ISSN 0169-1708

KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT

TECHNISCHE RAPPORTEN

T.R. - 70

A. Denkema

Onderzoek naar de homogeniteit van reeksen zonneschijnduren in Nederland gedurende het tijdvak 1951-1980

De Bilt, 1985

Publikatienummer: K.N.M.I. TR - 70 (KD)

Koninklijk Nederlands Meteorologisch Instituut
Klimatologische Dienst
Postbus 201
3730 AE DE BILT
Nederland

U.D.C.: 551.521.11 (492)
519.2

ISSN 0169-1708

Onderzoek naar de homogeniteit van reeksen zonneschijnduren in Nederland gedurende 1951-1980

A. Denkema

Inhoud

English summary on next page

A. Denkema,

Statistical tests on the homogeneity of series of sunshine duration in the Netherlands in the period 1951-1980

Summary

In this report several tests on homogeneity have been applied to series of yearly sunshine duration, measured in the period 1951-1980 at 35 stations in the Netherlands.

Although the series of 13 stations showed evidence of non-homogeneity, the magnitude of it proved to be too small for practical applications.

1. Inleiding

Eind december 1980 bedroeg het aantal zonneschijnstations met inbegrip van de hoofdstations 35. Aan het eind van de veertiger jaren was dit aantal nog ongeveer 20. In het begin van de vijftiger jaren vindt dan echter een forse uitbreiding plaats. Als ten behoeve van de berekening van normalen van zonneschijnduur over het tijdvak 1951-1980 vele ontbrekende gegevens zijn ingeschat, kan over een volledig bestand van 30 jaren zonneschijnduur op 35 stations worden beschikt.
In dit rapport zullen de resultaten van een onderzoek naar de homogeniteit van reeksen jaarsommen van de zonneschijnduur op deze stations besproken worden.

2. De gebruikte homogeniteitstoetsen

In de loop der jaren zijn diverse toetsen ontwikkeld, om de homogeniteit van tijdreeksen te onderzoeken. In deze paragraaf zal een drietal toetsen worden behandeld, namelijk het quotiënt van Von Neumann, een toets op het maximum en het minimum van partiële sommen en de toets van Worsley.

2.1. Het quotiënt van Von Neumann

Een toets, die in de klimatologische literatuur nogal eens gebruikt wordt, berust op het zgn. quotiënt van Von Neumann

$$Q = \frac{\sum_{i=1}^{n-1} (Y_{i+1} - Y_i)^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}, \qquad (1)$$

waarin $\overline{Y}$ het gemiddelde van de n waarden Y_i is. De toets mag gebruikt worden als de Y_i's normaal verdeeld zijn en onafhankelijk zijn (Buishand, 1981; WMO, 1966). Ten aanzien van jaarsommen van de zonneschijnduur houdt dit geen beperking in, daar ze niet gecorreleerd zijn. Onder de nulhypothese van een constant gemiddelde volgt de grootheid Q een symmetrische verdeling met gemiddelde $\mu_Q = 2$ en variantie

$$\sigma_Q^2 = \frac{4\,(n-2)}{(n-1)(n+1)} .$$

Is n niet te klein (zeg $n > 20$), dan kan deze verdeling van Q goed worden benaderd door een normale verdeling, zodat de grootheid

$T = \frac{Q - 2}{\sigma_Q}$ bij benadering standaardnormaal verdeeld is.

Daar bij het optreden van een stijgend of dalend verloop in de waarnemingsreeks de noemer van (1) sterker wordt beïnvloed dan de teller, zal Q dan in het algemeen relatief kleine waarden aannemen. In dit geval wordt dus linkseenzijdig getoetst.

2.2. Maxima en minima van partiële sommen

Een andere toets, gebaseerd op de maxima en minima van partiële sommen, veronderstelt evenals die van Von Neumann onderling onafhankelijke waarden Y_i, die normaal verdeeld zijn.
Voor praktijktoepassingen mag echter de kansverdeling van de Y_i's wel enigszins van de normale afwijken (Buishand, 1981).
De partiële sommen, gecorrigeerd op het gemiddelde (Eng.: "adjusted partial sums"), worden als volgt gedefinieerd:

$$\begin{cases} S_k^* = \sum_{i=1}^{k} (Y_i - \bar{Y}) & k = 1,\ldots\ldots,n \\ S_o^* = 0 & \end{cases} \qquad \ldots\ldots(2)$$

Blijkbaar is $S_n^* = \sum_{i=1}^{n} Y_i - n\bar{Y} = 0.$

De maxima en minima van de "adjusted partial sums" duiden we aan als M_n en m_n:

$$M_n = \max_{o \leq k \leq n} S_k^* \qquad \ldots\ldots(3)$$

$$m_n = \min_{o \leq k \leq n} S_k^* \qquad \ldots\ldots(4)$$

en de range als R_n^*

$$R_n^* = M_n - m_n \qquad \ldots\ldots(5)$$

Als toetsingsgrootheid wordt de "rescaled adjusted range" R gebruikt:

$$R = R_n^*/s, \qquad \ldots\ldots(6)$$

waarin s voor de steekproefstandaardafwijking staat. Zij is gevoelig voor een verschuiving in het gemiddelde. Als men hoge waarden voor R vindt, wijst dit op een niet-homogene reeks.
Het is verre van eenvoudig, met analytische methoden de volledige kansverdeling van R af te leiden. Met behulp van een Monte-Carlo methode zijn echter, voor diverse waarden van n, percentielen van R met voldoende nauwkeurigheid verkregen (Buishand, 1982).

2.3. Toets van Worsley

Ook bij de derde toets, die toegepast is op de hypothese H_0: "de tijdsreeks $Y_1 \ldots Y_n$ is afkomstig uit een normaal verdeelde populatie $N(\mu,\sigma^2)$" met als alternatief, dat op een onbekend moment de verwachting μ met een, eveneens onbekend, bedrag Δ is toegenomen, wordt gebruik gemaakt van partiële sommen (Worsley, 1979; Buishand, 1982). We definiëren:

$$S_k^{**} = S_k^{*}/s \qquad .....(7)$$

(Eng.: rescaled adjusted partial sums) en:

$$V = \max_{1 \leqslant k \leqslant n-1} \left\{ \frac{1}{\sqrt{k(n-k)}} \left| S_k^{**} \right| \right\} \qquad .....(8)$$

Als toetsingsgrootheid voert Worsley in:

$$W = \frac{V.\ \sqrt{n-2}}{\sqrt{1 - V^2}} \qquad .....(9)$$

In genoemde publikatie bespreekt Worsley de eigenschappen van de nulverdeling van W. Daarin is ook een tabel opgenomen met de 90, 95 en 99 percentielen (procentpunten) van W onder de nulhypothese. Deze punten zijn voor de steekproefgrootten n = 3(1)10 exact berekend, voor n = 15(5)50 is gebruik gemaakt van een Monte-Carlo methode.

3. Toepassing op jaarsommen van de zonneschijnduur

3.1. Variantie-reductie

Een aanzienlijke reductie van de variantie wordt verkregen door over te gaan van de oorspronkelijke reeks van jaarsommen $\{x_i\}$ van een individueel station op de verschillenreeks

$$\{Y_i\} = \{x_i - g_i\}$$

waarin g het gemiddelde is van alle overige stations in het jaar i. Deze variantiereductie komt het onderscheidingsvermogen van de drie toetsen ten goede (Buishand, 1981).
Uit berekeningen blijkt, dat voor het grootste deel van de 35 onderzochte stations s_Y tussen de 25 en 45 uren ligt. Voor de reeks van De Bilt geldt voor het tijdvak 1951-1980:
s_Y = 34,0 uren en s_X = 148,3 uren.

3.2. Resultaten van het onderzoek

In tabel 1 zijn de resultaten van de drie toetsen weergegeven, van Von Neumann in kolom Q, van de toets, die op partiële sommen berust in kolom R en tenslotte van de toets Worsley in de kolom W. Bij de getallen zijn soms kruisjes gezet, die de volgende betekenis hebben:

x enige aanwijzing voor een inhomogeniteit; de toetsingsgrootheid is significant bij een onbetrouwbaarheidsdrempel van 0,10, maar niet bij een onbetrouwbaarheidsdrempel van 0,05;

xx sterkere aanwijzing voor een inhomogeniteit; de toetsingsgrootheid is significant bij een onbetrouwbaarheidsdrempel van 0,05, maar niet bij een onbetrouwbaarheidsdrempel van 0,01.

xxx sterke aanwijzing voor een inhomogeniteit; de toetsingsgrootheid is significant bij een onbetrouwbaarheidsdrempel van 0,01.

Tellen we het aantal kruisjes bij elkaar, dan duidt een totale score 0 op geen enkele aanwijzing voor een inhomogeniteit, terwijl een score 9 er juist een zeer sterke aanwijzing voor is. Als er van uitgegaan wordt, dat een score van 5 of meer kruisjes een duidelijke evidentie voor inhomogeniteit is, dan zijn er 13 stations, waarvan dit gezegd kan worden. Deze 13 stations zijn Lelystad, Numansdorp, Schiermonnikoog, Valkenburg ZH, Schiphol, Twente VB, Vlissingen, Zierikzee, Amsterdam, Harderwijk, IJmuiden, Ouddorp en Dedemsvaart. In figuur 1 is de ligging van deze stations aangegeven met Δ , de overige stations (dus met een homogene reeks) met • Het valt op, dat beide soorten stations min of meer random over Nederland verdeeld zijn.

3.3. Enige mogelijke oorzaken van inhomogeniteiten

De vaststelling van het feit, dàt er inhomogeniteit is, werpt direct de vraag naar de oorzaak ervan op. Deze vraag blijkt echter moeilijk te beantwoorden. De meest voor de hand liggende oorzaak lijkt verandering van opstelling te zijn. Zo heeft de zonneschijnmeter van Vlissingen tot mei 1958 op het 1700 m landinwaarts gelegen vliegveldje van Souburg gestaan. Het gemiddelde verschil (Souburg - gemiddelde overige stations) over de jaren 1951-1957 blijkt evenwel 87,4 uren groter te zijn dan het gemiddelde verschil (Vlissingen-gemiddelde overige stations) over de jaren 1959-1980, terwijl men van een dicht aan zee gelegen station juist het grootste verschil zou verwachten. Om dit uit te zoeken, is het nodig na te gaan in welke uren van de dag en maanden verschillen optreden. Een mogelijke oorzaak zou mist boven zee kunnen zijn.
Te Numansdorp was de zonneschijnmeter van 24 oktober 1951- 18 april 1961 opgesteld op het grasveld vóór de woning van de waarnemer en van 18 april 1961 - april 1972 op de havenpier. In mei 1972 werd de zonneschijnmeter verplaatst naar de tuin van een nieuwe waarnemer en in juni 1975 naar een kwekerij op 2,5 km ten noordoosten van deze tuin. De gemiddelde verschillen (Numansdorp - gemiddelde overige stations) bedroegen voor de eerste twee tijdvakken (1952-1960 en 1962-1971) resp. 58,2 en 6,1 uren, dus ook hier weer vlak aan het water minder zon dan landinwaarts. De waarnemer merkte op, dat er dikwijls nevel lag over het Hollands Diep, maar niet over het land.
De zonneschijnstroken zijn in de loop der jaren uiteraard door verschillende personen "uitgetrokken", d.i. van dag tot dag werd het aantal uren zon bepaald. Hoewel een persoonseffect niet uitgesloten is, zal dit effect bij overgang op de verschillen reeks $Y_i = x_i - g_i$ geheel geëlimineerd zijn. Wel zou een mogelijke oorzaak nog gelegen kunnen zijn in het gebruik van verschillende glazen bollen en papiersoorten. Braak (1937) vond bij parallelmetingen in 1927, dat een nieuw type bol 5% meer zonneschijn opleverde dan de oude bol. Uit de stationsdocumentatie blijkt, dat te Amsterdam in mei 1967 de grote bol werd vervangen door een kleiner type. Berekenen we over de tijdvakken 1955-1966 en 1968-1980 de gemiddelde verschillen (Amsterdam - gemiddelde overige 34 stations), dan wordt gevonden resp. 72,3 en 2,4 uren, hetgeen in overeenstemming is met het feit, dat de grote bol iets meer zon registreert dan de kleine bol.

4. Gevolgen van inhomogeniteiten

4.1. Het effect op het isoheliënpatroon

Nu we gezien hebben, dat er enige "verdachte" stations zijn, doet zich de vraag voor, of de geconstateerde inhomogeniteit groot genoeg is, om het isoheliënpatroon drastisch te verstoren. Uit figuur 2, waarin de isoheliën van de jaarnormalen 1951-1980 getrokken zijn met gebruikmaking van alle 35 stations, blijkt wel, dat hier geen sprake van is. Uiteraard is dit afhankelijk van de gekozen niveaus van de isoheliën. De 13 in § 3.2 genoemde stations passen toch nog goed in het isoheliënbeeld. De oorzaak hiervan is dat het effect van de meest ernstige inhomogeniteiten op het dertigjarig gemiddelde ongeveer 30 uren per jaar bedraagt. Dit bedrag is klein ten opzichte van de gradiënt in de gemiddelde zonneschijnduur over Nederland.

4.2. Het effect op de berekende globale straling

In hun Wetenschappelijk Rapport W.R. 82-5 leidden Frantzen en Raaff de volgende relatie af tussen globale straling en zonneschijnduur voor het kustgebied:

$$G/Q_A = 0{,}21 + 0{,}61\ n/N \qquad (10)$$

Hierin is G de maandgemiddelde dagintensiteit van de globale straling;
Q_A is de gemiddelde maandelijkse dagintensiteit van de aan de rand van de atmosfeer invallende straling
n is de maandsom van de zonneschijnduur
N is de maximaal mogelijk duur van de zonneschijn.

In paragraaf 3.3 vonden we een inhomogeniteit van 87,4 uren in de zonneschijnduur van Vlissingen. Dit komt overeen met ruim 5% van de normale jaarsom. Nu bedraagt n/N voor Vlissingen gemiddeld 0,35, zodat volgens (10) geldt:
$\overline{G/Q_A} = 0{,}42$. Een inhomogeniteit van 87,4 uren heeft tot gevolg dat het gemiddelde van n/N met ongeveer 0,02 verandert. Uit (10) volgt, dat $\overline{G/Q_A}$ hierdoor met ongeveer 0,01 verandert (gemiddeld nog geen 3%). Dergelijke kleine veranderingen in de berekende globale straling zijn voor de meeste toepassingen (o.a. verdampingsberekeningen) niet relevant.

4.3. Conclusie

Hoewel er bij 13 reeksen sterke aanwijzingen bestaan voor inhomogeniteiten, blijken deze voor de praktijk toch van weinig betekenis te zijn. Blijkbaar beschikt ons land over voldoende goede reeksen zonneschijnduurgegevens.

5. Dankbetuiging

De schrijver betuigt gaarne zijn dank aan Dr. Ir. Tj. A. Buishand voor het kritisch doorlezen van het manuscript.

6. Literatuur

C. Braak (1937), Het klimaat van Nederland.
F. Zonneschijn en bewolking. Mededelingen en Verhandelingen No. 40. (KNMI-publ.)

W.M.O. (1966), Climatic change. Technical Note No. 79.

K.J. Worsley (1979), On the likelihood ratio-test for a shift in location of normal populations. Journal of the American Statistical Association 74 365-367

Tj. A. Buishand (1981), The analysis of homogeneity of long-term rainfall records in the Netherlands. Wet. Rapport 81-7. (KNMI-publ.)

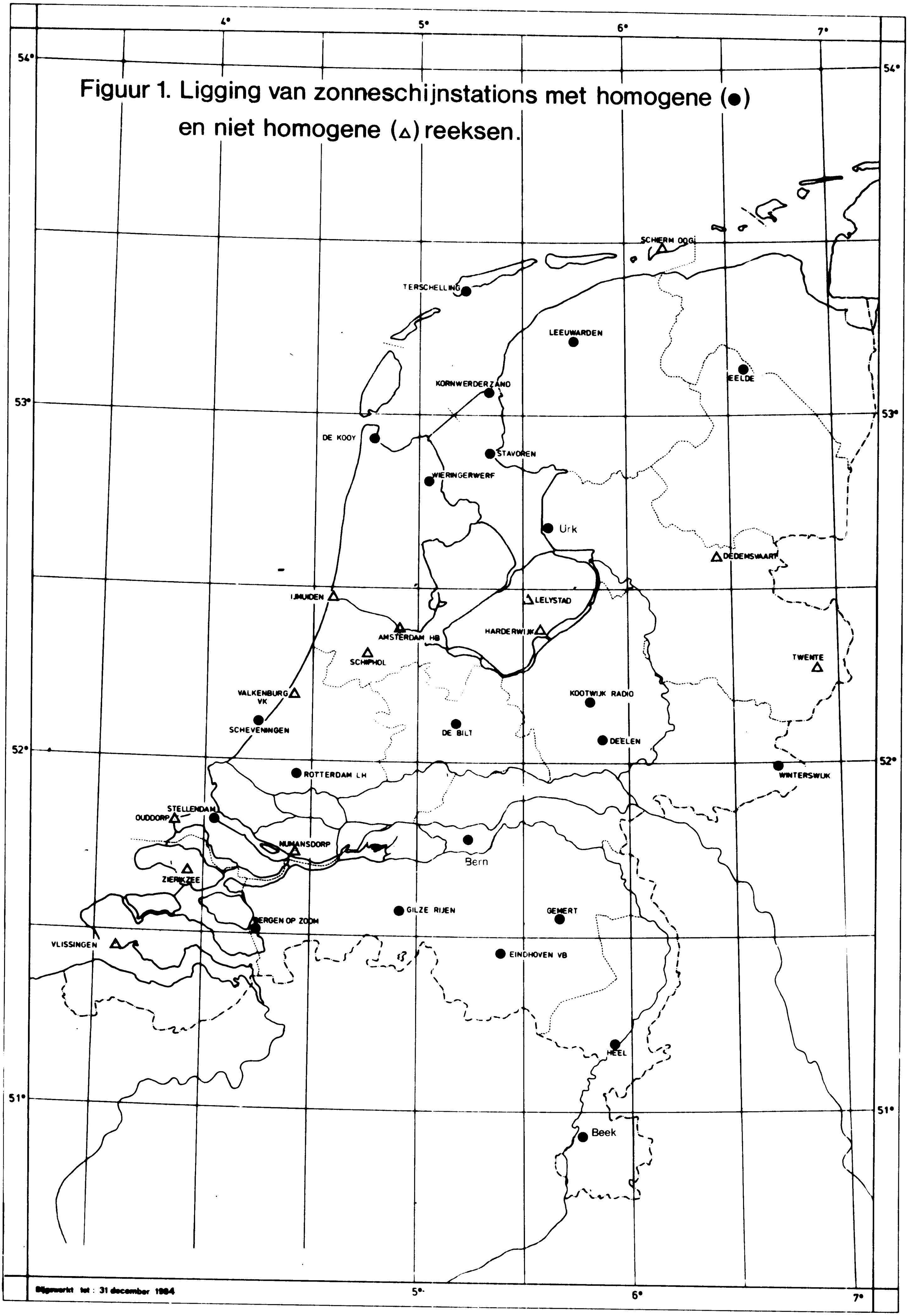
Tj. A. Buishand (1982), Some methods for testing the homogeneity of rainfall records.
Journal of Hydrology 58 11-27.

A.J. Frantzen en W.R. Raaff (1982). De relatie tussen de globale straling en de relatieve zonneschijnduur in Nederland. Wet. Rapport 82-5. (KNMI-publ.)

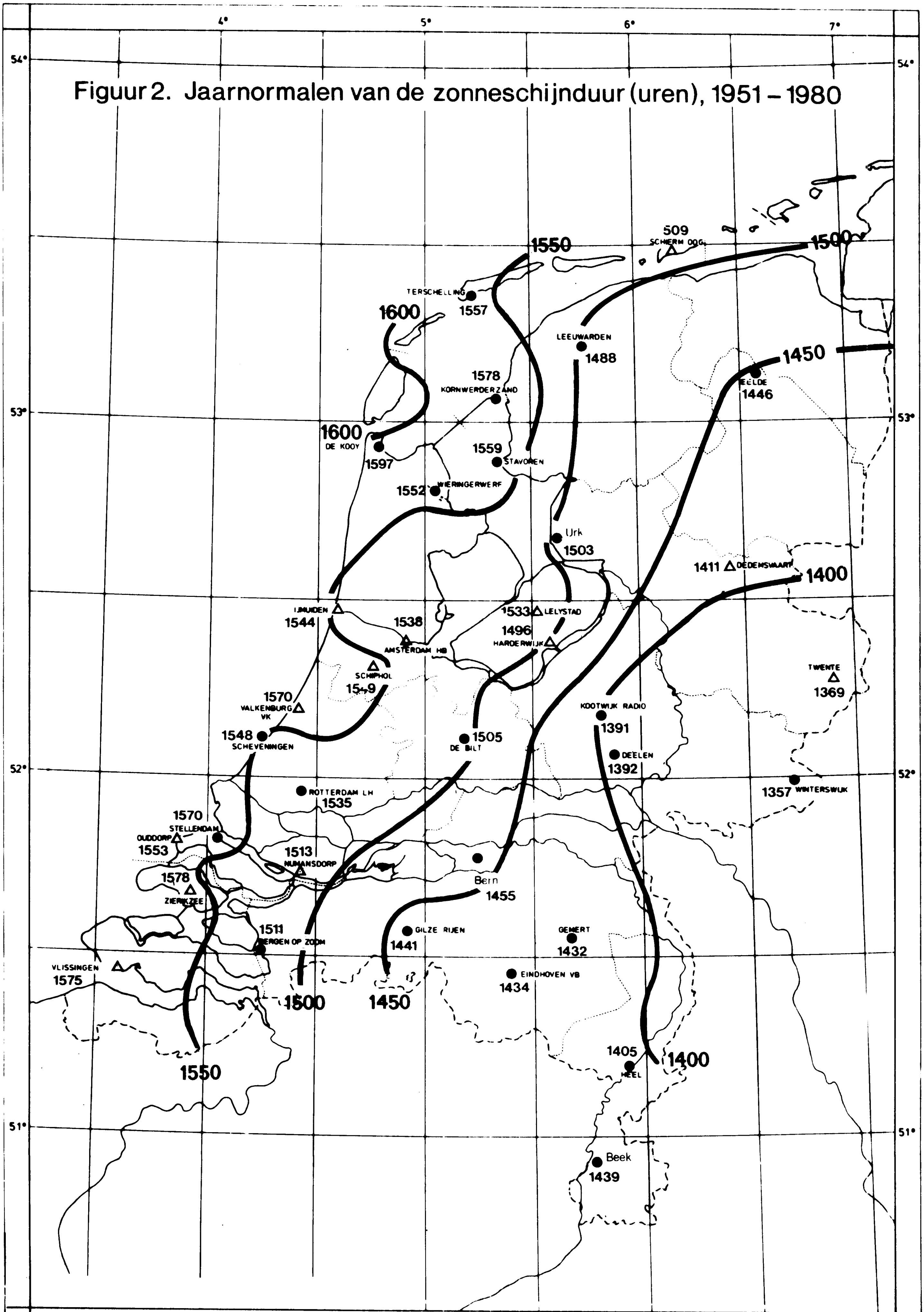
Tabel 1

Uitkomsten van toetsen op homogeniteit. De kruisjes achter de waarden van de toetsingsgrootheid geven een indicatie voor de mate van significantie van inhomogeniteiten (hoe meer kruisjes, hoe sterker significant, zie tekst). De score in de laatste kolom geeft het totaal aantal kruisjes.

Nr.	Stations	Q	R	W	score
8	Lelystad	0,484xxx	9,34xxx	2,30	6
33	Gemert	0,907	7,36	2,11	0
36	Numansdorp	0,491xxx	10,41xxx	5,31xxx	9
135	Kornwerderzand	0,900	7,44	3,41xx	2
153	Stavoren	0,912	5,95	3,04x	1
166	Schiermonnikoog	0,604xx	10,85xxx	3,34xx	7
210	Valkenburg ZH	0,485xxx	8,45xx	4,54xxx	8
235	De Kooy	0,970	6,70	1,86	0
240	Schiphol	0,568xxx	9,73xxx	4,58xxx	9
250	Terschelling	0,900	7,91x	2,49	1
260	De Bilt	0,661xx	6,68	2,77	2
270	Leeuwarden	0,778	7,32	3,04x	1
275	Deelen	0,867	4,74	1,69	0
280	Eelde	0,692xx	6,12	2,34	2
290	Twente VB	0,628xx	9,56xxx	4,53xxx	8
310	Vlissingen	0,455xxx	8,88xx	5,38xxx	8
325	Zierikzee	0,473xxx	8,31xx	2,93x	6
344	Rotterdam LH	0,791	7,47	2,60	0
350	Gilze-Rijen VB	0,935	4,97	1,45	0
370	Eindhoven VB	0,780	7,41	3,76xx	2
380	Beek LH	1.049	4,60	1,48	0
412	Amsterdam	0,283xxx	11,46xxx	7,32xxx	9
414	Bern	0,877	6,59	2,71	0
417	Stellendam	1,047	5,43	1,75	0
420	Harderwijk	0,670xx	9,51xxx	4,09xxx	8
421	Kootwijk	1,036	4,94	1,31	0
422	IJmuiden	0,518xxx	9,86xxx	4,71xxx	9
423	Heel	0,768x	6,19	1,54	1
427	Ouddorp	0,485xxx	9,69xxx	2,55	6
429	Scheveningen	0,791	6,84	3,04x	1
430	Urk	0,829	5,47	1,16	0
432	Dedemsvaart	0,633xx	6,09	4,00xxx	5
433	Winterswijk	0,769x	7,54	2,91x	2
438	Wieringerwerf	0,866	5,63	1,64	0
441	Bergen op Zoom	0,777	6,77	3,12x	1



Figuur 1. Ligging van zonneschijnstations met homogene (●) en niet homogene (△) reeksen.



Figuur 2. Jaarnormalen van de zonneschijnduur (uren), 1951 – 1980