JC

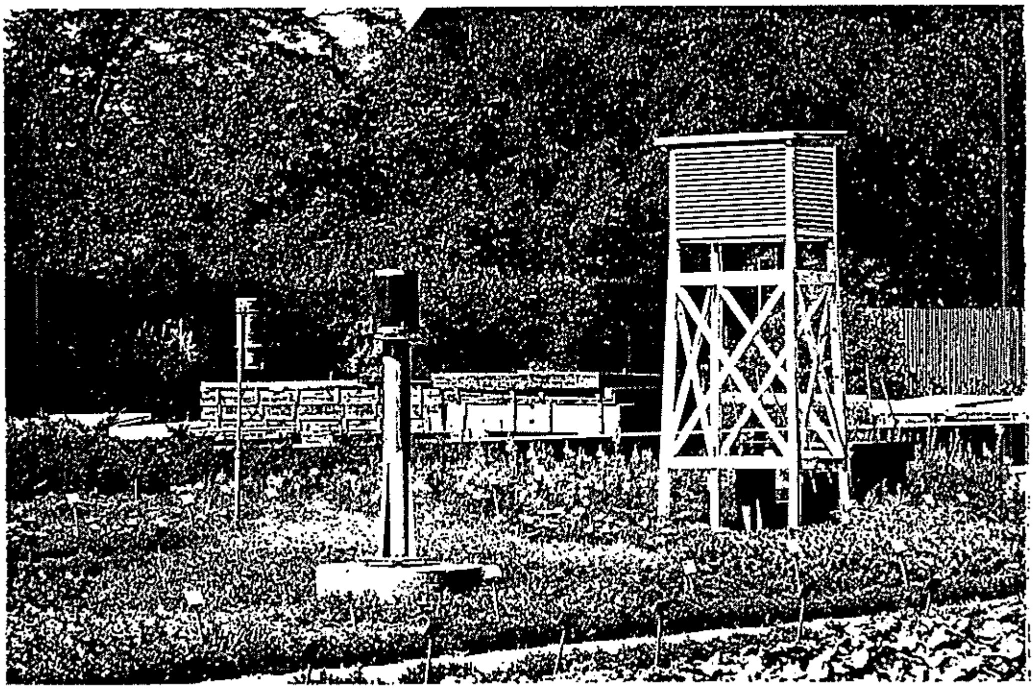
DANISH METEOROLOGICAL INSTITUTE

TECHNICAL REPORT

93-12

Homogeneity test of climatological data

P. Steffensen, F. Lynge Larsen and J. Cappelen



DMI

COPENHAGEN 1993

PREFACE

This report describes the methods used at The Danish Meteorological Institute (DMI) in homogeneity testing of various climatological elements. The method is based on the Standard Normal Homogeneity Test (SNHT) developed by Hans Alexandersson at SMHI. The test is described and an example concerning precipitation is used to show the efficiency of the testing procedure. Data from the station history archives are used to evaluate the test results and a correction is applied to the data.

CONTENTS Page

1.INTRODUCTION

Time series of meteorological elements are essential for studies of climatological fluctuations and changes. Due to changes in measuring methods, instrumentation or the environment however, the studies of climate may produce artificial breaks or trends (i.e. inhomogeneities) that overshadow the real variations. Long time series therefore should be tested statistically and corrected in case of inhomogeneities, before being used in analysis of the climate.

A homogeneous time series containing climatological data could be defined as a sequence of values which follow a reference series at all times within certain limits and where the reference time series is assumed to be homogeneous. In other words: Homogeneity in a climatological time series can only be tested by comparing with other series. A homogeneous time series is a series not influenced by:

1) Sudden artificial effects such as change in instrumentation and observers, relocation of the station etc., defined as breaks.
2) Long time natural/artificial effects such as a slow change in the surroundings (urbanization, vegetation) defined as trends.

Several methods including the SNHT can be used to test hypotheses concerning homogeneity. Contrary to most other methods the SNHT can not only detect the inhomogeneity but also give the significance level of the inhomogeneity and supply the parameters used for repairing the inhomogeneous time series. A disadvantage of the present version of the SNHT is that it can only be used in connection with time series with a single break and not for series with multiple breaks or trends. This problem has however been solved recently by H. Alexandersson [Alexandersson 93] and this improvement will be incorporated in one of the next floppy version of SNHT. Working with multiple breaks in the current version of SNHT it is necessary to divide the series into two or more series, each series containing one inhomogeneity at most.

2.METHOD OF HOMOGENEITY TESTING

Before testing observations of climatological elements as for instance precipitation and temperature we have to make some assumptions concerning the distributions of the observations. Considering precipitation, the relationship between two observations at the same time is assumed to be cumulative and thus the ratio between the two observations is assumed to be normally distributed. In the case of temperature measurement the relationsship between two observations at the same time from different stations is assumed to be additive and therefore the difference between the two observations is considered to be normally distributed.

To test the homogeneity of a station we form a sequence $\{q_i\}$ of ratios (in the case of cumulative elements as precipitation) or differences (in the case of additive elements as temperature) between the test station and a reference value. The reference value is made from the surrounding reference stations which are all assumed to be homogeneous. Each reference station may be weighted depending on how well it corresponds to the test station. If the test station is homogenuous the sequence $\{q_i\}$ is normally distributed. From $\{q_i\}$ a standardized sequence $\{z_i\}$ is formed and a likelihood ratio test is carried out on $\{z_i\}$.

2.1 The reference stations

2.1.1 Cumulative elements

Following climatological elements are assumed to be of cumulative nature:

- Precipitation.
- Hours of bright sunshine.
- No. of days with storm, fog, snow etc.

Considering these elements we can assume that

$$\frac{y_i}{g_i(x_{ji})} \in N(\mu,\sigma^2)$$

where

y_i : Observation at time i from the test station.

$g_i(x_{ji})$: Weighted mean of the observations from the reference stations at time i:

$$g_i(x_{ji}) = \frac{\sum_{j=1}^{k_i} w_j \frac{x_{ji}}{x_j}}{\sum_{j=1}^{k_i} w_j}$$

where

x_{ji} : Observation from reference station j at time i.
x_j : Mean of observations from reference station j.
w_j : Weight of the reference station j.
k_i : Number of reference stations at time i.

2.1.2 Additive elements

Elements assumed to be of additive nature are:

- Temperature.
- Air pressure.
- Cloud cover.

In the case of additive elements we have that:

$$y_i - g_i(x_{ji}) \in N(\mu,\sigma^2)$$

where
y_i : Observation at the test station at time i.
$g_i(x_{ji})$: Weighted mean of the reference stations at time i:

$$g_i(x_{ji}) = \frac{\sum_{j=1}^{k_i} w_j(x_{ji}-x_j)}{\sum_{j=1}^{k_i} w_j}$$

where
x_{ji} : Observation at reference station j at time i.
x_j : Mean of reference station j.
w_j : Weight of the reference station j.
k_i : Number of reference stations at time i.

2.1.3 Weighting

There are several ways of defining the weight w_j of reference station j. The weight should depend on how well the reference station corresponds to the test station.

Firstly the weight can depend on the correlation coefficient c_j^2 between the test station and the reference station j:

$$w_j = c_j^2 = \frac{s_{yx_j}^2}{s_y s_{x_j}}$$

where:

s_y^2, s_x^2, s_{yx}^2: Standard deviations of test and reference stations and covariance between test and reference station and

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-y)^2 \ , \quad s_{x_j}^2 = \frac{1}{n_j-1}\sum_{i=1}^{n_j}(x_{ji}-x_j)^2 \ ,$$

$$s_{yx_j}^2 = \frac{1}{m-1}\sum_{i=1}^{m}(y_i-y)(x_{ji}-x_j)$$

y_i : Observation no. i from the test station,
y : Mean of observations from test station,
x_{ji} : Observation no. i from reference station j,
x_j : Mean of observations from reference station j,
n,n_j,m: No. of observations from test station, reference station and test and reference station at the same time, respectively.

Obviously a large correlation coefficient means that the reference station corresponds well to the test station and therefore the reference station is weighted high.

Secondly the weight of reference station j can be definded as the reciprocal of the standard deviation s_j^2 of the reference station j:

$$w_j = \frac{1}{s_j^2} = \frac{1}{\frac{1}{n_j-1}\sum_{i=1}^{n_j}(x_{ji}-x_j)^2}$$

This means that a reference station with large variations is "scaled down" so that it corresponds to other reference stations with smaller variations.

Thirdly the weight can be defined as a combination of the above:

$$w_j = \frac{c_j^2}{s_j^2}$$

Finally the weight can depend exponentially on the distance l_j between the test station and reference station j, e.g.:

$$w_j = e^{-dl_j}$$

This way a reference station near the test station is weighted higher than a reference station farther away. However, this is probably reflected in the correlation coefficient.

The weighting of the reference stations can also be based on other meteorological conditions or general knowledge of the individual reference station.

2.2 Description of the test

We now form a new series $\{q_i\}$ of normally distributed ratios from cumulative and additive observations, respectively:

$$q_i = \frac{y_i}{\frac{\sum_{j=1}^{k_i} w_j \frac{x_{ji}}{x_j}}{\sum_{j=1}^{k_i} w_j}} \quad , \quad q_i = y_i - \frac{\sum_{j=1}^{k_i} w_j(x_{ji}-x_j)}{\sum_{j=1}^{k_i} w_j}$$

We want to test the hypotheses H_0 against the alternative H_1:

H_0: The test station is homogeneous, i.e. the series $\{q_i\}$ is homogeneous mean value and variance independent of i.
H_1: The test station is inhomogeneous. The series $\{q_i\}$ has one mean value the first v years and another mean value the last $(n-v)$ years. In other words $\{q_i\}$ has a single break between $i=v$ and $i=v+1$.

Forming a standard normally distributed series $\{z_i\}$:

$$z_i = \frac{q_i - \bar{q}}{s_q}$$

where

$$\bar{q} = \frac{1}{n}\sum_{i=1}^{n} q_i$$

and

$$s_q^2 = \frac{1}{n-1}\sum_{i=1}^{n} (q_i - \bar{q})^2$$

the hypotheses can be written as:

$$H_0\text{: } z_i \in N(0,1)$$

$$H_1\text{: } \begin{matrix} z_i \in N(\mu_1,\sigma_1^2), & i \le v \\ z_i \in N(\mu_2,\sigma_2^2), & i > v \end{matrix}, \quad v \in [1,n-1], \quad \sigma_1^2 < 1, \quad \sigma_2^2 < 1$$

We now use the likelihood ratio test to test H_0 against the alternative H_1. The propability function for a normally distributed value z with unit variance:

$$z \in N(\mu,1)$$

is

$$f(z,\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\mu)^2}$$

and with n observations $z_1,\ldots,z_n$ this gives us the likelihood function

$$L(\mu) = \prod_{i=1}^{n} f(z_i,\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z_i-\mu)^2} = (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n}(z_i-\mu)^2}$$

The test ratio is then given by

$$c(\mu_1,\mu_2,v) = \frac{\max_{\mu_1,\mu_2,v} L(\mu_1,\mu_2,v)}{\max_{\mu=0} L(\mu)}$$

$$= \frac{\max_{\mu_1,\mu_2,v}[(2\pi)^{-\frac{n}{2}}(e^{-\frac{1}{2}\sum_{i=1}^{v}(z_i-\mu_1)^2-\frac{1}{2}\sum_{i=v+1}^{n}(z_i-\mu_2)^2})]}{(2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n} z_i^2}}$$

To find the maximum of the nominator we take the logarithm and the derivative of the likelihood functions which give us:

$$-\sum_{i=1}^{v}(z_i-\mu_1)-\sum_{i=v+1}^{n}(z_i-\mu_2) = 0 \Rightarrow$$

$$\mu_1 = \frac{1}{v}\sum_{i=1}^{v} z_i = \overline{z_1} \wedge \mu_2 = \frac{1}{n-v}\sum_{i=v}^{n} z_i = \overline{z_2}$$

We can now find the maximum value of the test ratio:

$$c(v) = \frac{\max[(2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{v}(z_i-\overline{z_1})^2-\frac{1}{2}\sum_{i=v+1}^{n}(z_i-\overline{z_2})^2}]}{(2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n} z_i^2}} < C \Rightarrow$$

$$\max_v[e^{-\frac{1}{2}\sum_{i=1}^{v}(z_i-\overline{z_1})^2-\frac{1}{2}\sum_{i=v+1}^{n}(z_i-\overline{z_2})^2+\frac{1}{2}\sum_{i=1}^{n} z_i^2}] < C \Rightarrow$$

$$\max_v[-\sum_{i=1}^{v} z_i^2+2\overline{z_1}\sum_{i=1}^{v} z_i-\sum_{i=1}^{v}\overline{z_1}^2-\sum_{i=v+1}^{n} z_i^2+2\overline{z_2}\sum_{i=v+1}^{n} z_i-\sum_{i=v+1}^{n}\overline{z_2}^2+\sum_{i=1}^{n} z_i^2] < 2\ln C \Rightarrow$$

$$\max_v[2\overline{z_1}v\overline{z_1}-v\overline{z_1}^2+2\overline{z_2}(n-v)\overline{z_2}-(n-v)\overline{z_2}^2] < 2\ln C \Rightarrow$$

$$\max[v\overline{z_1}^2+(n-v)\overline{z_2}^2] < 2\ln C$$

If the test parameter T is defined as

$$T(v) = v\overline{z}_1^2 + (n-v)\overline{z}_2^2$$

the idea is to find the maximum value of $T(v)$, T_{max}, by varying v:

$$T_{\max} = \max_v[T(v)]$$

It is not possible to find the distribution of T_{max} under H_0 analytically but critical values $T_{1-\alpha}$ may be computed by simulation. Tables of critical values $T_{1-\alpha}(n)$ depending on the significance level α and the number of observations n, can be found in [Alexandersson 86] and [Potter 81] and the table below is made from these.

n	10	15	20	30	40	70	100	150	200
T_{75}	4.7	4.9	5.0	5.3	5.4	5.9	6.0		
T_{90}	6.0	6.5	6.7	7.0	7.3	7.9	7.9	8.1	8.2
T_{95}	6.8	7.4	7.8	8.2	8.7	9.3	9.3	9.4	9.5
T_{99}	7.9	9.3	9.8	10.7	11.6	12.2	12.5		

Table 2.1: Critical values of $T_{1-\alpha}$ for the SNHT.

The critical values for the 10% significance level for instance is denoted T_{90}. If we define a significance level of 10% and $T_{max} > T_{90}$, H_0 is rejected. However, there will be a 10% risk of having rejected a homogeneous series because of random variation. Defining a significance level of 5% instead will reduce this risk to 5% but make the possibillity of rejecting H_0 smaller.

3.HOMOGENEITY TEST IN PRACTICE

3.1 The test procedure

First all reference stations are run against the test station, and secondly one reference station at a time is picked out in order to check the homogeneity of the single reference station. If the result differ radically from the former run the reference station picked out may be inhomogeneous.

First phase.

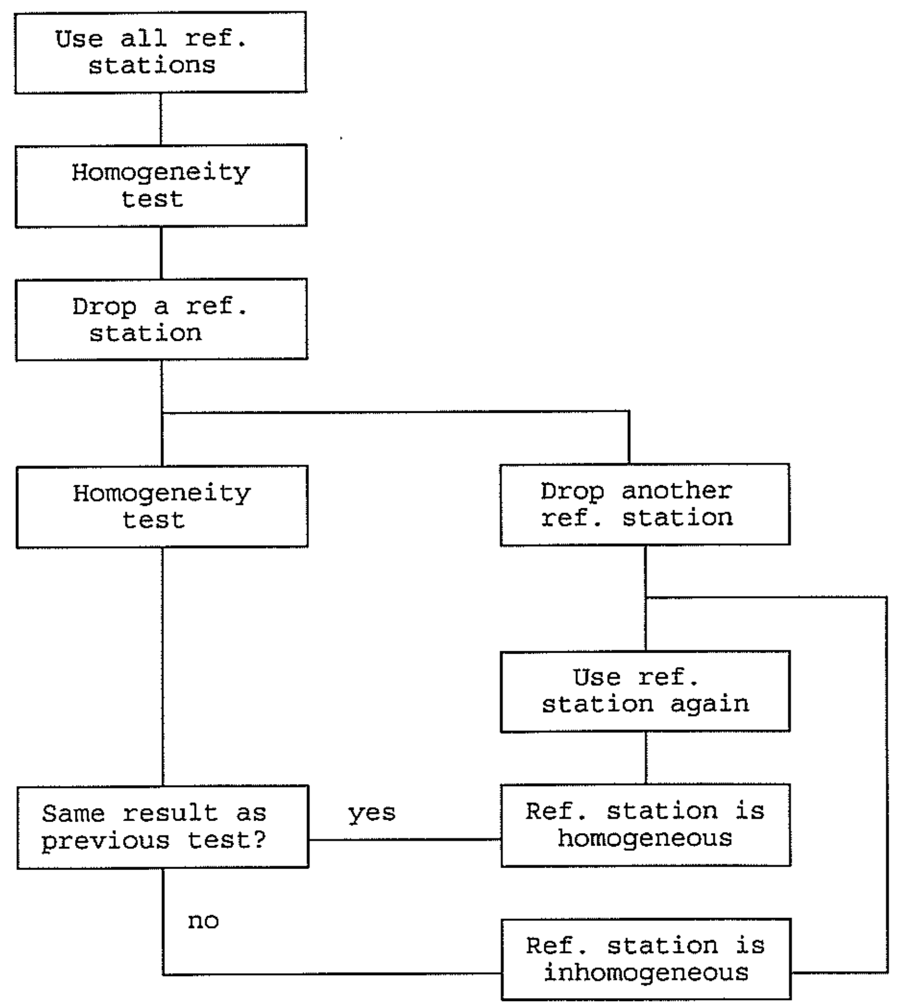


Figure 3.1: Flow diagram for reference station checking.

After this the remaining homogeneous reference stations are used in the further process. If necessary the test series may be corrected and the homogeneity test is repeated, hopefully for the last time.

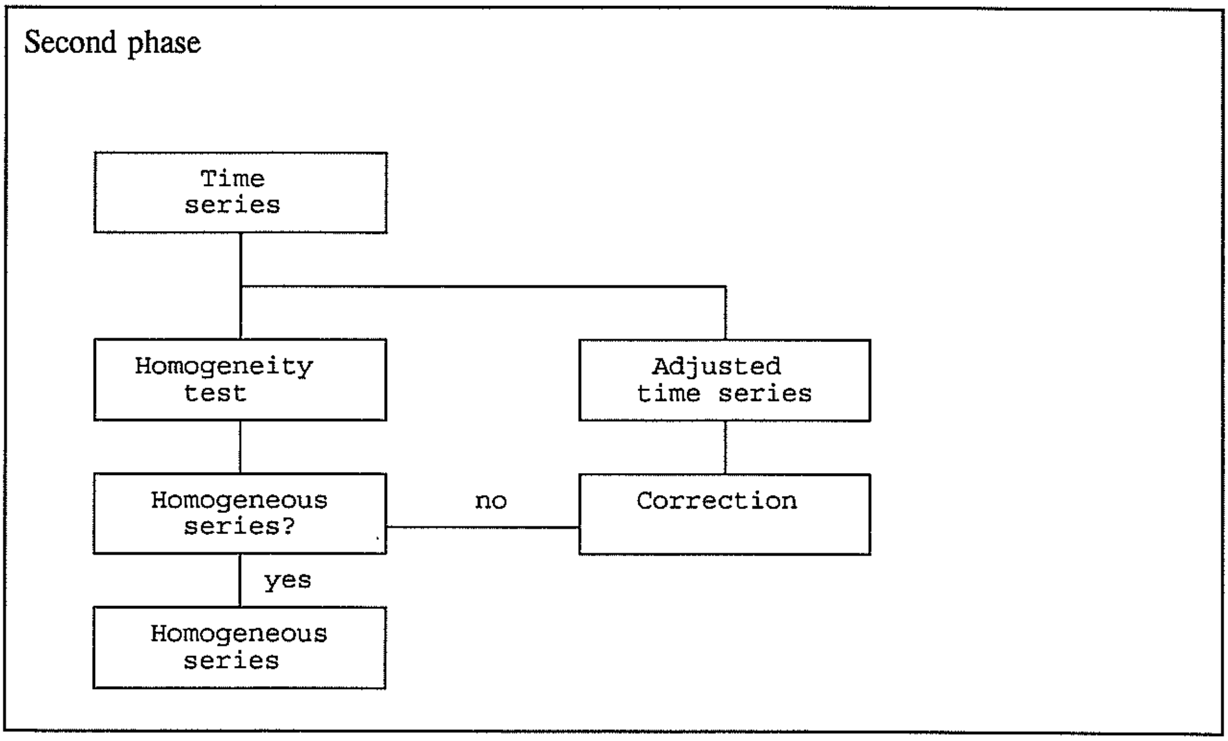


Figure 3.2: Flow diagram for homogeneity testing.

3.2 General remarks

When testing homogeneity of long climatological time series it is very important to use the SNHT with caution. The test should be used only as a guideline. It must never be used alone in a test of a station but should be held together with the station history and general knowledge of the area where the test station is situated.

Firstly it might be a problem constructing a reliable pseudo reference station. Very few stations satisfy the requirements of both being homogeneous over a period of several decades and lying in the neighbourhood of the test station. The test should include several reference stations to avoid a major influence from an inhomogeneous reference station.

Secondly the test is liable to show breaks in the beginning or at the end of a series. This is because the series before a "break" in the beginning and after a "break" at the end is based on very few observations. Generally the time series should be of a length of at least 20 years and the break should occour more than 5 years from the beginning and the end of the series.

In [Frich Cappelen 92] it is suggested the rule below should be followed by the participants of the NACD project:

A series is inhomogeneous if

1) the inhomogeneity is significant at a 10% level, and is explained in station history or

2) the inhomogeneity is significant at a 5% level, and at least 5 years from start/end of the series and no change is indicated in the station history.

4.AN EXAMPLE OF HOMOGENEITY TESTING

The precipitation series (1961-1990) from station 30380 Landbohøjskolen, situated in the centre of Copenhagen, was tested against 8 surrounding stations. The 8 reference stations are all located within 15 km from the test station. In table 4.1 the stations are listed.

Station number	Name	Lattitude deg.min.	Longitude deg.min.	Altitude m.a.s.l.
30380	**Landbohøjskolen**	**55 41 N**	**12 32 E**	**9**
30220	Virum	55 47 N	12 30 E	31
30230	Store Hareskov	55 46 N	12 26 E	48
30240	Søndersø	55 46 N	12 21 E	15
30250	Bogøgård	55 47 N	12 18 E	8
30300	Nybølle	55 42 N	12 16 E	13
30310	Islevbro	55 42 N	12 27 E	10
30320	Tinghøj	55 44 N	12 30 E	48
30390	Torsbro II	55 37 N	12 16 E	16

Table 4.1: Test station 30380 Landbohøjskolen and reference stations 1961-90.

DMI/B 1993.10. 5 page 1

THE STANDARD NORMAL HOMOGENEITY TEST

Precipitation

SUMMARY

Weight : 1.00
Test station and element : 30380 R
Ref. station and element : 30220 R
Ref. station and element : 30230 R
Ref. station and element : 30240 R
Ref. station and element : 30250 R
Ref. station and element : 30300 R
Ref. station and element : 30310 R
Ref. station and element : 30320 R
Ref. station and element : 30390 R
Data period : 1961 - 1990

Ratio between test station and pseudo station analyzed

	Winter	Spring	Summer	Autumn	Year
Number of years	29	30	30	30	30
Maximum t-value	13.7	7.9	4.4	3.0	13.1
Year of max t-value	1970	1989	1972	1982	1971
Significant 0.90	Yes	Yes	No	No	Yes
Critical t-value	6.7	6.7	6.7	6.7	6.7
Significant 0.95	Yes	Yes	No	No	Yes
Critical t-value	7.9	7.9	7.9	7.9	7.9
DEVIATIONS FROM MEAN VALUE OF TEST STATION					
Mean before break	143.55	********	208.17	183.88	662.56
Mean after break	146.07	********	169.02	170.81	620.07
Q VALUES					
Mean before break	159.65	********	195.75	183.23	666.25
Mean after break	137.51	********	176.87	170.78	613.67
Q ratio (after/before)	.86	********	.90	.93	.92

Figure 4.1: A summary of the output from the standard normal homogeneity test applied at 30380 Landbohøjskolen and all reference stations weighted equally. The asterix means that the inhomogeneity was found within 5 years from the end of the series, which is not acceptable.

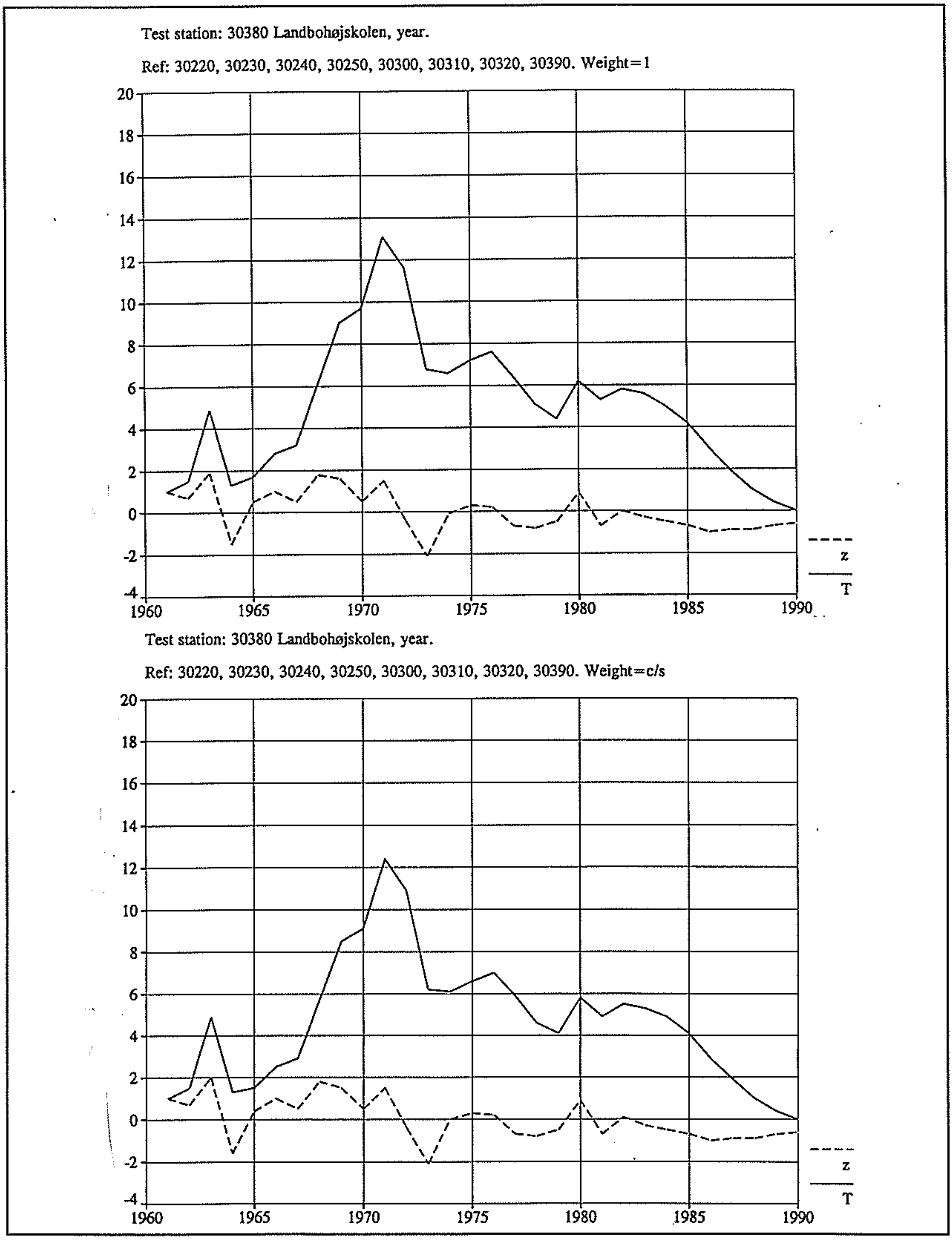


Figure 4.2: Plots of T- and z-values from homogeneity test of 30380 Landbohøjskolen. The upper plot with all reference stations weighted equally and the lower plot with the reference stations weighted with the correlation coefficient devided with the standard deviation. In this example there is hardly any difference.

As shown in figure 4.1 the maximum T-value for the yearly precipitation sum is 13.1 and significant at a 5% level. In table 2.1 we see that it is significant even at a 1% level since $13.1 > T_{99}(30) = 10.7$. The last line of the output in figure 4.1 shows that the amount of precipitation was reduced by 8% after the homogeneity break.

A part of the output gives the opportunity to vizualize the observations, the z-values and the T-values, see figure 4.2. The plots show a clear homogeneity break in 1971 as a "peak" in the T-curve, and it is also seen in the z-values, which have different means before and after the break.

The homogeneity break is significant at at least a 5% level and following the rules for correcting inhomogeneous series;

1) the inhomogeneity is significant at the 10% level, and is explained in station history or

2) the inhomogeneity is significant at the 5% level, and at least 5 years from start/end of the series, when there is no indication from the station history that a change has occured,

we could then correct the inhomogeneity without consulting the station history.

Anyhow the photographs indicate that at least until september 1971 a hedge had been growing next to the precipitation gauge. Judged from the photographs we assume that the hedge has been removed october/november 1971, leaving the precipitation gauge more exposed, see figure 4.3. This explains the higher amount of precipitation before the break.

An 8% correction before the break gives a new adjusted time series, see figure 4.4, and the homogeneity test is repeated. The results are shown in figure 4.5 and figure 4.6 now indicating that the yearly precipitation series is homogeneous.

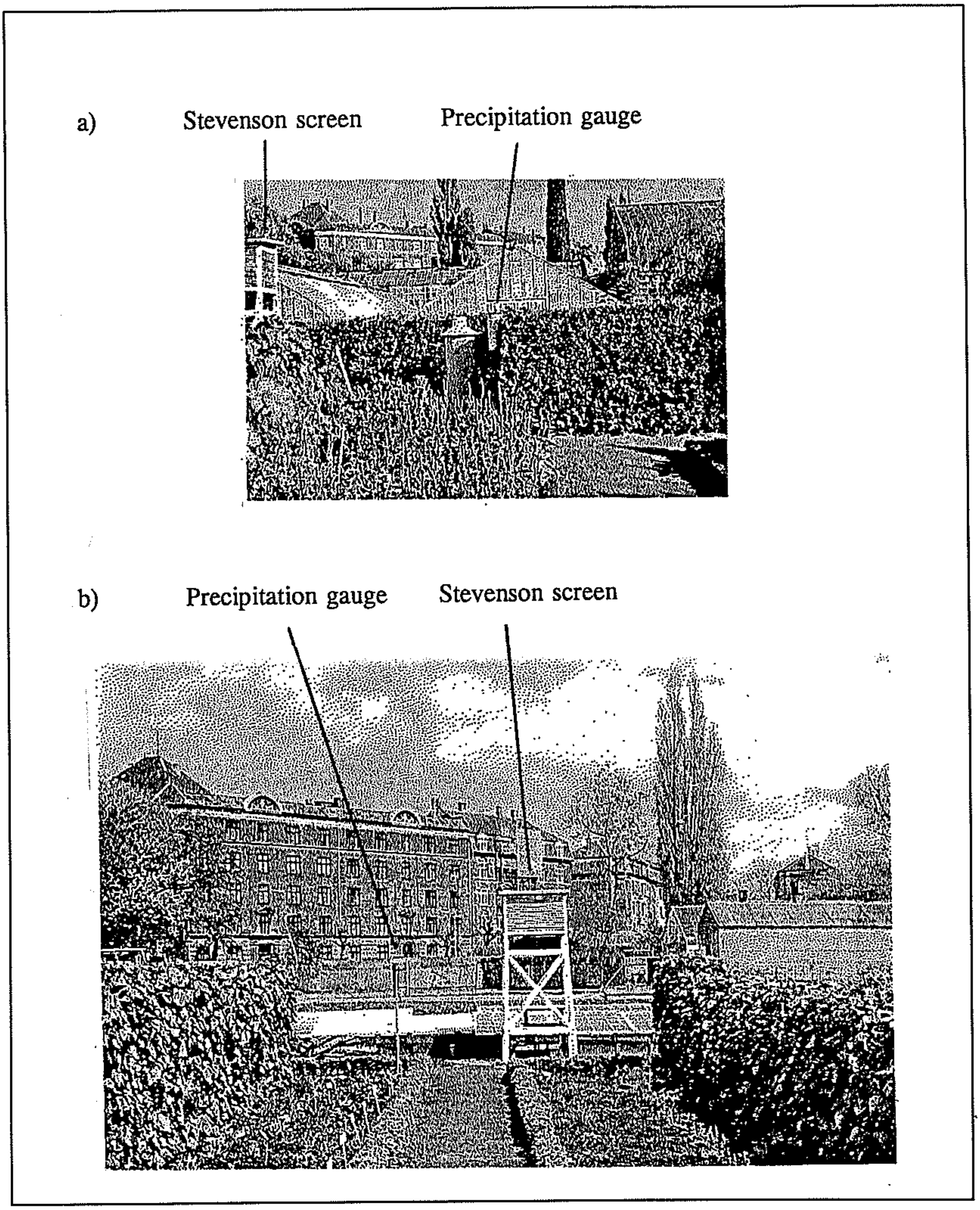


Figure 4.3: Precipitation gauge (Hellmann) and Stevenson screen as seen from south a) 15.9.1971 and b) 13.5.1987. Notice that the hedge by the precipitation gauge is removed and that the Stevenson screen has been moved whereas the precipitation gauge has not.

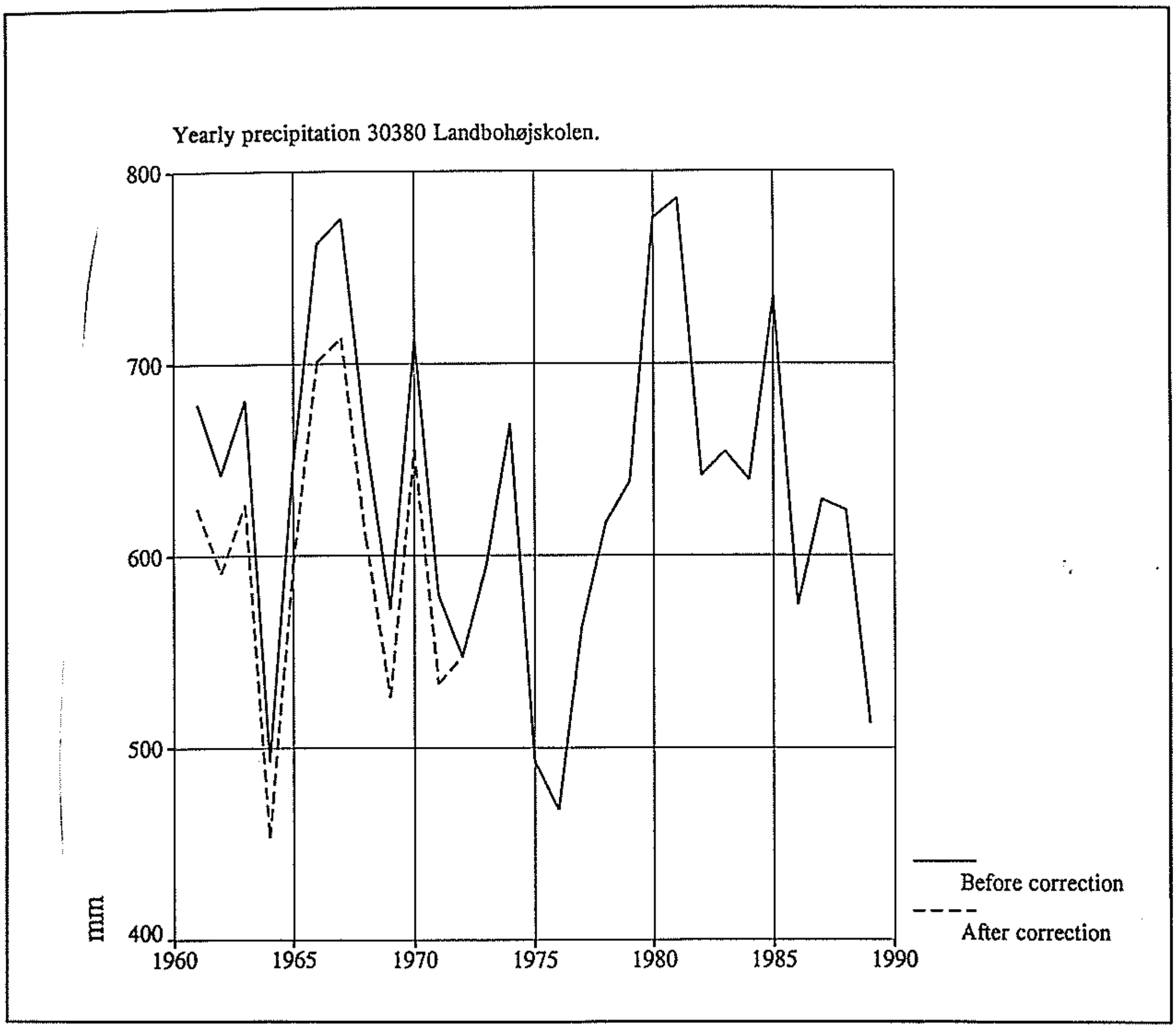


Figure 4.4: Adjusted series 30380 Landbohøjskolen.

DMI/B 1993.10. 6 page 1

THE STANDARD NORMAL HOMOGENEITY TEST

Precipitation

SUMMARY

Weight : 1.00
Test station and element : 30380 R
Ref. station and element : 30220 R
Ref. station and element : 30230 R
Ref. station and element : 30240 R
Ref. station and element : 30250 R
Ref. station and element : 30300 R
Ref. station and element : 30310 R
Ref. station and element : 30320 R
Ref. station and element : 30390 R
Data period : 1961 - 1990

Ratio between test station and pseudo station analyzed

	Winter	Spring	Summer	Autumn	Year
Number of years	28	30	30	30	29
Maximum t-value	6.7	5.3	5.0	3.3	6.2
Year of max t-value	1970	1989	1989	1966	1980
Significant 0.90	Yes	No	No	No	No
Critical t-value	6.6	6.7	6.7	6.7	6.7
Significant 0.95	No	No	No	No	No
Critical t-value	7.8	7.9	7.9	7.9	7.9

DEVIATIONS FROM MEAN VALUE OF TEST STATION

Mean before break	132.04	********	********	140.56	594.43
Mean after break	140.56	********	********	181.25	594.04
Q VALUES					
Mean before break	144.10	********	********	164.73	614.06
Mean after break	133.07	********	********	176.71	577.54
Q ratio (after/before)	.92	********	********	1.07	.94

Figure 4.5: Output from the Standard normal homogeneity test after the correction. There is no significant inhomogeneity.

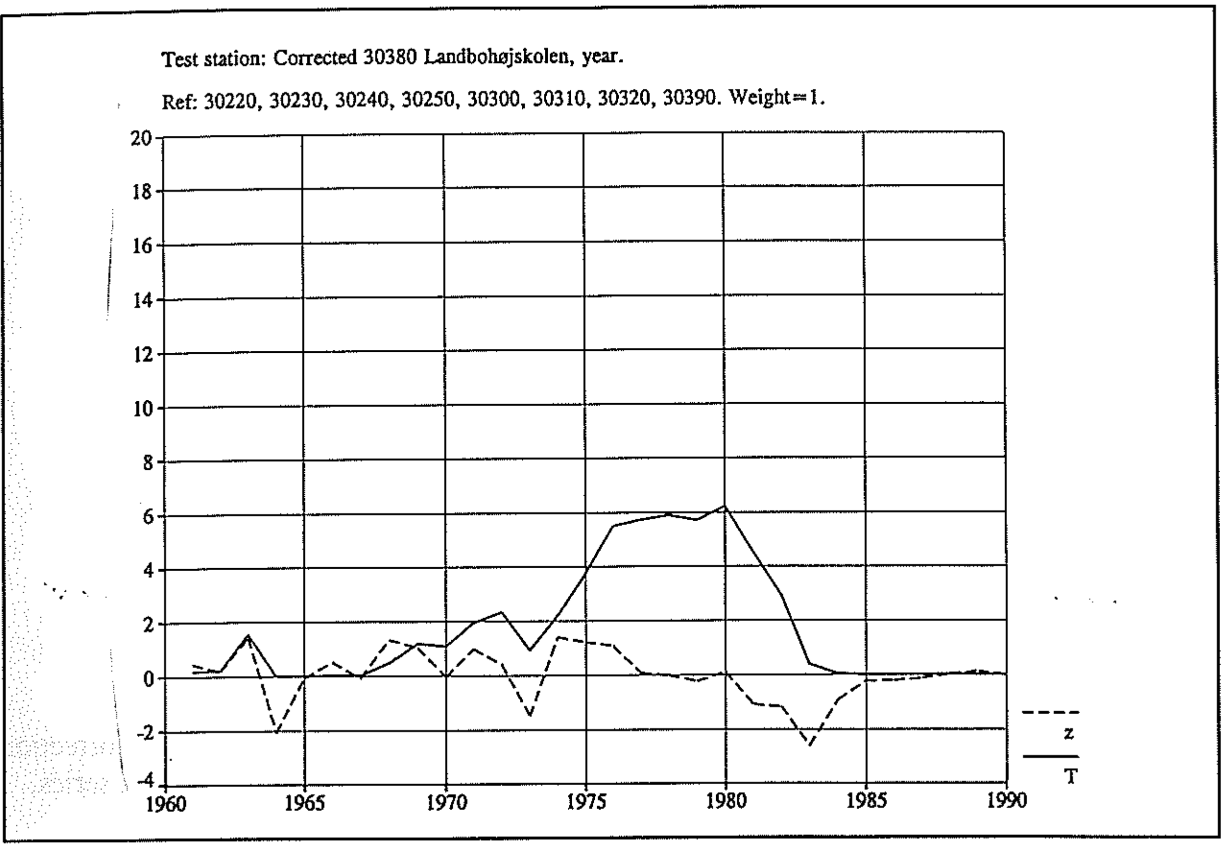


Figure 4.6: Plots of z- and T-values after the corrections and all reference stations weighted equally. There is no significant inhomogeneity in the series now.

REFERENCES

[Alexandersson 86]:
Alexandersson, H., 1986:
A homogeneity test applied to precipitation data.
Journal of Climatology, vol. 6, p.661-675.

[Alexandersson 93]:
Alexandersson, H., 1993:
Urban effects on Nordic temperature data - a statistical test.
Submitted to Journal of Climate.

[Conradsen 81]:
Conradsen, K., 1984:
En introduktion til statistik.
IMSOR.

[Frich Cappelen 92]:
Frich, P. and Cappelen, J., 1992:
Proceedings from Third Nordic Climate Workshop: Homogeneity testing and correction of spatial and temporal inhomogeneities in North Atlantic climatological data series.
DMI Technical Report 92-15.

[Hanssen-Bauer 91]:
Hanssen-Bauer, I., 1991:
Homogeneity test of precipitation data. Description of the methods used at DNMI.
DNMI-Rapport 13/91 KLIMA.

[Potter 81]:
Potter, K. W., 1981:
Illustration of a new test for detecting a shift in mean in precipitation series.
Monthly Weather Review, vol. 109, p. 2040-2045.