KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT

WETENSCHAPPELIJK RAPPORT
SCIENTIFIC REPORT
81 - 7

T. A. Buishand

The analysis of homogeneity of long-term rainfall records in the Netherlands

De Bilt 1981

Publikatienummer: K.N.M.I. W.R. 81-7 (FM)

Koninklijk Nederlands Meteorologisch Instituut,
Fysisch Meteorologisch Onderzoek,
Postbus 201,
3730 AE De Bilt,
The Netherlands.

U.D.C.: 551.582(492) :
551.501.777 :
551.501.45

THE ANALYSIS OF HOMOGENEITY OF LONG-TERM RAINFALL RECORDS IN THE NETHERLANDS

T.A. Buishand

Abstract

For 24 long-term rainfall records (with an average length of 100 years) it was investigated how far the measured amount of rainfall was subjected to shifts in the mean. Lack of homogeneity was tested by the classical von Neumann ratio and the range of the rescaled cumulative deviations from the mean. The tests were applied to the annual amounts or to the annual differences between the rainfall amount of the station under consideration and the average amount of the other stations with long-term rainfall records.

There was strong evidence of departures from homogeneity. For each record a short description is given of the nature of possible changes in the measured amount of rainfall.

Contents

1. Introduction

Long, reliable rainfall records are often of great importance to make sensible decisions in hydrologic design. Though many long-term rainfall records are available now, it frequently occurs that they are not homogeneous. Non-homogeneities can be due to changes in the type or height of the raingauge, changes of observers or changes in the site of the rainfall station.

The possibility that a long-term rainfall record is not homogeneous makes its use less attractive. In some cases, however, the value of a rainfall record can be improved by making corrections for changes in the mean. A proper adjustment of rainfall records always requires a good knowledge of causes and magnitudes of departures from homogeneity. Usually an extensive statistical analysis is necessary to obtain a good insight into lack of homogeneity.

This report gives a detailed description of departures from homogeneity for 24 long-term rainfall records in the Netherlands. First, in Chapter 2 statistical methods for testing lack of homogeneity are explained. Then the results of the tests are given in Chapter 3. Finally, in Chapter 4, the nature of possible non-homogeneities is described for each station separately.

2. Statistical methods for evaluating the homogeneity of rainfall records

In this chapter some techniques are described to test the homogeneity of a sequence $Y_1, Y_2, \ldots, Y_n$. The sequence $\{Y_i\}$ is called stationary when its statistical properties do not change with time. This implies for instance that all Y_i's have the same distribution.

Long-term rainfall sequences usually consist of observations of the amount of rainfall over periods of 24 hours. Such sequences of daily rainfall amounts are not stationary because of seasonal variation of its statistical characteristics. But in a homogeneous rainfall record derived quantities like the consecutive annual amounts and the rainfall amounts for a particular calendar month in successive years form a stationary sequence. The homogeneity of a rainfall record is usually investigated by testing the stationarity of some derived sequence.

Tests for stationarity can be done for various statistical properties of the sequence. Often stationarity is tested for the mean only. The main reason for this is that the mean is less variable than quantities like the standard deviation, annual maxima, etc..

Two test-statistics for a stationary process are discussed in detail, namely the von Neumann ratio and a statistic based on the cumulative deviations from the mean. Both tests can only be used when the Y_i's are independent. This is not a serious restriction since homogeneity is usually tested on monthly, seasonal or annual amounts which are approximately uncorrelated. Though the distributions of the test-statistics are derived under the assumption that the Y_i's are normally distributed, the tests can still be applied in cases of slight departures from normality (Buishand, 1977).

2.1. The von Neumann ratio

In the climatological literature the use of the von Neumann ratio is often recommended as a test for homogeneity against unspecified alternatives (Sneijers, 1957; WMO, 1966). The test-statistic has the following form:

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

in which $\overline{Y}$ stands for the average of the Y_i's.

Under the null hypothesis of a constant mean it can be shown that $E(N) = 2$. For a non-homogeneous record the mean of N is usually smaller than 2. For instance, assume that there is a change of level:

$$E(Y_i) = \begin{cases} \mu, & i = 1,\ldots,m \\ \mu+\Delta, & i = m+1,\ldots,n. \end{cases} \qquad (2)$$

For this model it can be shown (Buishand, 1977):

$$E\{\sum_{i=1}^{n-1} (Y_i - Y_{i+1})^2\} = 2(n-1)\sigma_Y^2 + \Delta^2$$

and

$$E\{\sum_{i=1}^{n} (Y_i - \overline{Y})^2\} \approx (n-1)\{\sigma_Y^2 + q(1-q)\Delta^2\}$$

where $q = m/n$ and σ_Y^2 stands for the variance of the Y_i's. Therefore

$$E(N) \approx \frac{E\{\sum_{i=1}^{n-1} (Y_i - Y_{i+1})^2\}}{E\{\sum_{i=1}^{n} (Y_i - \overline{Y})^2\}}$$

$$\approx \frac{2}{1 + q(1-q)\ \Delta^2/\sigma_Y^2} \qquad (3)$$

provided that n, q and 1-q are not too small. From (3) one sees that $E(N) < 2$ if $\Delta \neq 0$.

Small values of the statistic N are an indication for a non-homogeneous record. For $n \gtrsim 20$ critical values for the test-statistic can be obtained from a normal approximation of the distribution of N under the null hypothesis:

$$N_p \approx 2 - 2u_p \sqrt{\frac{n-2}{(n-1)\ (n+1)}} \qquad (4)$$

with N_p : the p-th percentile of N,
u_p : the p-th percentile of a standard normal variate.

In the main part of this publication n = 73. For this value of n one gets for the critical value of a test at the 5% level:

$$N_{0.05} \approx 2 - 2 \times 1.645 \sqrt{\frac{71}{72 \times 74}} = 1.620 \ .$$

2.2. Cumulative sum tests

Cumulative sum techniques are very popular in the analysis of homogeneity (Craddock, 1979). The technique used in this study is based on the adjusted partial sums:

$$S_0^* = 0 \ ; \ S_k^* = \sum_{i=1}^{k} (Y_i - \overline{Y}), \qquad k=1,\dots,n. \qquad (5)$$

Note that $S_n^* = 0$. For a stationary sequence one may expect that the S_k^*'s fluctuate around zero since there is no systematic pattern in the deviations of the Y_i's from their average value $\overline{Y}$. Quite another picture arises when there is a change in the mean. For instance, assume that $\Delta < 0$ in Eq. (2). Then for $i \leq m$ the Y_i's are usually larger than $\overline{Y}$, while for $i > m$ the Y_i's tend to be smaller than $\overline{Y}$. This leads to positive values of the S_k^*'s. A typical example is given in Fig. 1. A similar argument yields that most values of S_k^* will be negative when $\Delta > 0$.

It is, however, not easy to see from a graph whether the cumulative deviations S_k^* differ significantly from those of a stationary random series. Since the successive S_k^*'s are highly correlated the fluctuations of the S_k^*'s are quite regular. Even for a sequence of random numbers it is known from the theory of random walks that the probability is quite large that nearly all S_k^*'s have the same sign. For this reason climatologists sometimes avoid the use of cumulative deviations (WMO, 1966). This is not necessary since it is possible to construct suitable test-statistics on the adjusted partial sums (Buishand, 1982).

A statistic which can be used is the rescaled adjusted range. The adjusted range is defined by (see Fig. 1):

$$\tilde{R} = \max_{0 \leq k \leq n} S_k^* - \min_{0 \leq k \leq n} S_k^* \qquad (6)$$

and the rescaled adjusted range is obtained by dividing $\tilde{R}$ by the sample standard deviation:

$$R = \tilde{R}/D_Y \qquad (7)$$

with

$$D_Y^2 = \sum_{i=1}^{n} (Y_i - \overline{Y})^2/n.$$

The statistic R is sensitive to shifts in the mean. High values for the test-statistic are an indication for a non-homogeneous record.

For a test on homogeneity one needs to know the distribution of R under te null hypothesis (that is under the assumption of a homogeneous record). Though it is possible to give an expression for the

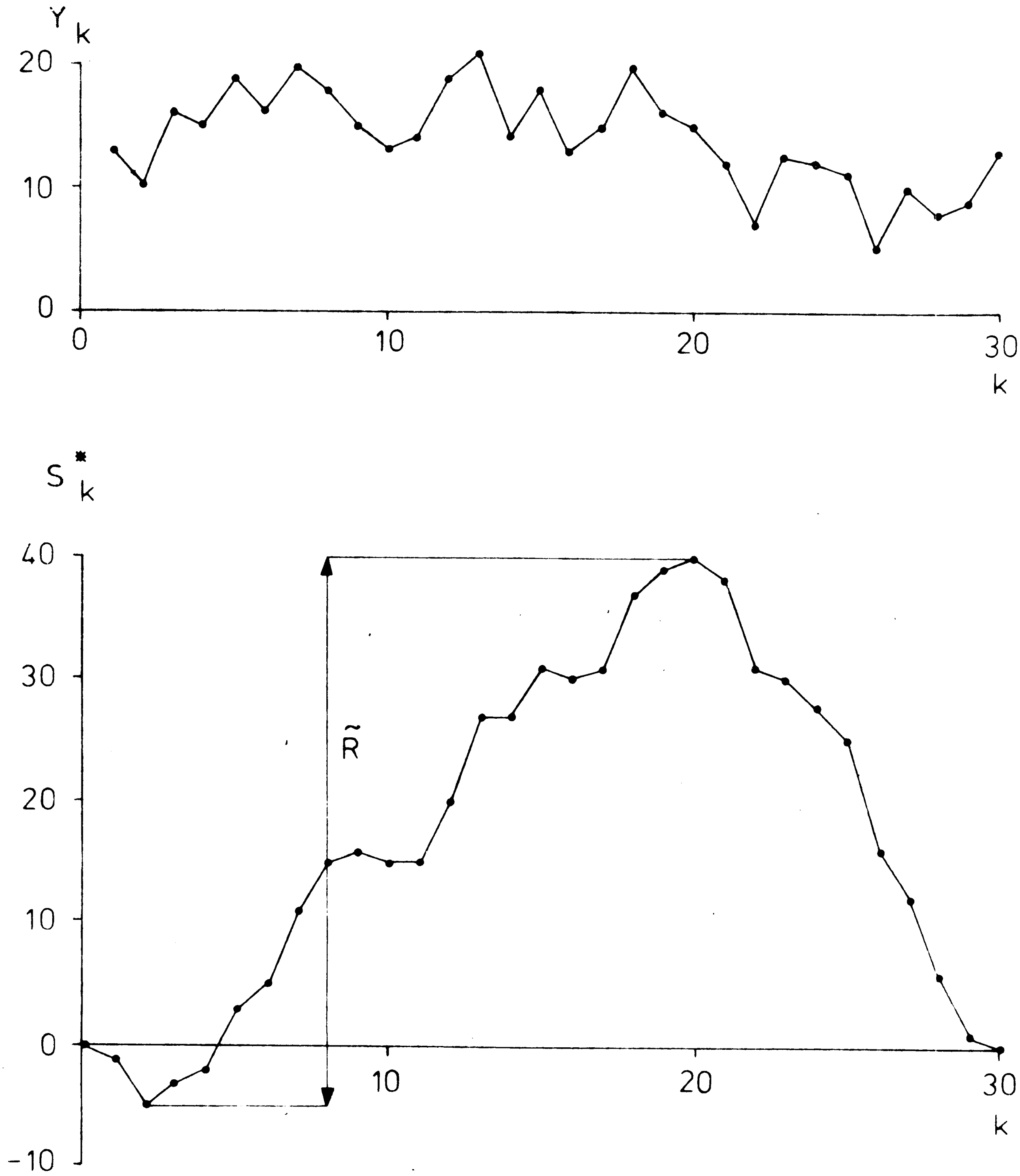

Fig. 1. Non-homogeneous time series with adjusted partial sums. $\tilde{R}$ denotes the adjusted range.

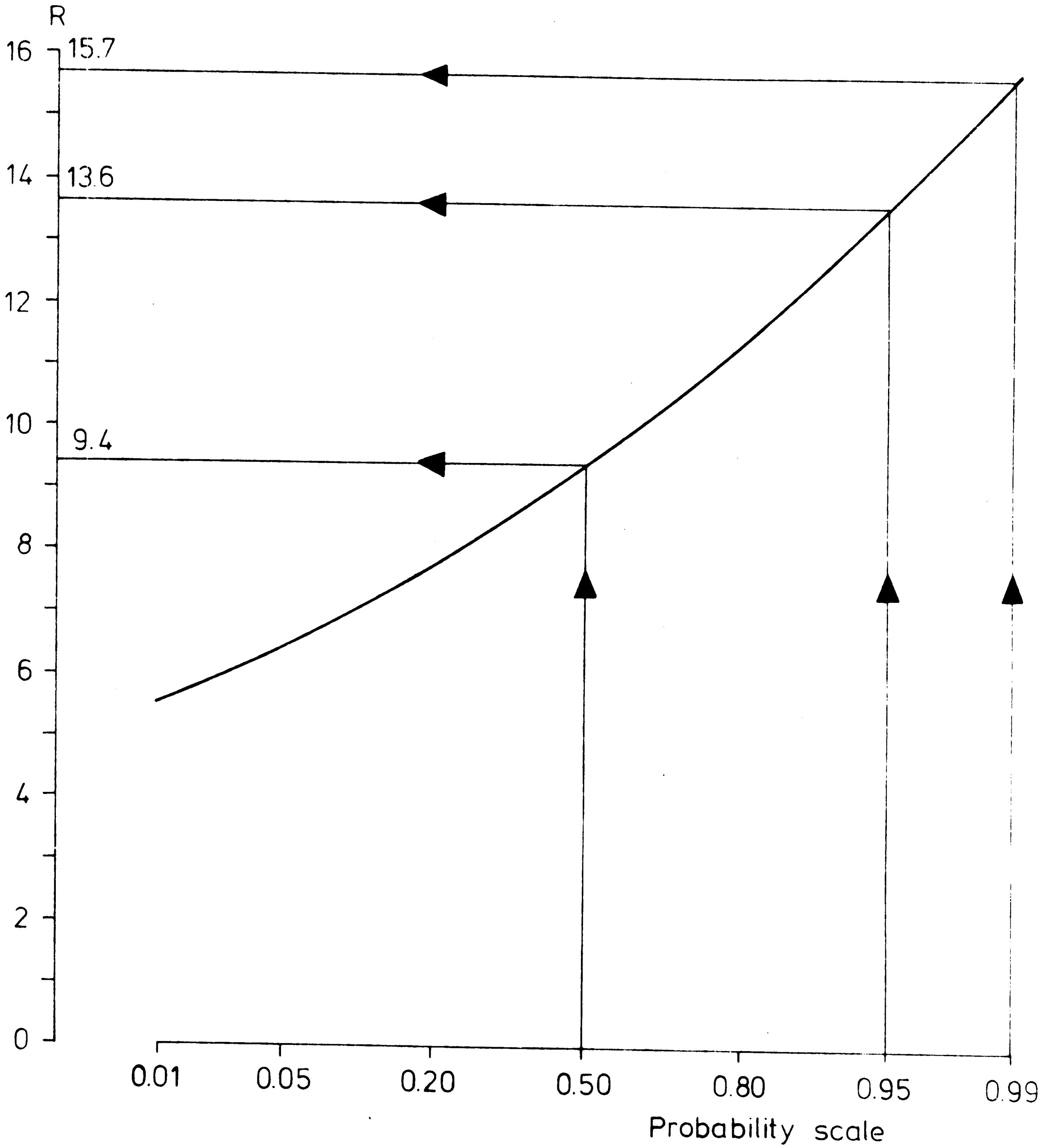

Fig. 2. Normal probability plot of the rescaled adjusted range R for sequences of length n = 73.

mean of R (Anis and Lloyd, 1976), it is not easy to derive the complete distribution by analytic methods. Percentage points of R are usually obtained by Monte Carlo methods. As an example Fig. 2 shows the distribution of R for n=73. The given distribution of R is based on 19,999 synthetic sequences of independent Gaussian random numbers. From Fig. 2 it is seen that for n=73 the median of R is about 9.4, the 95 and 99 percentage points are 13.6 and 15.7, respectively. For other values of n percentage points of R can be obtained from Wallis and O'Connell (1973) or Buishand (1982).

2.3. The power of tests on homogeneity

In the previous section two statistics for testing homogeneity were defined. The probability of detecting changes in the mean of a sequence $Y_1, Y_2, \ldots, Y_n$ by statistical methods depends on how serious these changes are. When only a small change occurs during a short period the probability is low that the null hypothesis of a constant mean is rejected. On the other hand, for feasible test-statistics it is required that they can discover relevant departures from homogeneity.

In this section the power of the statistics N and R is discussed in the situation that the Y_i's have a normal distribution with mean

$$E(Y_i) = \begin{cases} \mu & i = 1, \ldots, m_1 \\ \mu + \Delta & i = m_1 + 1, \ldots, m_2 \\ \mu & i = m_2 + 1, \ldots, n \end{cases} \qquad (8a)$$

and variance

$$\text{var } Y_i = \sigma_Y^2 \qquad i = 1, \ldots, n. \qquad (8b)$$

The model assumes that there is a change in the mean of magnitude Δ during $m_2 - m_1$ years, for instance due to another location of the raingauge site. The statistics N and R can be used to test the null hypothesis:

$$H_0 : \Delta = 0$$

against the alternative hypothesis:

$$H_1 : \Delta \neq 0.$$

When N is smaller than some critical value N_α then H_0 is rejected in favour of H_1; for the statistic R values greater than the critical value R_α lead to rejection of H_0. The magnitude of the critical values is determined by the significance level:

$$\alpha = \Pr(H_0 \text{ is rejected} | \Delta = 0).$$

In this section α is taken equal to 0.05. Then for n = 73 the critical values for the statistics N and R are 1.620 (Section 2.1) and 13.6 (Section 2.2) respectively.

For a particular test-statistic the probability of rejecting H_0 depends on the value of Δ, the standard deviation σ_Y, the number of observations n and the position of the change-points m_1 and m_2. The dependence on Δ and σ_Y can be combined into one single parameter $\Delta' = \Delta/\sigma_Y$.

Comparisons between test-statistics are usually based on their <u>power function</u>:

$$P(\Delta', m_1, m_2, n) = \Pr(H_0 \text{ is rejected} | \Delta'; m_1, m_2, n).$$

If Δ = 0 then $P(\Delta', m_1, m_2, n) = \alpha = 0.05$; for $\Delta \neq 0$ and m_1, m_2 and n fixed the power function increases monotonically with the absolute value of Δ'. When |Δ'| is growing then $P(\Delta', m_1, m_2, n)$ tends to 1, that is H_0 is rejected with probability 1.

Fig. 3 gives simulated power functions of N and R for n = 73 and two different combinations of m_1, and m_2, namely:

$$m_1 = 24;\ m_2 = 49 \rightarrow m_2 - m_1 = 25$$
$$m_1 = 31;\ m_2 = 41 \rightarrow m_2 - m_1 = 10.$$

Since the power functions are symmetric in Δ', non-negative values of Δ' are considered only. Each point on the curves in Fig. 3 is based on 1,999 sequences of Gaussian random numbers (the same set of pseudo uniform random numbers was used for the two test-statistics and for distinct values of Δ').

From the figure one reads that for Δ' = 1, $m_1 = 24$ and $m_2 = 49$ the probability of rejecting H_0 is 0.85 for the rescaled range R and 0.45 for the von Neumann ratio N. This means that R is a more powerful statistic than N for testing Δ = 0 against $\Delta \neq 0$. For $m_1 = 31$ and $m_2 = 41$

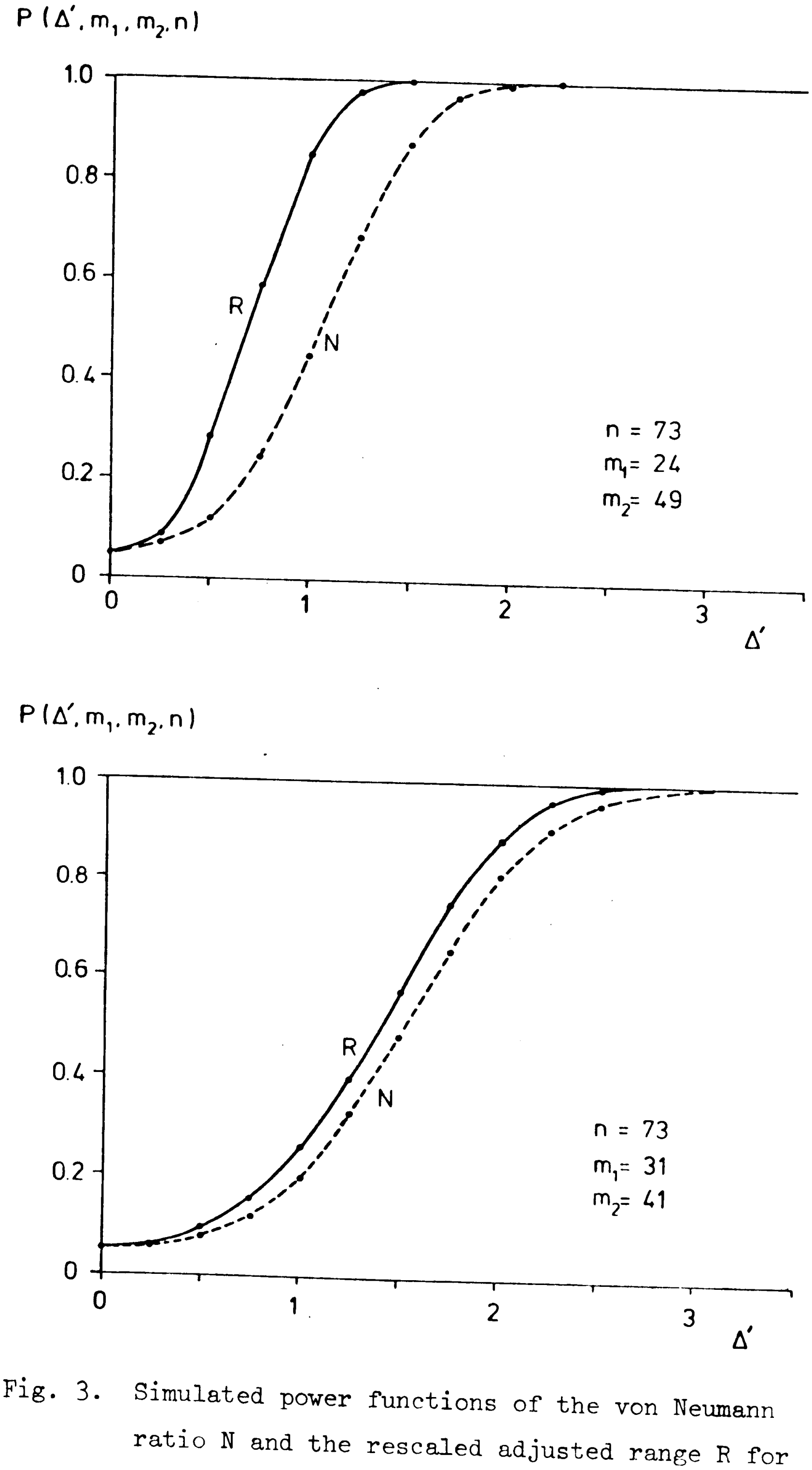

Fig. 3. Simulated power functions of the von Neumann ratio N and the rescaled adjusted range R for testing $\Delta = 0$ against $\Delta \neq 0$ in (8) at the 5% level.

there is only a small difference in power between R and N.

With respect to the position of the change-points m_1 and m_2 it can be shown that the shape of the power functions of N and R is mainly determined by $m_2 - m_1$. For instance, the power functions of the combinations $m_1 = 0$, $m_2 = 25$ and $m_1 = 24$, $m_2 = 49$ are nearly identical. Changes of level during a short period ($m_2 - m_1$ small) are often not discovered. For instance, if $m_2 - m_1 = 10$ it is seen from Fig. 3 that for $\Delta' = 1$ the probability of rejecting H_0 is only 0.25 when the statistic R is used. As was seen previously this probability is 0.85 if $m_2 - m_1 = 25$!

The rescaled range R is a suitable statistic in a situation that there is possibly some period, with unknown beginning and end, for which the mean of the Y_i's differs from the mean of other Y_i's. Therefore it is not surprising that in Fig. 3 the power function of R lies above the power function of N. For other departures from homogeneity N could be more powerful than R.

2.4. Estimation of the position of change-points

Graphs of cumulative deviations are often used to determine the position of change-points. It is then assumed that something has happended at points where the cumulative sum plot shows a clear change of slope. When dealing with adjusted partial sums a change-point may coincide with the position of a minimum or maximum of S_k^*.

Let M be te value of k for which S_k^* reaches its maximum:

$$S_M^* = \max_{0 \leq k \leq n} S_k^*$$

and K be the value of k for which S_k is minimal:

$$S_K^* = \min_{0 \leq k \leq n} S_k^* .$$

A small complication occurs when all S_k^*'s have the same sign for $k = 1, \ldots,$ $n - 1$. In this situation either M (negative S_k^*'s) or K (positive S_k^*'s) is set to zero. So M and K can take values $0, 1, \ldots, n - 1$.

Consider again Eq. (8a) and assume that Δ is positive. Then it might be expected that S_k^* reaches its minimum in the neighbourhood of m_1, and its maximum in the neighbourhood of m_2. That is K and M will be close to m_1 and m_2, respectively. The frequency distributions of K and M were investigated for the samples on which Fig. 3 was based. Only were those samples taken into account for which the null hypothesis was rejected at the 5% level. Let

$$P_M(k) = \Pr(M = k | H_0 \text{ is rejected}), \qquad k = 0, \ldots, n - 1$$

and

$$P_K(k) = \Pr(K = k | H_0 \text{ is rejected}), \qquad k = 0, \ldots, n - 1.$$

The shape of $P_M(k)$ and $P_K(k)$ depends on the value of the standardized shift $\Delta' = \Delta/\sigma_Y$. In the first instance it will be assumed that $\Delta' = 1$. For annual rainfall data in the Netherlands this value of Δ' corresponds to a change in the mean of about 7%, see Section 3.3.

From Fig. 4 it is seen that the mode of M and K always coincides with a change-point. For the probability that $K = m_1$ one reads:

$$P_K(m_1) = 0.29 \qquad (m_1 = 24, m_2 = 49)$$
$$P_K(m_1) = 0.10 \qquad (m_1 = 31, m_2 = 41)$$

and for the probability that $M = m_2$

$$P_M(m_2) = 0.27 \qquad (m_1 = 24, m_2 = 49)$$
$$P_M(m_2) = 0.12 \qquad (m_1 = 31, m_2 = 41).$$

In the situation that $m_1 = 31$ and $m_2 = 41$ there is only a small probability that M or K coincides with a change-point. Moreover, for this combination of m_1 and m_2 the positions of the minima and maxima of S_k^* are very dispersed. These facts are a consequence of the small number of observations with a different mean.

From Fig. 4 it is also noted that the distributions of M and K are not symmetric. For the statistic K the distribution is long-tailed to the left (negative skewness) whereas the distribution of M has a long right tail (positive skewness). Therefore cumulative sum plots may lead to a slight overestimation of the length of the period with a change of level.

For the distributions given in Fig. 4 the parameter Δ' was 1. If Δ' is larger than 1, then the values of M and K are more concentrated around the change-points. For instance if $\Delta' = 2$ the probability that

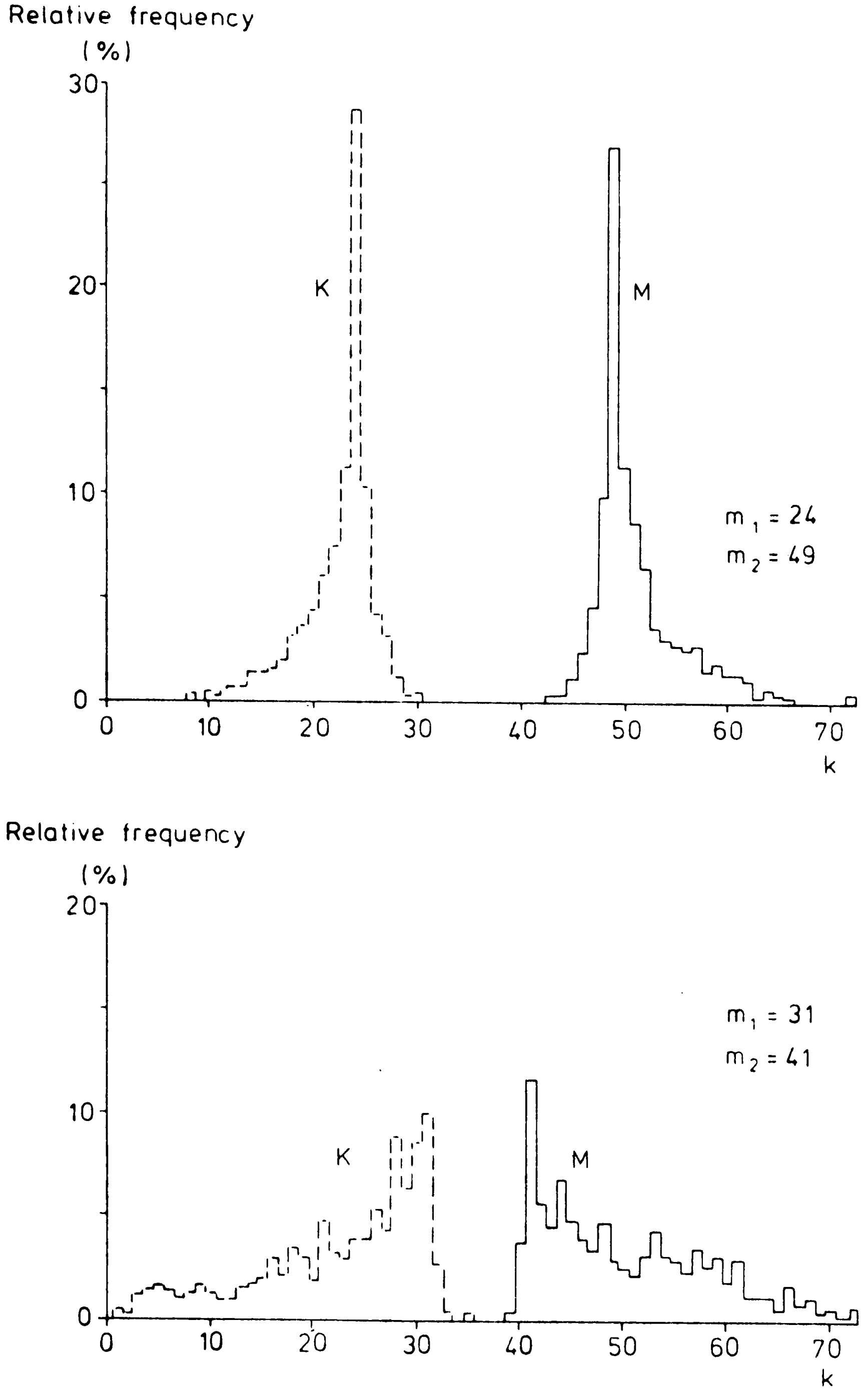

Fig. 4. Frequency distributions of estimates for the change-points m_1 and m_2 in (8) under the condition that the null hypothesis is rejected at the 5% level ($n = 73$, $\Delta' = 1$). M is the position of the maximum of the adjusted partial sums(estimate of m_2) and K is the position of the minimum (estimate of m_1).

$K = m_1 = 24$ or $M = m_2 = 49$ is nearly 0.60. This is about twice the value that is obtained when $\Delta' = 1$. For small values of Δ' the realizations of M and K are widely scattered. When there are only slight departures from homogeneity estimates of change-points are very poor.

A comment on maximum likelihood estimates

The likelihood ratio test for the situation of a single change-point, Eq. (2), is discussed by Worsley (1979) and Buishand (1982). For this model the maximum likelihood estimate of m coincides with the position of the maximum or minimum of the weighted adjusted partial sums:

$$Z_k^* = \frac{1}{\sqrt{k(n-k)}} S_k^*, \qquad k = 1, \ldots, n-1. \qquad (9)$$

Here a large weight is given to S_1^* and S_{n-1}^*; the weights are relatively small for k in the neighbourhood of n/2. Because of this weighting procedure good estimates of m are obtained in situations that the shift occurs near one of the end-points of the sequence (Buishand, 1982).

Weighted adjusted partial sums do not give better estimates of the change-points for the cases considered in Fig. 4. To obtain better estimates for the change-points in the situation that $m_1 = 31$ and $m_2 = 41$ it is necessary to give large weight to the S_k^*'s in the middle of the sequence, that is when k is in the interval (m_1, m_2). This does not occur when the Z_k^*'s are used.

The maximum likelihood estimates of the change-points m_1 and m_2 in Eq. (8) are the values of k_1 and k_2 for which

$$F(k_1, k_2) = \frac{|S_{k_1}^* - S_{k_2}^*|}{\sqrt{\{k_2 - k_1\}\{n - (k_2 - k_1)\}}} \qquad (10)$$

reaches its maximum $(0 \leq k_1 < k_2 \leq n)$. The derivation of (10) follows the same lines as in the situation of one change-point.

In contrast with the rescaled adjusted range the method of maximum likelihood is sensitive to departures from normality. For instance, assume that the j - th observation Y_j is an outlier. Then for $k_1 = j - 1$, $k_2 = j$:

$$|S^*_{k_1} - S^*_{k_2}| = |Y_j - \bar{Y}| \text{ is large.}$$

But also

$$\sqrt{\{k_2 - k_1\}\{n - (k_2 - k_1)\}} \text{ is minimal.}$$

Consequently $F(j - 1, j)$ is large and it may occur that due to the single outlier Y_j the null hypothesis of a constant mean is rejected.

Let K_1 and K_2 be the values of k_1 and k_2 for which $F(k_1,k_2)$ reaches its maximum. Fig. 5 gives the distribution of the "maximum likelihood" estimates K_1 and K_2 in the situation that $m_1 = 31$ and $m_2 = 41$. To reduce the influence of outliers the maximization was done under the restriction that $3 \leqslant k_1 \leqslant k_2 - 3$ and $k_2 \leqslant n - 3$.

Fig. 5 is based on the same set of random numbers as Fig. 4. Again the empirical distributions of the estimates for the change-points were derived from samples for which the null hypothesis was rejected at the 5% level.

From a comparison of Figs. 4 and 5 one sees that for $m_1 = 31$ and $m_2 = 41$ the statistics K_1 and K_2 are superior to M and K. The weights in Eq. (10) are such that frequently a change-point is found in the interval (31,41) which leads to a considerable reduction of bias.

It is rather easy to determine the position of the maximum and minimum of the adjusted partial sums. To obtain maximum likelihood estimates of the change-points m_1 and m_2 it is necessary to maximize a function of two discrete variables (k_1 and k_2) which requires much more computational labour. Moreover, as a test for $\Delta = 0$ against $\Delta \neq 0$ the likelihood ratio test is not always more powerful than the test based on the rescaled adjusted range R. For instance, in the case that $m_1 = 24$ and $m_2 = 49$ the statistic R is superior to the likelihood ratio statistic. Because of these disadvantages the method of maximum likelihood is not considered further in this paper.

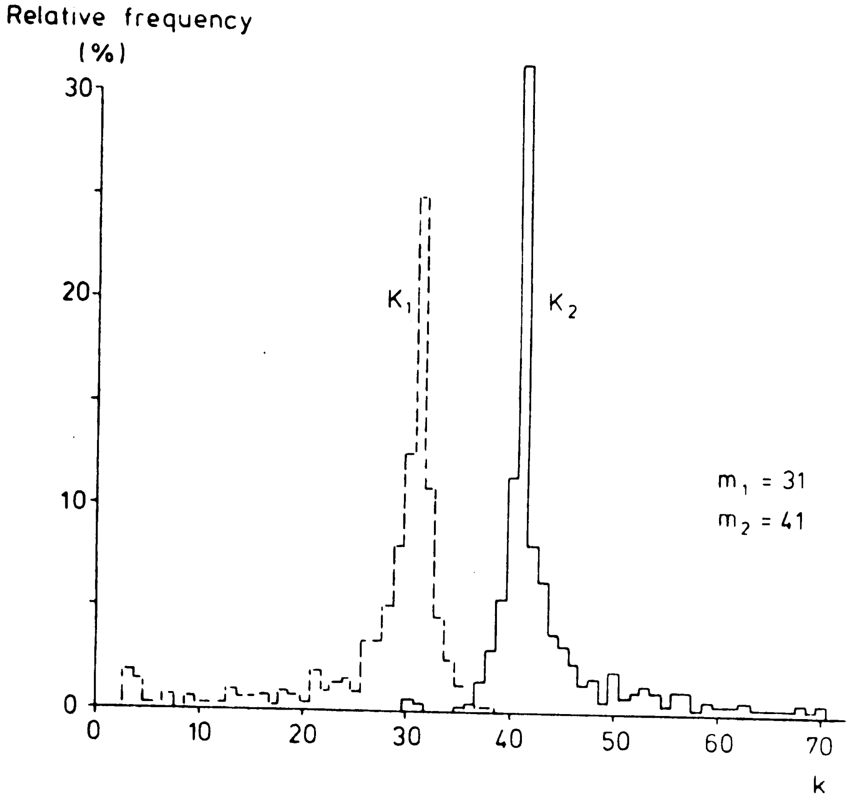

Fig. 5. Frequency distributions of "maximum likelihood" estimates for the change-points m_1 and m_2 in (8) under the condition that the null hypothesis is rejected at the 5% level ($n = 73$, $\Delta' = 1$). K_1 is an estimate of m_1 and K_2 is an estimate of m_2.

3. Results of testing homogeneity

In the Netherlands a rather dense network of raingauges has been in operation since the second half of the 19-th century. In the beginning various types of instruments were in use and there was no uniform height of the rim of the raingauge. During the first decade of this century a standardgauge was introduced. This gauge had an orifice of 400 cm^2 and had its rim at a height of 1.50 m. About 1950 the height of the rim was lowered to 0.40 m to reduce the aerodynamic error. Since 1962 a new type of raingauge has been introduced with an orifice of 200 cm^2.

Daily rainfall data on magnetic tape are available for all stations from 1951 onwards. For the period prior to 1951 only the data of 24 selected stations were put on magnetic tape. The homogeneity of the records from these 24 stations was examined with the techniques discussed in the previous section.

Fig. 6 gives the geographical location of the stations with long-term rainfall records on magnetic tape. The records have an average length of 100 years. The beginning date of the observations ranges from January 1847 (Groningen) to October 1911 (Dwingeloo). Rainfall observations in Heusden and Utrecht were terminated in January 1942 and July 1959, respectively. The other stations in Fig. 6 are still operational now.

The rainfall record of Heusden was supplemented with rainfall observations from the nearby station Andel. No attempts were made to supplement interruptions of other records.

3.1. Application of the von Neumann ratio to annual amounts

In a flat country like the Netherlands there are only small local differences in the rainfall regime. For the annual amounts the mean and standard deviation are usually close to 750 and 120 mm, respectively. This means that for a change in the mean of 10% the standardized shift Δ' is about 0.6. From Fig. 3 it is seen that for this value of Δ' the probability of rejecting H_0 is usually less than 0.5. In general it is not possible to get more powerful tests by taking the consecutive rainfall amounts of a particular month or season instead of the annual amounts (Mitchell, 1961; Buishand, 1977).

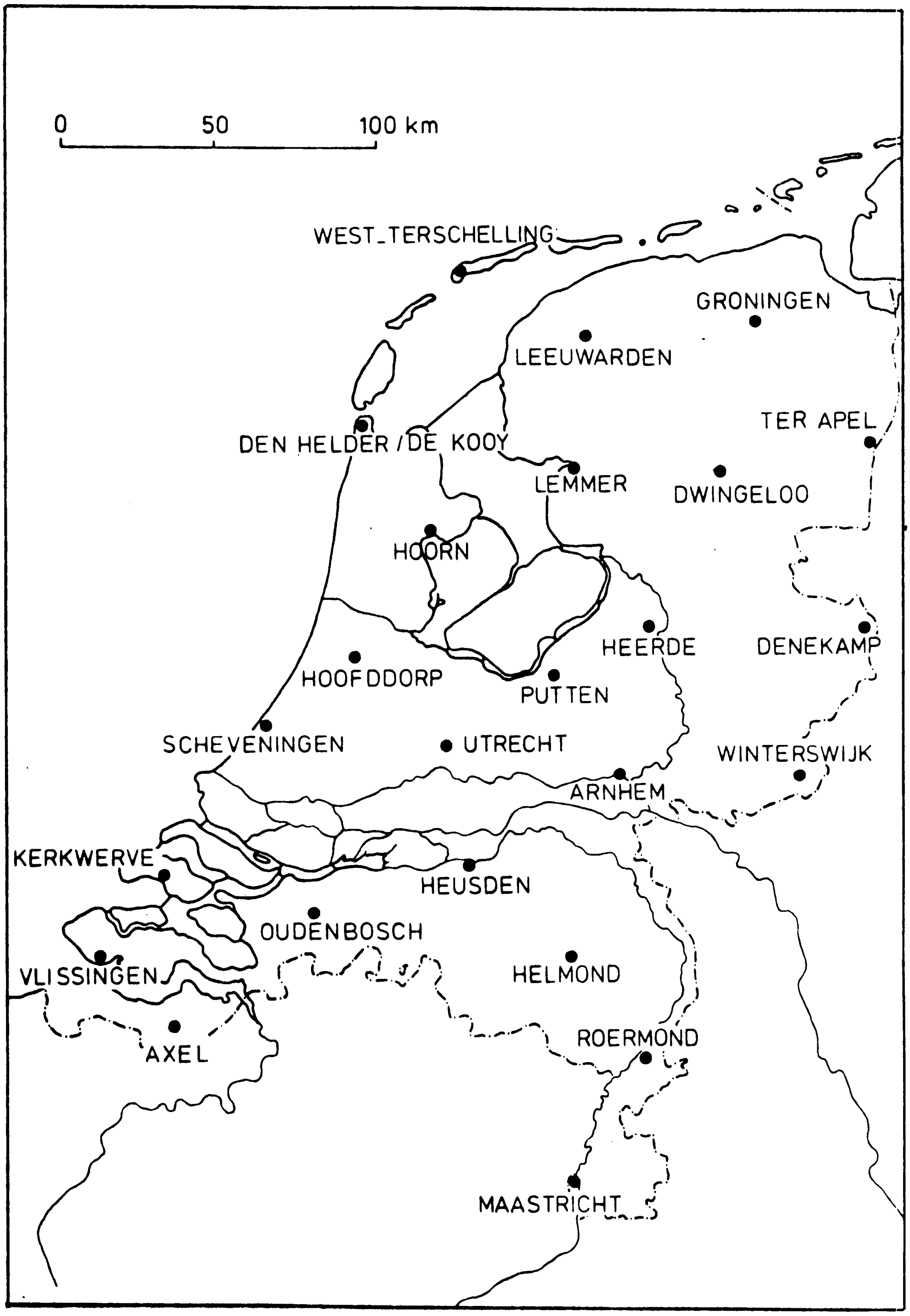

Fig. 6. Geographical location of stations used in the analysis of homogeneity.

Table 1

Realizations of the von Neumann ratio N for testing lack of homogeneity for sequences of annual amounts.

Station	Number of years	N		Station	Number of years	N
Den Helder/De Kooy	129	1.57**		Utrecht	87	1.93
West-Terschelling	104	1.90		Arnhem	85	1.63*
Lemmer	100	1.31**		Putten	112	1.59*
Leeuwarden	104	1.72		Winterswijk	99	1.90
Groningen	133	1.68*		Vlissingen	125	1.04**
Ter Apel	88	1.96		Kerkwerve	101	1.58*
Dwingeloo	68	1.98		Axel	75	1.97
Heerde	87	2.00		Heusden/Andel	108	1.82
Denekamp	75	1.79		Oudenbosch	91	1.64*
Hoorn	96	1.58*		Helmond	102	1.92
Hoofddorp	113	2.01		Roermond	111	1.71
Scheveningen	103	1.43**		Maastricht	127	1.25**

** Significant at the 1% level.

* Significant at the 5% level, but not at the 1% level.

As a preliminary test the von Neumann ratio was applied to the sequences of annual amounts. The results are given in Table 1. Despite the low power of the statistic N there are strong indications for departures from homogeneity. For nearly one half of the stations the null hypothesis is rejected at the 5% level. For Vlissingen the von Neumann ratio is only 1.04 whereas for Maastricht a value of 1.25 is found. These low values are partly due to the anomalous position of the raingauge in the 19-th century, see Chapter 4.

3.2. The use of year-by-year differences

In the previous section the von Neumann ratio was applied directly to the sequence of consecutive annual amounts. A better test can be obtained by taking the differences between the annual amount of the station under consideration and the average annual amount of a number of surrounding stations. That is for the p-th station the sequence $\{Y_i\}$ has the form

$$Y_i = X_{pi} - A_{pi} \ , \qquad i = 1, \ldots, n \qquad (11)$$

with X_{pi}: amount of rainfall during year i at station p,

A_{pi}: average amount of rainfall for year i at a number of surrounding stations.

Since there exists a strong positive correlation between X_{pi} and A_{pi} the variance σ_Y^2 of Y_i is smaller than the variance σ_X^2 of X_{pi}. A change of level of magnitude Δ in the X_{pi}'s also implies a change of level of the same magnitude in the Y_i's. But since

$$\Delta_Y' = \Delta/\sigma_Y > \Delta/\sigma_X = \Delta_X'$$

the Y_i's are preferred for testing $\Delta = 0$ against $\Delta \neq 0$.

In this study the von Neumann ratio N and the rescaled adjusted range R were applied to sequences of year-by-year differences for the period 1907-1979. For a particular rainfall station in Fig. 6 with a complete record $X_{p1}, \ldots, X_{p,73}$ during this period, the sequence $\{A_{pi}\}$ was constructed as follows:

$$A_{pi} = \sum_{\substack{q=1 \\ q \neq p}}^{22} X_{qi}/21, \qquad i = 1, \ldots, 73. \qquad (12)$$

Table 2

Realizations of the von Neumann ratio N and the rescaled adjusted range R for testing lack of homogeneity.

Station	N	R	Station	N	R
Den Helder/De Kooy	1.88	14.1*	Arnhem	1.07**	19.1**
West-Terschelling	1.56*	13.1	Putten	1.46**	14.6*
Lemmer	1.43**	17.3**	Winterswijk	1.60*	13.2
Leeuwarden	1.71	11.4	Vlissingen	1.60*	16.9**
Groningen	1.78	7.6	Kerkwerve	1.75	11.9
Ter Apel	2.19	10.5	Axel	1.80	14.9*
Dwingeloo	1.49*	16.3**	Heusden/Andel	1.72	12.5
Heerde	2.02	14.7*	Oudenbosch	1.81	8.2
Denekamp	1.81	15.8**	Helmond	1.57*	20.6**
Hoorn	2.12	9.5	Roermond	2.15	8.9
Hoofddorp	1.70	10.0	Maastricht	1.62	19.6**
Scheveningen	1.75	18.0**			

The tests were applied to the sequence of year-by-year differences, Eqs. (11) and (12), for the period 1907-1979 (Dwingeloo 1912-1979).

** Significant at the 1% level.

* Significant at the 5% level, but not at the 1% level.

Here the upper bound of the summation is 22 instead of 24 since Dwingeloo and Utrecht were not in operation during the whole period 1907-1979. The sequence of Utrecht was not considered further; for Dwingeloo the sequence $\{A_{pi}\}$ consisted of the average annual amounts of the 22 stations with complete records during the period 1907-1979.

Table 2 gives the realizations of the von Neumann ratio N and the rescaled adjusted range R for the period 1907-1979. The cumulative deviations from the mean for this period are given in Fig. 7. From Table 2 it is seen that for many stations the null hypothesis of a constant mean is rejected. The most serious departures from homogeneity occur in the rainfall data of Lemmer, Dwingeloo, Arnhem, Putten, Vlissingen, Helmond and Maastricht. For these stations there are marked changes in the slope of the cumulative sum plots. The nature and causes of these changes are discussed in Chapter 4.

3.3. Concluding remarks

The year-by-year differences have a standard deviation of about 55 mm. This means that for a change in the mean of 7% the standardized shift Δ' is $0.07 \times 750/55 \approx 1$. For this value of Δ' the power of the statistics R and N was discussed in Section 2.3. In the situation that $m_1 = 31$ and $m_2 = 41$ (change in the mean during 10 years only) the probability of rejecting H_0 is less than 0.30. Moreover, for such short periods with a different mean the indices of the minima and maxima of the adjusted partial sums give very poor estimates for the positions of the change-points.

The tests on the year-by-year differences were applied to data for the period 1907-1979 only, since serious problems were met in constructing a good sequence $\{A_{pi}\}$ for the period prior to 1907. One reason for these difficulties is that the observations of the stations in Fig. 6 have unequal starting dates. Especially during the second half of the previous century there is a large variation in the number of operational stations. Further, for a number of stations the average amount of rainfall over the past 70 years is much higher than the average amount over earlier periods. This is illustrated in Table 3 where the 1881-1906 annual average is compared with the 1907-1979

Table 3

Average annual amounts and realizations of Student's t-statistic for a two-sided test on a difference in mean between the periods 1881-1906 and 1907-1979.

Station	Mean (mm) 1881-1906	Mean (mm) 1907-1979	Difference (mm)	t
Den Helder/De Kooy	681	700	19	0.73
West-Terschelling	699	739	40	1.50
Lemmer	654	732	78	2.68**
Leeuwarden	748	737	-10	-0.43
Groningen	696	767	71	3.05**
Hoofddorp	770	760	-10	-0.36
Scheveningen	645	764	119	4.18**
Putten	689	788	99	3.26**
Winterswijk	745	759	14	0.47
Vlissingen	546	715	168	5.96**
Kerkwerve	621	701	80	3.14**
Heusden/Andel	684	723	39	1.46
Helmond	652	732	81	3.04**
Roermond	617	657	40	1.51
Maastricht	598	679	81	2.71**

** Significant at the 1% level.

* Significant at the 5% level, but not at the 1% level.

annual average for all stations in Fig. 6 with complete records since 1881. From the realizations of Student's t-statistic in Table 3 one sees that for 8 of the 15 stations there is strong evidence of a difference in mean. For these stations the annual mean of the most recent period is on average nearly 100 mm higher than the mean of the earlier period. This large difference in the mean annual amount for about one half of the stations gives rise to non-homogeneous sequences of areal averages $\{A_{pi}\}$ over the whole period 1881-1979.

An interesting question is how far changes in the mean of the X_{pi}'s may influence the homogeneity of the sequence of averages $\{A_{pi}\}$. If these changes occur at random, then in general the sequence $\{A_{pi}\}$ is still approximately a stationary sequence, since it is based on a large number of stations (Mitchell, 1961). Though changes in the raingauge site are usually distributed at random in time and amoung stations, this is not the case for the change in height of the raingauge which took place around 1950. Because of this country-wide change in the height of the raingauge the sequence $\{A_{pi}\}$ is non-homogeneous, but it is questionable whether it has a serious effect on the results of the statistical tests. If the change of height of the raingauge has the same effect on the average amount of rainfall at all stations, then the sequence of year-by-year differences $\{Y_i\}$ should still be homogeneous. In reality, the effect of lowering the raingauge differs from station to station due to local differences in the wind-climate and the degree of sheltering of the raingauge. Estimates of the magnitude of the wind-effect for all stations in the Netherlands were given by Braak (1945). From these estimates it can be deduced that the change in height of the raingauges has resulted in a shift in the mean which ranges from nearly 0 to 0.8 σ_Y for the stations in Fig. 6. These values are too small for a serious effect on the statistical tests in the previous section.

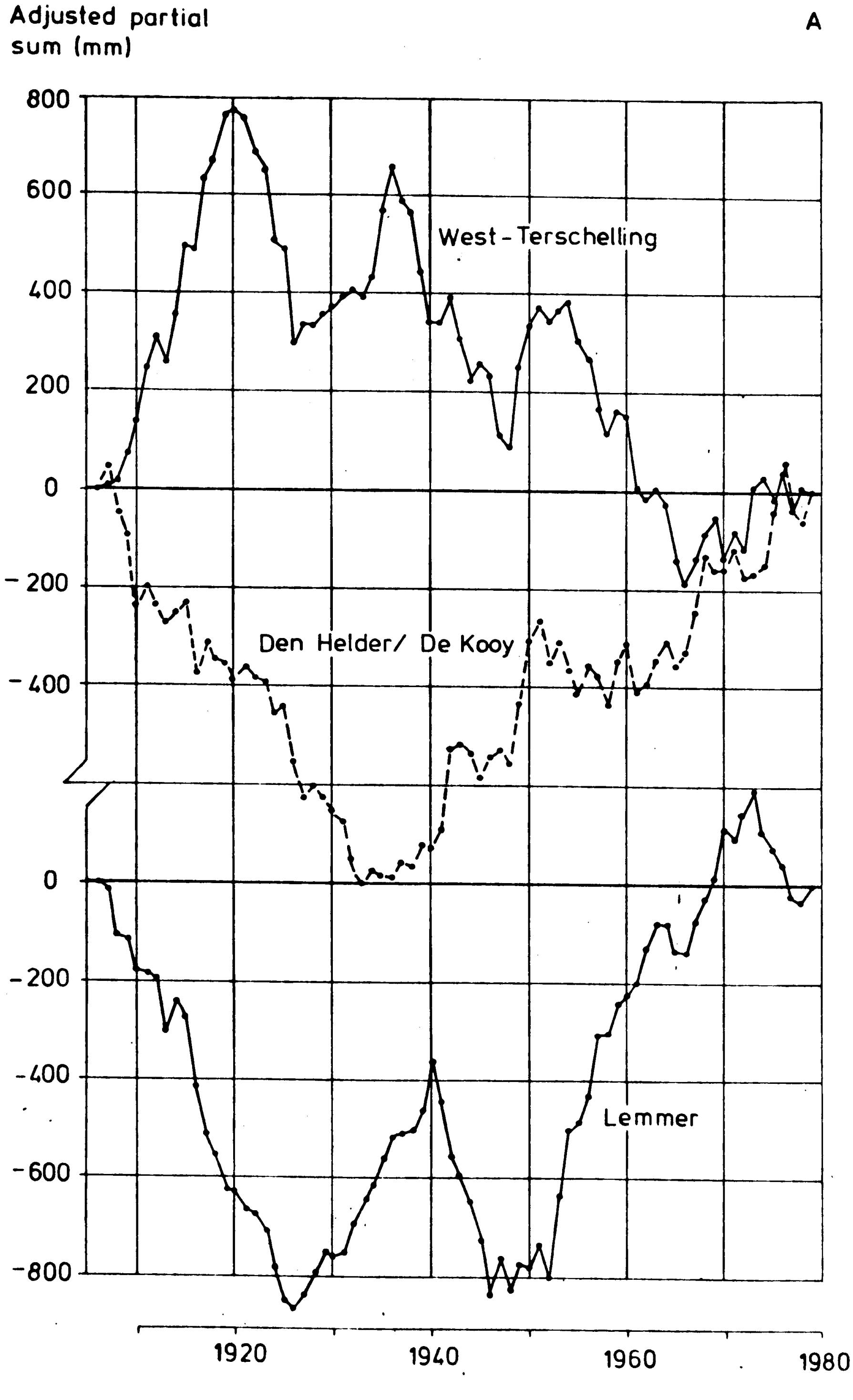

Fig. 7. Adjusted partial sums of year-by-year differences, Eqs. (11) and (12), for the period 1907-1979.

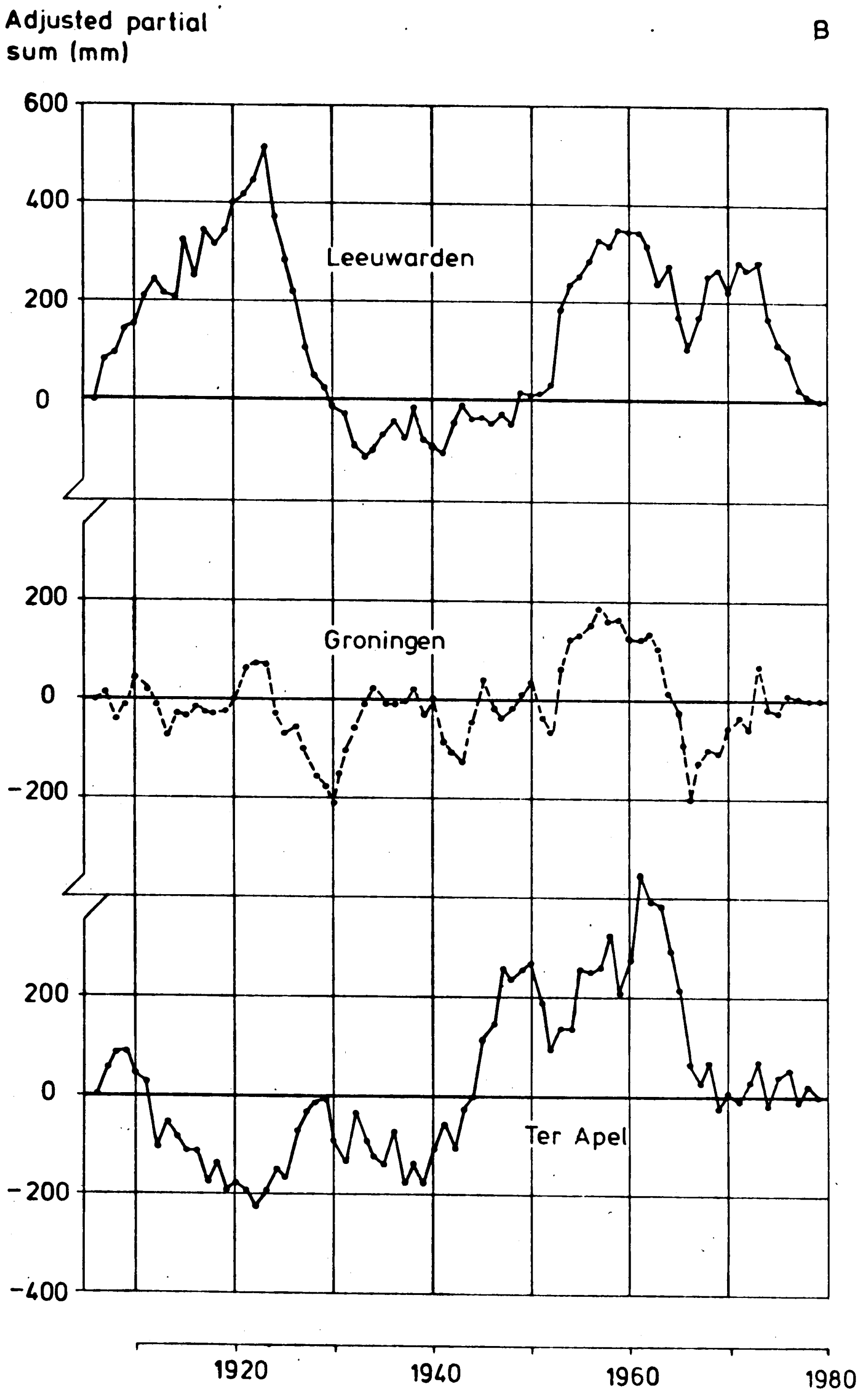

Fig. 7. (Continued)

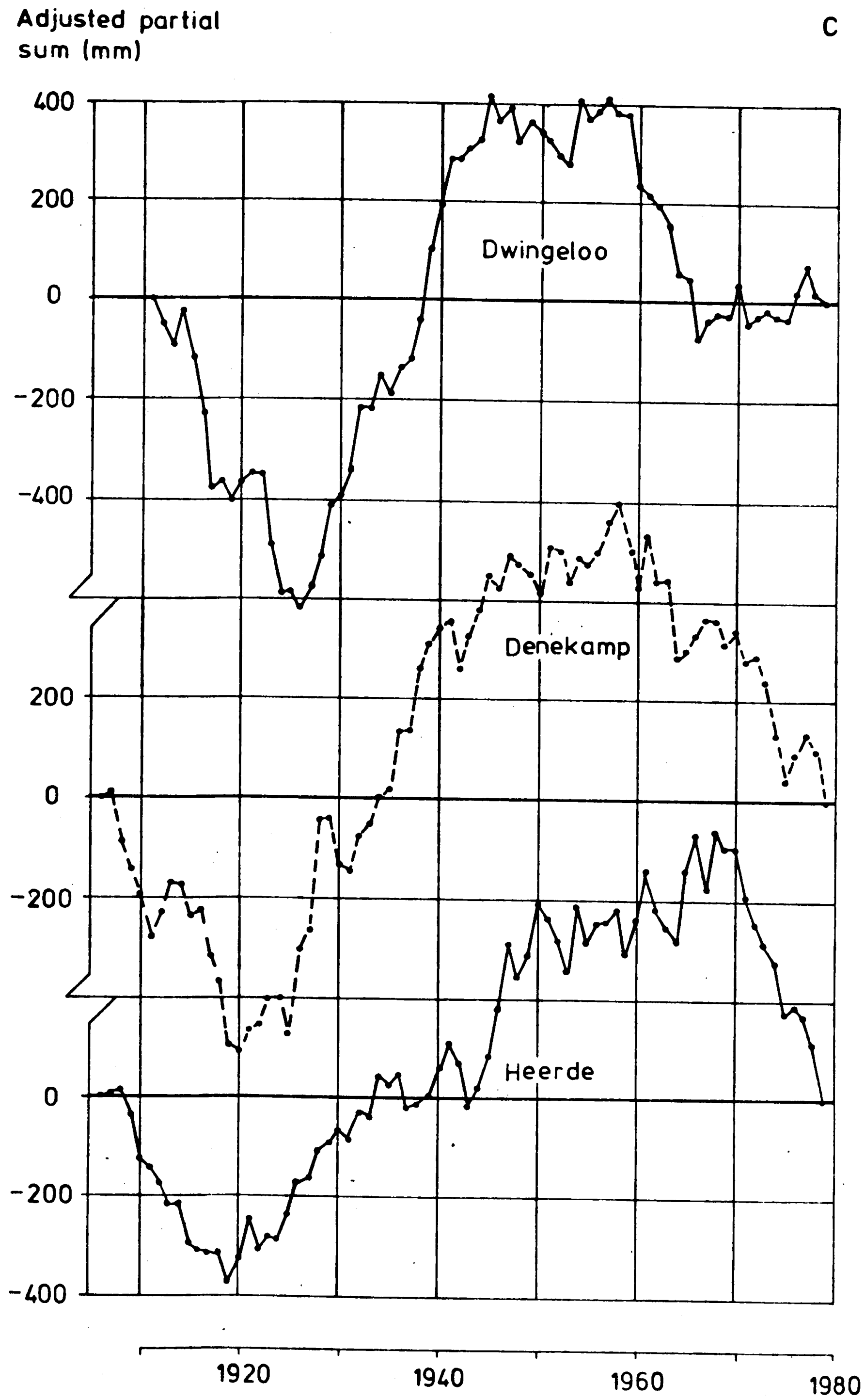

Fig. 7. (Continued)

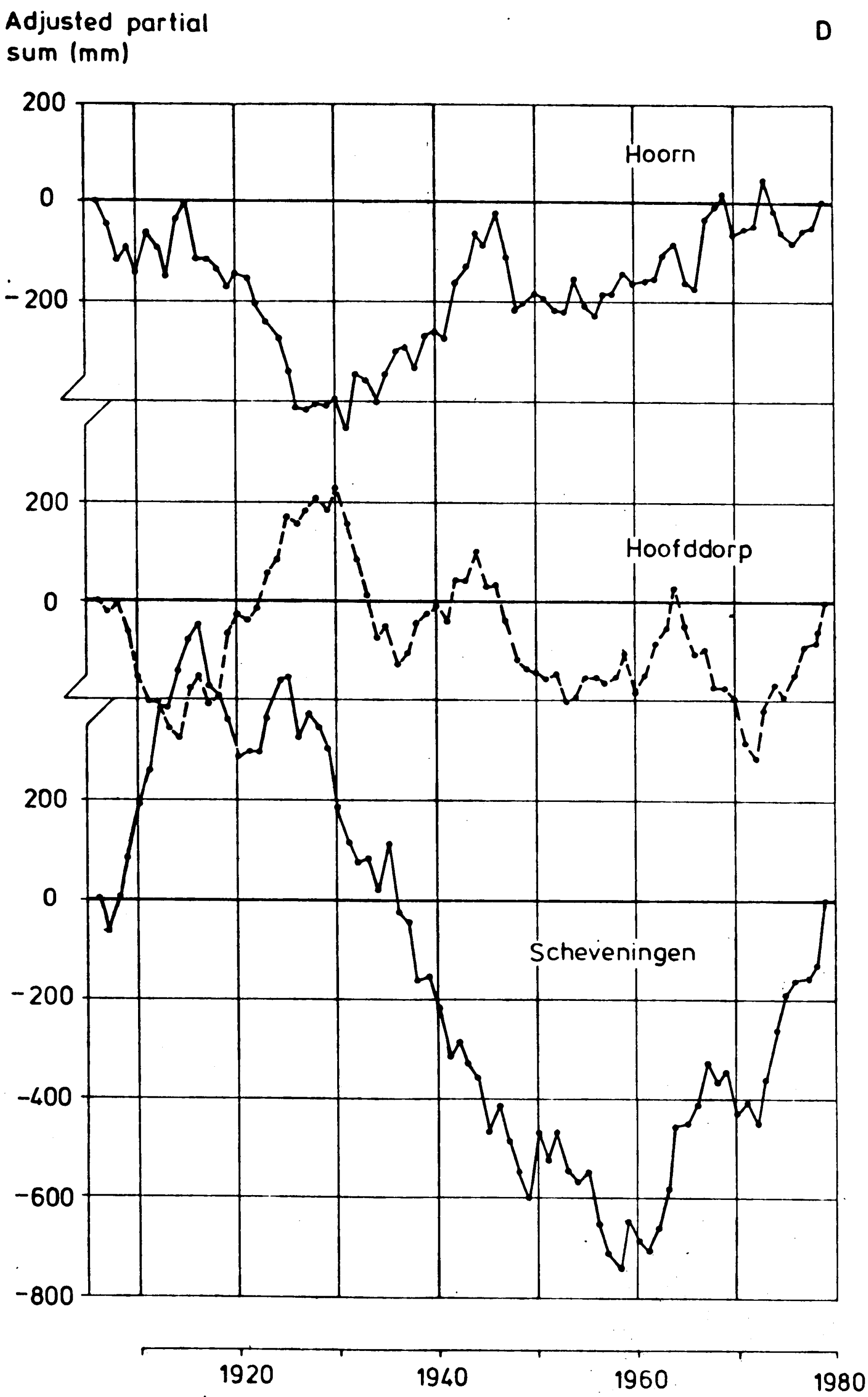

Fig. 7. (Continued)

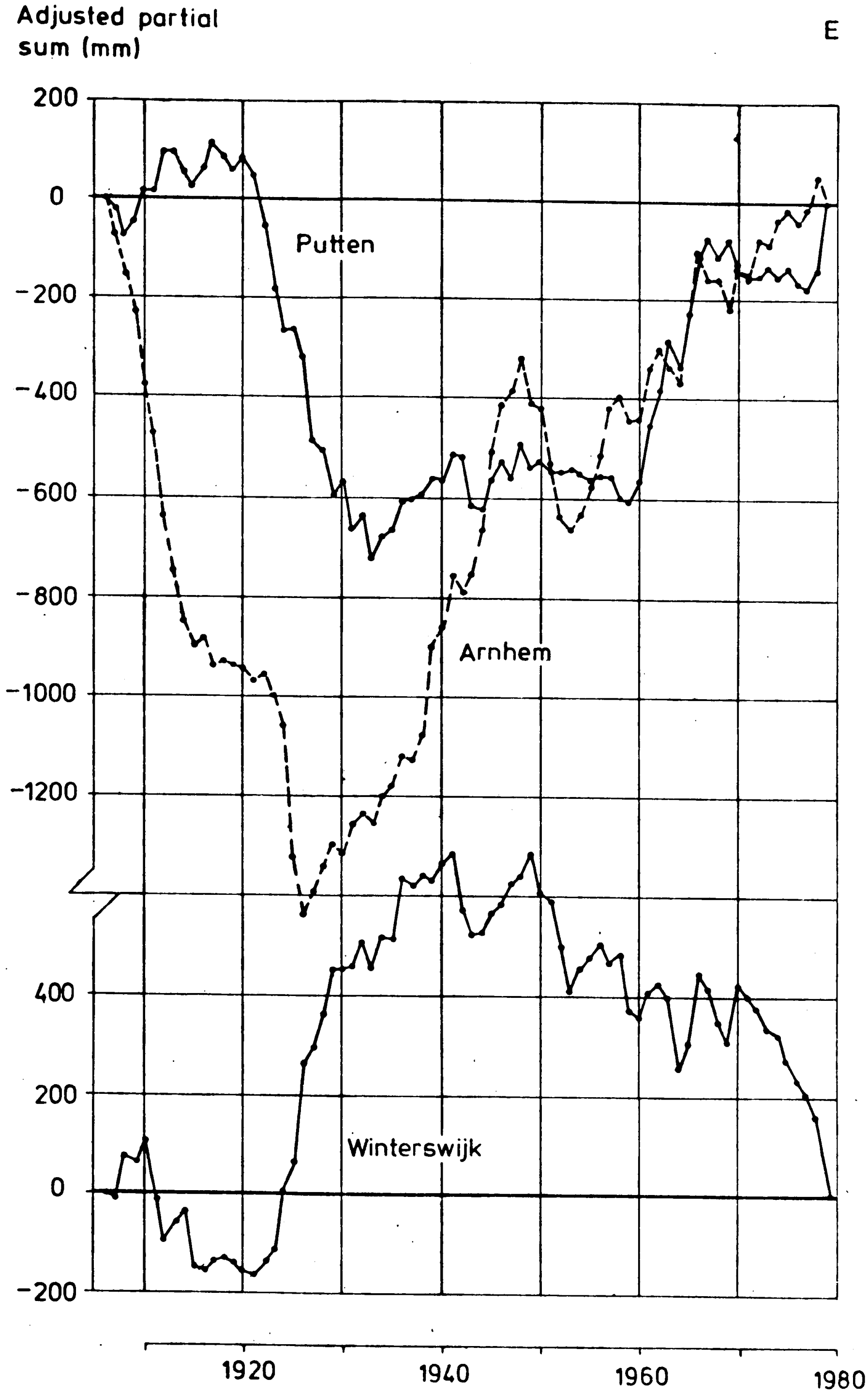

Fig. 7. (Continued)

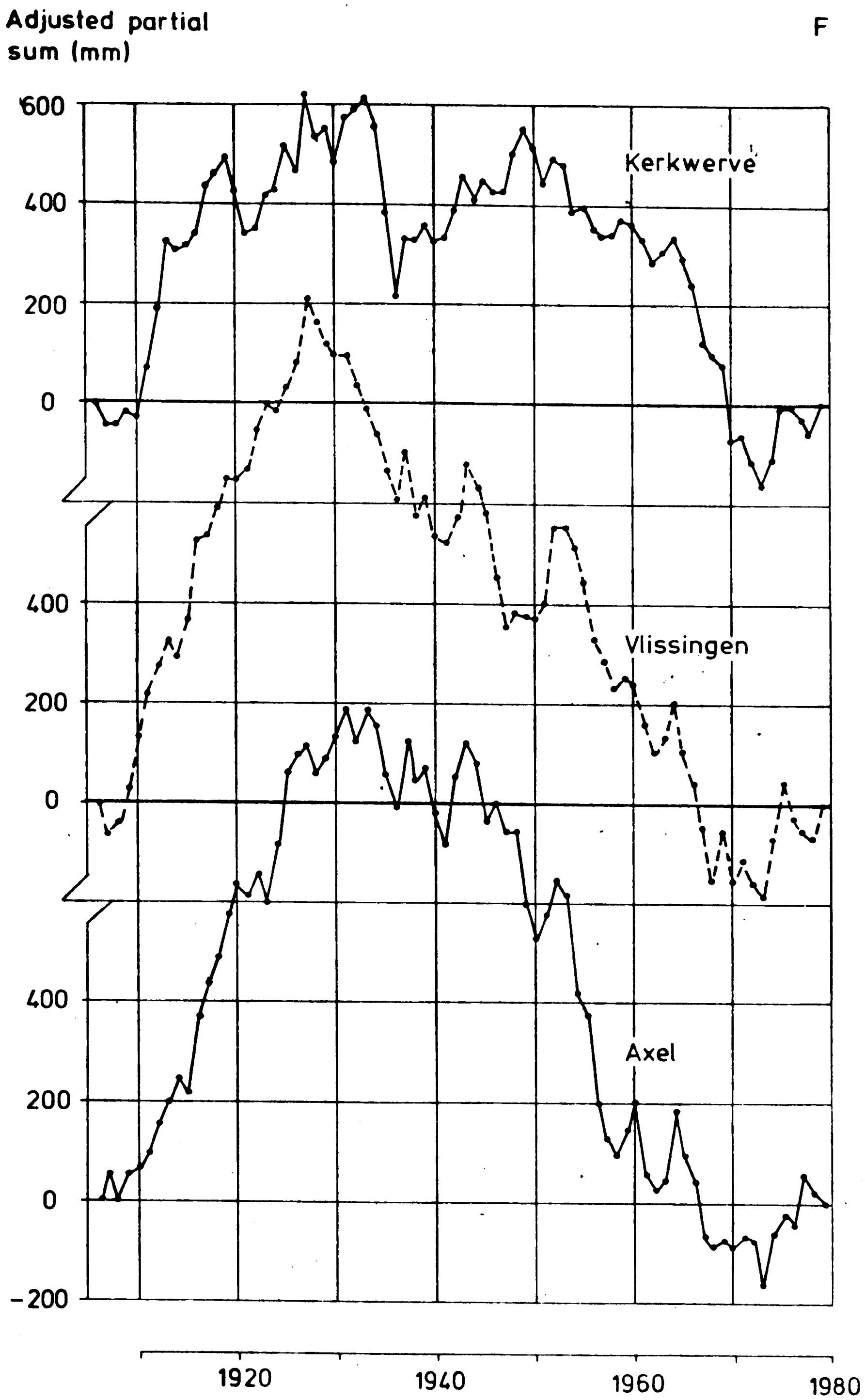

Fig. 7. (Continued)

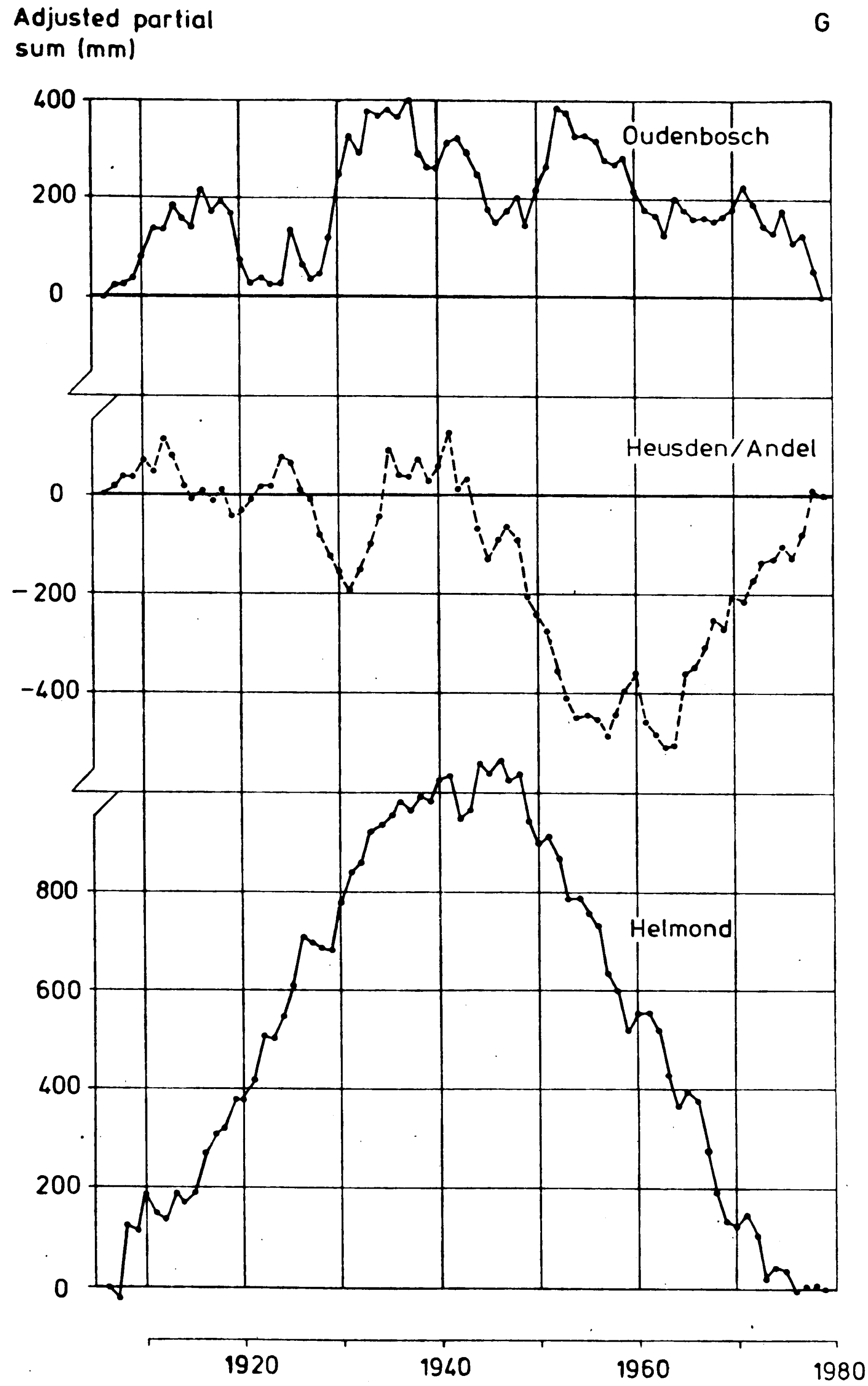

Fig. 7. (Continued)

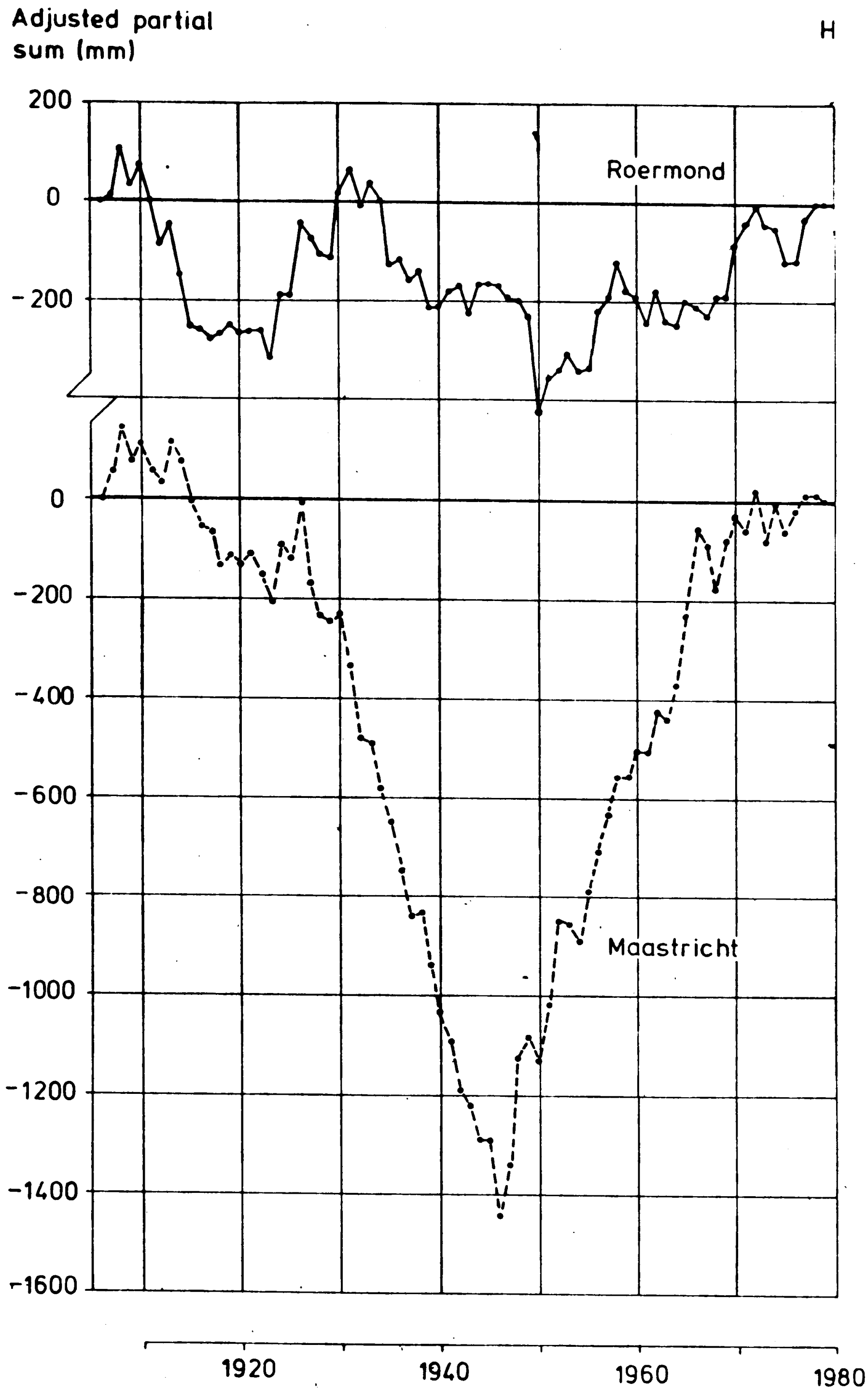

Fig. 7. (Continued)

4. Description of changes in the mean

The results of the statistical tests in the Tables 1, 2 and 3 show that for many long-term rainfall records in the Netherlands there is strong evidence of non-homogeneity. Cumulative sum plots were used in combination with the station histories to find an explanation for departures from homogeneity. For each station possible changes in the mean were investigated further with rainfall data from selected neighbouring stations.

In this chapter the nature and causes of changes in the mean are described for each station separately. The sequence in which the stations are treated is the same as in Table 1.

Den Helder/De Kooy

The series is a combination of the rainfall records from the principal climatological stations Den Helder (January 1851-July 1972) and De Kooy (August 1972-December 1979). Simultaneous measurements in Den Helder and De Kooy for the period March 1958-July 1972 indicate that for the Kooy the annual average lies about 4% above the average of Den Helder.

In the Tables 1 and 2 significant departures from homogeneity were found. The main reason for this is the low mean amount during the period 1905-1935 (about 1935 one sees in Fig. 7A a marked change in the slope of the cumulative sum plot). The station history does not give an indication for this change in the mean amount of rainfall.

West-Terschelling

The record of daily rainfall data starts in January 1876. The earliest data of this record are observations from the rainfall station Terschelling. The rainfall station West-Terschelling has been in operation since March 1885.

There is hardly statistical evidence of departures from homogeneity. The sparse station information does not give an explanation for the low value of the von Neumann ratio in Table 2. Since it concerns a very small departure form homogeneity it is also difficult to say something about change-points on the basis of the cumulative sum plot in Fig. 7A (cf. Section 2.4).

Lemmer

The series consists of observations from the rainfall station Lemmer (March 1879 – July 1941) and a rainfall station near the pumping station of the Noordoostpolder (August 1941 – December 1979). This last station is known as Lemmer Gemaal NOP or Lemmer Gemaal Buma. The record of the rainfall station Lemmer is of poor quality. Due to the sparse information on the raingauge site of this station no explanations were found for departures from homogeneity.

Leeuwarden

The record of daily rainfall data on magnetic tape starts in January 1876. The earliest data of this record are observations from the ordinary climatological station Leeuwarden. The observations at this station were terminated in 1920 but rainfall observations were continued at another location. Since November 1950 the observations have been done at the military airport of Leeuwarden.

From 1920 to 1950 the raingauge was encircled by buildings and trees. A slight improvement in the situation of the raingauge site took place in September 1935. A comparison with records from the neighbouring stations Makkum and Sint Annaparochie over the periods 1911-1919 and 1921-1930 shows for the annual mean of Leeuwarden a relative decrease of 40 mm for the period 1921-1930. For this period there is also a rapid fall in the cumulative sums of the year-by-year differences in Fig. 7B.

From June 1974 onwards the rainfall data on magnetic tape were derived from registrations of an autographic raingauge. The readings from the self-recording gauge are lower than those from the standard raingauge. For the period November 1977-February 1980 the difference is 12% during winter (October-March) and 8% during summer (April-September).

Groningen

The record of daily rainfall data on megnetic tape starts in January 1847. Up to February 1952 the record consists of observations at the old principal climatological station Groningen. The location of this station was changed in 1906. In 1952 the observations at the principal climatological station were formally terminated, but rainfall observations have been continued privately (Pelleboer, 1977). The rainfall data on magnetic tape, however, have been derived from a rainfall station at another location.

In Table 3 the 1881-1906 mean strongly differs from the 1907-1979 mean. In frequency tabels of k-day period amounts the data prior to 1906 were left out of consideration. For the period 1952-1976 the annual mean of the official rainfall record is 50 mm larger than the mean of private rainfall observations.

Ter Apel

The record of daily data from the rainfall station Ter Apel starts in September 1891. From the results of the statistical tests in the Tables 1, 2 and 3 it can be concluded that there is no evidence of departures from homogeneity.

Dwingeloo

An uninterrupted sequence of rainfall observations started in October 1911. In Table 2 significant departures from homogeneity were found. About 1925 there is a marked change in the slope of the cumulative sum plot (Fig. 7C). Though the measurements were done at the same location during the period 1911-1955, there were some changes in the raingauge site. Since no further details on these changes are given in the station archive, it is not possible to find an explanation for departures from homogeneity during this period.

Heerde

The record of daily data from the rainfall station Heerde starts in September 1892. In Table 2 the value of the rescaled adjusted range is significant at the 5% level. An examination of the cumulative sum plot and the station information did not give an indication for this small departure from homogeneity.

Denekamp

The record of daily rainfall data on magnetic tape starts in January 1905. The record consists of observations from the rainfall station Denekamp.

In Table 2 the value of the rescaled adjusted range is significant at the 1% level. The cumulative sum plot in Fig. 7C shows a relative increase of the mean rainfall amount after a station relocation in 1922. For the periods 1905-1921 and 1923-1941 the annual amounts of Denekamp were compared with the average annual amounts of the neighbouring stations Hengelo and Enschede. From this comparison it was found that after 1922 the change in the mean annual amount of Denekamp is only about 30 mm.

Hoorn

The record of daily rainfall data starts in May 1883. The earliest data of this record are observations from the rainfall station Hoorn. In December 1904 an ordinary climatological station was founded at the horticultural experiment station of Hoorn. With a short interruption from November 1946-September 1948 rainfall observations have been done at this site.

The low value of the von Neumann ratio in Table 1 is mainly due to the change of the raingauge site in 1904. There are no indications for departures from homogeneity in the rainfall record from 1904 onwards.

Hoofddorp

Rainfall observations in Hoofddorp started in 1861. From January 1867 the daily dara are available on magnetic tape. From the results of the statistical tests in the Tables 1, 2 and 3 it can be concluded that there is no evidence of departures from homogeneity.

Scheveningen

The record of daily data from the rainfall station Scheveningen starts in January 1877. In August 1907 and November 1933 station inspections revealed that the site was unsheltered. After the decrease in elevation of the raingauge in January 1946 use has been made of an earth wall to reduce the wind-error.
In the Tables 1, 2 and 3 significant departures from homogeneity were found. Though there are clear changes in the slope of the cumulative sum plot in Fig. 7D, it is not possible to find an explanation for changes in the mean amount of rainfall from the information in the station archive.

Utrecht

The series is a combination of the rainfall records from the old main observatory (January 1849-April 1897; January 1898-November 1898) and a rainfall station near the gas-works (September 1919-July 1959). No indications were found for departures from homogeneity.

Arnhem

In the previous century daily rainfall measurements in Arnhem were done during the period March 1867-March 1880. The rainfall observations started again in June 1906.

The observations in the beginning of this century were done at a rather open site near the River Rhine. The height of the rim of the raingauge was about 2.5 m. In 1926 there were several changes in the raingauge site. The observations near the River Rhine were terminated in January 1927. About this date there is a marked change of slope in the cumulative sums of the year-by-year differences in Fig. 7E. To determine the magnitude of the change in the mean amount of rainfall after 1926 the Arnhem data were compared with the data from the neighbouring stations Kilder/Doetinchem, Wageningen and Oranje Nassau's Oord (Renkum) for the periods 1908-1926 and 1927-1945. From this comparison it can be concluded that there is an increase of nearly 70 mm in the mean annual amount of Arnhem after the station relocation in January 1927.

The observations for the period January 1951-June 1952 are of poor quality due to some leakage of the raingauge. Reliable neighbouring stations for this period are Oosterbeek and Velp.

Putten

The record of daily data from the rainfall station Putten starts in March 1867. Up to 1950 the record is of poor quality with many doubtful observations. In the Tables 1, 2 and 3 significant departures from homogeneity were found. Due to the sparse station information it is not possible to give an explanation for changes in the mean amount of rainfall.

Winterswijk

The record of daily data from Winterswijk starts in July 1880. Since January 1894 the station is operational as an ordinary climatological station. During the period 30 September 1944-21April 1945 the measurements were temporally done in Meddo. Other station relocations took place in January 1904, September 1907, April 1923, March 1940 and April 1977.

In Table 2 the von Neumann ratio is significant at the 5% level. From the cumulative sums of the year-by-year differences (Fig. 7E) it is seen that after the station relocation in 1923 there is a rapid increase in the mean annual amount of Winterswijk. For the periods 1908-1922 and 1924-1939 the annual amounts of Winterswijk were compared with the average annual amounts of the neighbouring stations Aalten, Kilder/Doetinchem, Borculo and Lochem. From this comparison it turns out that there is only a small increase in the mean annual amount of Winterswijk after the change of site in April 1923. For the annual amounts the change in the mean is about 20 mm.

Vlissingen

The record of daily rainfall data from the principal climatological station Vlissingen starts in January 1855. During the period 15 August 1947-30 April 1958 the measurements were temporally done in West-Souburg. Other changes of site took place in January 1928, November 1929 (or January 1930), November 1943 and August 1945.

The oldest rainfall data of Vlissingen are of very poor quality (The annual amount of 1857 is only 40.5 mm!). Up to November 1905 the rainfall data were derived from a raingauge on the roof of the observatory (8 to 10 meters above ground-level).
During a period of one year the measurements from the raingauge on the roof were compared with those from a raingauge at 1.50 m. This comparison showed that the rainfall amounts in the gauge on the roof were too low; the difference between the rainfall totals in the two gauges was about 20%. This percentage corresponds roughly with the difference in mean rainfall over the periods 1881-1906 and 1907-1979 given in Table 3. A consequence of the large change in the mean amount of rainfall after the decrease in elevation of the raingauge in November 1905 is a very low value of the von Neumann ratio in Table 1.

In Table 2 there is strong statistical evidence of departures from homogeneity. A change of slope in the cumulative sums of the year-by-year differences in Fig. 7F suggests an effect of the changes in the raingauge site in the period 1928-1930. For the periods 1912-1927 and 1930-1941 the annual amounts of Vlissingen were compared with the average annual amounts of the neighbouring stations Sint Kruis, Ossenisse, Kapelle, Krabbendijke, Noordgouwe, Haamstede and Brouwershaven. A relative decrease of about 50 mm in the mean annual amount of Vlissingen was found for the period 1930-1941.

Kerkwerve

The record of daily data from the rainfall station Kerkwerve starts in April 1878. From Table 3 it is seen that the 1881-1906 mean differs significantly from the 1907-1979 mean. The station history does not give an indication for the low mean value during the period 1881-1906. An extensive analysis of homogeneity for the period 1951-1980 (Buishand, 1982) reveals serious departures from homogeneity in the Kerkwerve record which are possibly due to a decrease in the mean amount of rainfall after a station relocation in March 1959.

Axel

The record of daily rainfall on magnetic tape starts in January 1905. The record consists of observations from the rainfall station Axel.

In Table 2 the value of the rescaled adjusted range is significant at the 5% level. The station history does not give an indication for this small departure from homogeneity.

Heusden/Andel

An uninterrupted sequence from May 1871 was obtained by combining the records of Heusden (May 1871-January 1942) and Andel (February 1942-December 1979). Simultaneous measurements in Heusden and Andel for the period 1923-1941 indicate that for Andel the annual average lies about 4% above the average of Heusden. There are no other indications for departures from homogeneity.

Oudenbosch

The record of daily data from Oudenbosch starts in February 1888. Since January 1893 the station is operational as an ordinary climatological station. An inspection in September 1904 revealed that the raingauge was situated near a couple of trees. This resulted in a small change in the location of the raingauge.

The small value of the von Neumann ratio in Table 1 is mainly due to the less favourable measurement conditions before 1905. From the results of the statistical tests in Table 2 it can be concluded that there is no evidence of departures from homogeneity for the period 1907-1979.

Helmond

The record consists of daily rainfall data from the rainfall station Helmond for the periods January 1869-July 1872 and March 1880-December 1979. The station history does not give an indication for the low mean value for the period 1881-1906 (Table 3).

For the period 1907-1979 there is strong statistical evidence of departures from homogeneity (Table 2). The cumulative sum plot in Fig. 7G shows a marked change of slope about 1945. This is possibly due to the poor situation of the raingauge site during the period Feburary 1946-May 1974. Since the beginning of this period nearly coincides with the period in which the height of the raingauges was decreased in the Netherlands it is not possible to quantify a change in the mean amount of rainfall accurately.

Roermond

The record of daily data from the rainfall station Roermond starts in January 1869. From the results of the statistical tests in the Tables 1, 2 and 3 it can be concluded that there is no evidence of departures from homogeneity.

Maastricht

The record of daily data from Maastricht consists of observations from three climatological stations: the old principal climatological station in the Helmstraat (May 1852-1946), the rainfall station

Turennestraat/Brusselseweg (1947-1951) and the ordinary climatological station Caberg (1952-1979). To obtain a clear insight into changes in the mean amount of rainfall it is necessary to give a short description of the history of these stations.

During a long time there were two raingauges at the Helmstraat observatory: one gauge in the garden and another gauge on the top of the roof (since 1904 on a tower).

The rainfall observations in the garden started in May 1852. In May 1904 there was a decrease in the elevation of the gauge from 2 m to 1.5 m. It is not known exactly when the observations in the garden were terminated. The gauge was still operational in 1916 but an inspection in December 1930 revealed that there was no gauge in the garden.

The rainfall observations on the roof started in August 1868. Initially the height of the rim was 14.7 m. In May 1904 the raingauge was put on a small tower on the roof. It is not known, however, whether the raingauge was in the beginning on the top of the tower (29.6 m). There is a possibility that during the first years the raingauge was on a lower level (19.6 m). The observations on the tower were terminated in December 1952.

For the period August 1868-1916 there are two rainfall records for the Helmstraat observatory. There is only one record for the periods May 1852-July 1868 and 1917-1952. One problem is that it is not known whether the rainfall data for the period 1917-1929 are derived from the gauge on the tower or from the gauge on the roof.

The rainfall station Turennestraat/Brusselseweg was in operation during the period November 1946-November 1956. The transition from the Turennestraat to the Brusselseweg took place in November 1948. Rainfall observations at the ordinary climatological station Caberg started in July 1951.

For the rainfall record on magnetic tape the observations from the station with the lowest elevation were taken. From the results in the Tables 1, 2 and 3 it is seen that this combination of rainfall data does not lead to a homogeneous record. This is partly because during the period 1930-1946 (and possibly also 1917-1929) there were only rainfall observations with the gauge on the tower of the Helmstraat observatory. For this period there is a rapid fall in the cumulative sums of the year-by-year differences in Fig. 7H.

Rainfall observations before 1904 are of rather poor quality due to the unfavourable measurement conditions. For the raingauge in the garden of the Helmstraat observatory the 1853-1903 mean is only 584 mm. From comparisons with neighbouring stations in Belgium Hartman (1913) concluded that 675 mm was a better estimate for the annual mean of Maastricht.

For the period 1905-1916 the annual means of the raingauges in the garden and on the tower of the Helmstraat observatory are 669 and 616 mm, respectively (days with missing observations were not taken into account). So there is a difference of 53 mm between the two gauges.

For the period 1947-1952 the annual means for the Helmstraat observatory (raingauge at 29.6 m) and the rainfall station Turennestraat/ Brusselseweg are 677 and 793 mm, respectively. The difference is 116 mm, which is much larger than the difference for the period 1905-1916.

For the period July 1951-June 1956 the annual means for the rainfall station Turennestraat/Brusselseweg and the Caberg station are 719 and 746 mm, respectively. So the difference between these two stations is only 27 mm.

Acknowledgements

For comparing the annual amounts of a particular rainfall station with the annual amounts of neighbouring rainfall stations use was made of the collection of annual rainfall data of Dr. Ir. J.W. de Zeeuw of the Department of Land and Water Use of the Agricultural University of Wageningen. The author wishes to express his sincere gratitude to Mr. A. Denkema for his work with the rainfall data.

References

Anis, A.A. and Lloyd, E.H. (1976): The expected value of the adjusted rescaled Hurst range of independent normal summands. Biometrika, 63, pp. 111-116.

Braak, C. (1945): Influence of the wind on rainfall measurements. KNMI Meded. en Verh. 48, Rijksuitgeverij, 's-Gravenhage.

Buishand, T.A. (1977): Stochastic modelling of daily rainfall sequences. Commun. Agric. Univ. Wageningen 77-3, Wageningen.

Buishand, T.A. (1982): Some methods for testing the homogeneity of rainfall records. To appear.

Craddock, J.M. (1979): Methods of comparing annual rainfall records for climatic purposes. Weather, 34, pp. 332-346.

Hartman, C.M.A. (1913): Het klimaat van Nederland. A. Neerslag. KNMI Meded. en Verhand. 15, Rijksuitgeverij, 's-Gravenhage.

Mitchell, Jr., J.M. (1961): The measurement of secular temperature change in the Eastern United States. Research Paper No. 43, US Weather Bureau, Washington, D.C.

Pelleboer, J.H. (1977): Extremen in Noord-Nederland. Nieuwsblad van het Noorden, Groningen.

Sneijers, R. (1957): Sur la détermination de l'homogenéité des séries climatologiques. J. Scientifique de la Météorologie, VII-28, pp. 359-372.

Wallis, J.R. and O'Connell, P.E. (1973): Firm reservoir yield - How reliable are historic hydrological records? Hydrol. Sci. Bull., XVIII, pp. 347-365.

WMO (1966): Climatic change. Tech. Note 79, World Meteor. Org., Geneva.

Worsley, K.J. (1979): On the likelihood ratio test for a shift in location of normal populations. J. Amer. Statist. Assoc., 74, pp. 365-367.