+			+	14	4	1	3
Eindhoven	902								+		14	5	1	2
Megen	903	feb.'60	Oss				+				14	4	1	3
Deurne	908	juli '69	Helmond						-		14	5	1	2

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Dinther	911										14	4	1	3
Leende	912									-	14	5	1	2
Eersel (Witrijt)	915	juli '52	Eersel					+			10	5	1	3
Vaals	968									-	15	5	2	1
Stramproij	970									-	15	5	1	2
Beek (L.)	973				-					-	15	5	2	3
Buchten	974						-	-	-	-	15	5	1	2

N.B.: Indien ontbrekende gegevens voor een station werden aangevuld met gegevens van ditzelfde station, zijn de ter aanvulling gebruikte gegevens overgenomen uit de K.N.M.I. jaarboeken: "Regenwaarnemingen".

APPENDIX II

ENIGE STATISTISCHE OPMERKINGEN M.B.T. HOMOGENITEITSONDERZOEK

M.A.J. van Montfort (Vakgroep Wiskunde, LH) 1981-06-17

Inhoud

1. Inleiding
2. Een statistisch model
3. Statistische probleemstellingen
4. Commentaar

1. Inleiding

Ingegaan wordt op het volgende soort problemen. Gegeven zij een landkaart met daarop aangegeven metingen in meetpunten (bv. klimatologische stations). Men kan zich dan afvragen of daarvan redelijkerwijze een hoogtekaart gemaakt kan worden; dit is het geval als de metingen van dicht bij elkaar gelegen stations minder variatie vertonen dan die van ver van elkaar verwijderde stations (locale clustering). Een nog vagere formulering kan zijn dat men wil toetsen of alle stations dezelfde verwachting voor de meting hebben tegen het alternatief, dat dit niet zo is.De toe te passen statistische besluitwijze moet rekening houden met eventuele gecorreleerdheid der waarnemingen (bv. als gevolg van cross-correlatie).

Een variant van dit probleem kan zijn te onderzoeken of (vooraf gespecificeerde) gebieden onderling sterker verschillen dan redelijkerwijze te verwachten is op basis van hun binnen-gebied-variatie.

2. Een statistisch model

Laat z_i de meting van station i ($i = 1,\ldots,m$) zijn, en laat $z = (z_1,..z_m)'$ de waarnemingsvector zijn. Als model wordt gekozen dat $\underline{z}$ m-dimensionaal Normaal verdeeld is (eventueel na een geschikt gekozen transformatie) met onbekende verwachtingsvector ξ en (goeddeels) bekende covariantiematrix C; in symbolen:

$$\underline{z} \sim N(\xi, C) \qquad (1)$$

met

$$\underline{z} = (\underline{z}_1, \ldots, \underline{z}_m)' \in \mathbb{R}^m$$

$$\xi = (\xi_1, \ldots, \xi_m)' \in \mathbb{R}^m$$

C : m x m-covariantiematrix

Bij gegeven C (of bij gegeven correlatiematrix A) worden statistische opmerkingen over de verwachtingsvector ξ gemaakt. De probleemstelling geschiedt in woorden en in symbolen.

Enige aanvullende notatie. De ruimte $\mathbb{R}^m$ wordt gesplitst in twee onderling loodrechte deelruimten D en R (dus $D + R = \mathbb{R}^m$ en $D \perp R$). Dan geldt

$$\xi = \xi_D + \xi_R$$

Met betrekking tot ξ, ξ_D en ξ_R worden nulhypotheses H geformuleerd met als alternatief de ontkenning ervan. De toetsingsgrootheden worden genoteerd als $\underline{T}$, en blijken alle een rechtszijdige kritiek te behoeven.

3. Statistische probleemstellingen

Probleem 1: Alle verwachtingen nul?

$$H_1 : \xi = 0$$

Uit (1) volgt $C^{-\frac{1}{2}} \underline{z} \sim N(C^{-\frac{1}{2}} \xi, I) \underset{H_1}{\sim} N(0, I)$ en de toets berust op

$$\underline{T}_1 = \underline{z}' C^{-1} \underline{z} \underset{H_1}{\simeq} \underline{\chi}^2_m$$

Probleem 2: Alle verwachtingen onderling gelijk?

$$H_2 : \xi_R = 0 \text{ met } D = \langle s \rangle \text{ met } s = (1, \ldots, 1)'$$

gelijkwaardig met $\xi \in D$

$$\underline{T}_2 = \underline{z}' C^{-1} \underline{z} - \frac{(s' C^{-1} \underline{z})^2}{s' C^{-1} s} \underset{H_2}{\simeq} \underline{\chi}^2_{m-1}$$

Probleem 3: Som (of gemiddelde) der verwachtingen nul?

$$H_3 : \xi_D = 0 \text{ met } D = < s > \text{ met } s = (1,\ldots,1)'$$

$$\underline{T}_3 = \frac{(\underline{z}' s)^2}{s' C s} \underset{H_3}{\simeq} \underline{\chi}_1^2$$

Deze toetsingsgrootheid is plausibel als men zich realiseert dat H_3 gelijkwaardig is met $\sum_1^m \xi_i = 0$ en dat $\Sigma \xi_i$ geschat wordt met $\Sigma \underline{z}_i$, gelijk aan $\underline{z}$'s met var ($\underline{z}$'s) = s' C s (d.i. de som van de elementen van de covariantiematrix).

Probleem 4: Gebieden intern homogeen?

$$H_4 : \xi_R = 0 \text{ met } D = < e_1,\ldots,e_d > \text{ met dim } (D) = d$$

Speciaal geval hier: D op basis van een indeling in d klassen (lees: gebieden) met een orthogonale basis van klasse-karakteristieke vectoren $e_1,\ldots,e_d$. De te toetsen hypothese luidt dan in woorden: binnen ieder van de d (vooraf gespecificeerde) gebieden zijn de verwachtingen onderling gelijk.
Dan geldt onder H_4

$$E(\underline{z}) = \beta_1 e_1 + \ldots + \beta_d e_d$$

en

$$E(C^{-\frac{1}{2}} \underline{z}) = \beta_1 (C^{-\frac{1}{2}} e_1) + \ldots + \beta_d (C^{-\frac{1}{2}} e_d)$$

De normaalvergelijkingen (NV) luiden:

$$\begin{pmatrix} e_1' C^{-1} e_1 & \ldots & e_1' C^{-1} e_d \\ \cdot & & \cdot \\ \cdot & & \cdot \\ e_d' C^{-1} e_1 & \ldots & e_d' C^{-1} e_d \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \cdot \\ \cdot \\ \hat{\beta}_d \end{pmatrix} = \begin{pmatrix} e_1' C^{-1} z \\ \cdot \\ \cdot \\ e_d' C^{-1} z \end{pmatrix}$$

en leveren projectie^2 van $C^{-\frac{1}{2}} \underline{z}$ op $< C^{-\frac{1}{2}} e_1,\ldots,C^{-\frac{1}{2}} e_d >$.

$$\underline{T}_4 = \underline{z}' C^{-1} \underline{z} - \underline{\text{proj}}^2 \underset{H_4}{\simeq} \underline{\chi}_{m-d}^2$$

Bij het opstellen van de NV kan nuttig gebruik gemaakt worden van de vele nullen in $e_1,\ldots,e_d$; zo is $e_i' C^{-1} e_j$ de som van de elementen van een rechthoekige deelmatrix van C^{-1}.

Probleem 5: Is er verschil tussen intern homogene gebieden?

Stel $D = \langle e_1,\ldots,e_d \rangle$ met d klasse-karakteristieke vectoren

$$D_0 = \langle s \rangle \text{ met } s = e_1 + \ldots + e_d = (1,\ldots,1)'$$

dan geldt

$$\xi = \xi_D + \xi_R \text{ met } \xi_R = 0 \text{ wegens interne homogeniteit}$$

De vraag leidt dan tot

$$H_5 : \xi \in D_0$$

Stel

$$C = \sigma^2 A \text{ met a de bekend veronderstelde corr. matrix}$$

$$\tilde{z} = A^{-\frac{1}{2}} z,\ \tilde{D} = A^{-\frac{1}{2}} D \text{ etc.}$$

dan

$$\underline{T}_5 = \frac{(\underline{\tilde{z}}_{\tilde{D}}^2 - \underline{\tilde{z}}_{\tilde{D}_0}^2) / (d-1)}{(\underline{\tilde{z}}^2 - \underline{\tilde{z}}_{\tilde{D}}^2) / (m-d)} \underset{H_5}{\simeq} \underline{F}_{m-d}^{d-1}$$

Analoog ontstaat bij bekende C met $\underline{\hat{z}} = C^{-\frac{1}{2}} \underline{z}$, $\hat{D} = C^{-\frac{1}{2}} D$ etc.

$$\underline{T}_5 = \underline{\hat{z}}_{\hat{D}}^2 - \underline{\hat{z}}_{\hat{D}_0}^2 \underset{H_5}{\simeq} \underline{\chi}_{d-1}^2$$

4. Commentaar

C1. Als Normaliteit van het gekozen kenmerk discutabel is, kan een Normaliserende transformatie toegepast worden. Zo verdient bij aantallen n analyse van $\sqrt{n} + \sqrt{n+1}$ de voorkeur.

C2. De covariantiematrix C zal doorgaans niet bekend zijn. Bij homoscedastische grootheden (dus var $(\underline{z}_i) = \sigma^2$ voor $i = 1,\ldots,m$) geldt $C = \sigma^2 A$ met A de corr. matrix, en bij genormeerde grootheden geldt $C = A$. Dit leidt tot de deelproblemen van schatting van σ^2 en van A. Vervanging van C door $\hat{C}$ in de toetsingsgrootheden leidt hopelijk niet tot sterk afwijkende kritieke waarden.

C3. Een vaak gebruikt model voor de schatting van A gaat ervan uit, dat ρ_{z_i,z_j} - te noteren als ρ_{ij} - slechts afhangt van de onderlinge afstand d_{ij} van station i en station j (isotropie). Het model luidt dan bijv.

$$\rho_{ij} = \exp\,(-\lambda_1\, d_{ij}^{\lambda_2}) \text{ met } \lambda_1 > 0$$

of in een andere notatie $\rho_d = \exp\,(-\lambda_1\, d^{\lambda_2})$.
Een ander veelvuldig gebruikt model gaat uit van

$$\underline{z}_i - \underline{z}_j \simeq (\xi_i - \xi_j) + \sqrt{2\sigma^2\,(1-\rho_{ij})}\;\underline{x} \quad \text{mits } \sigma_i = \sigma_j = \sigma$$

waaruit volgt

$$E(\underline{z}_i - \underline{z}_j)^2 = (\xi_i - \xi_j)^2 + 2\sigma^2(1-\rho_{ij})$$

met als mogelijke benadering $(\xi_i - \xi_j)^2 = 0$ voor d_{ij} klein.

C4. Probleem 5 is (zonder acht te slaan op de waarnemingen) veel specifieker dan Probleem 2. Indien het mogelijk is gebieden te vormen, zal dit het onderscheidingsvermogen bij onderzoek naar inhomogeniteit sterk verhogen. De oplossing van Probleem 5 vereist niet kennis van de gemeenschappelijke variantie σ^2.

C5. Als bij probleem 4 de matrix C onbekend is, zou deze geschat moeten worden bij mogelijke aanwezigheid van verschillen tussen gebieden.

C6. De bij bekende C bij T_5 benodigde $proj^2$, te weten $(s'C^{-1}z)^2/s'C^{-1}s$, is reeds bij T_2 berekend. De onder T_4 vermelde NVlaten concludering m.b.t. verschillen (of contrasten) van gebieden toe. Van de verschillen $\beta_i - \beta_j$ geeft de oplossing der NV een schatting, en uit de ev. gebruikte inverse bij de oplossing volgt de variantie van deze schatting.