Geophysica (1998), 34(3), 159-178

Observations Inhomogeneities and Detection of Climate Change

The case of the Oulu (Finland) air temperature series

R. Sneyers, H. Tuomenvirta** and R. Heino***

*Institut royal météorologique de Belgique
**Finnish meteorological institute

(Received: June 1998; Accepted: October 1998)

Abstract

It is now well known that inhomogeneities in observations and climate instability have the common statistical property of producing generally in climatological series abrupt changes in the distribution of their elements (change-points) characterized by shifts in the mean and (or) in the scatter around this mean (dispersion). It follows that elimination of inhomogeneities in the observations should allow a better appreciation of the evolution of local climates and of the derived regional and global ones.

For undocumented inhomogeneities as for climate instability, among the different methods used for the detection of change-points, many have a questionable efficiency due to inappropriate testing or to incomplete account of the available information.

The purpose of this paper is to give and to justify a logical approach of the problem and to illustrate its efficiency for the search of modifications in the statistical characterics of climatological series created by observations inhomogeneities.

The interest of the Oulu (Finland) series consists in its length, the relative high latitude of its situation and the care with which the problem of inhomogeneities has been solved.

In the conclusion, the importance of the statistical aspect of climate changes is emphasized.

Key words: Climatological time series, homogenization, sequential statistical analysis, change-point search, joint significance levels, climate evolution

1. *Introduction*

If the general causes of inhomogeneites in observations have always explained in an evident manner the abrupt character of changes in the distribution of elements in climatological time series, inversely, the conviction that climate instability is characterized by abrupt changes is still not universal.

Actually, climate change has generally been made clear by the evolution of climatic features such as alteration of snow lines, extension or reduction of glaciers termini, lake level changes, etc. Moreover, the usual progressive way of these changes has made it generally unbelievable that climate itself could evolve abruptly. In particular, when the climate warming detected by *Lewis* (1947), *Lysgaard* (1948) and *Vandenplas* (1948) appeared statistically to be the result of abrupt seasonal temperature increases as well for Brussels as for Paris (*Sneyers,* 1958), the explanation given by the progressive urbanization of the observatories surroundings was generally preferred (*Dettwiller* 1970). However, *Karl et al*. (1988) have shown that, for the USA, the bias in the annual mean temperature due to the urbanization remains of the order of .1°C.

On the other hand (*Sneyers*, 1955), the inhomogneities which perturb the quality of climatological series are mostly the result of very punctual features such as instrumentation, observation methods or location changes. Exception made when resulting from the extention of a heat island, from progressive alteration of the environment around the station or from the recording device, they generally produce modifications having a step-like character and a seasonal variation.

Now, if climate changes and inhomogeneities in observations have similar forms, at the seasonal scale, they have often a different character. Actually, if modifications of the observation process act logically at the same time in the annual, seasonal or monthly series, as a general rule, this is not the case with a climate change. For instance, for the change discovered for Brussels-Uccle and Paris, the temperature increase occurred round 1910 for winter and spring and round 1930 for summer and autumn which leads, at the annual scale, to two successive increasing shifts placed round 1910 and 1930.

It follows that, if the same methodology is valid for the search of change-points in both cases, distinction between observations and climate causes is only possible when comparing the tested series with reference series placed in very similar climate conditions.

Obviously, detection and correction of inhomogeneities in observations are indispensable for ensuring reliable conclusions on the statistical properties of climatological series and thus, in climate evolution research.

After having justified the methodologies used here for solving the different problems raised, the purpose of this paper is to give the results of the analysis of the air temperature series of Oulu (Finland), before and after adjustments for inhomogeneities, in order to have an idea of the degree of improvement gained.

At the climatological point of view, the analysis of Oulu makes possible the comparison with results found for other locations and adds further useful information to earlier investigations made on the climate characteristics of the region by *Heino* (1992, 1994), *Tuomenvirta* and *Drebs* (1994) and *Tuomenvirta* and *Heino* (1996).

2. The homogeneity assessment of observations

Any observation data is generally considered as being the sum of two parts, the first one giving the right value of the measure of the observed variable and the other being the error on the measure due to the observation process (*Sneyers*, 1955). It follows that changes in the process producing the observation (instrumentation, method of observation) may change the statistical properties of the measure as well as of its error. In addition, when the location of observations changes, the value of the measure may be influenced differently by the local environmental conditions (micro-climate).

For evaluating the importance of such inhomogeneities and for computing the necessary corrections for homogenizing series of data, the comparison with simultaneous homogeneous series has to be placed in the case of correlated series (*Sneyers*, 1969, 1975, 1989). The simplest case is obviously the one when, among the parameters defining the distributions of the series, the single perturbed one to be adjusted is the mean.

Moreover, homogenization should only be made for those observational perturbations confirmed in the history of the tested series (metadata, *Tuomenvirta* and *Drebs*, 1994) or by other appropriate information suggested by climatological experience (*Sneyers*, 1972).

3. The statistical homogeneity (simple randomness) assessment and the search for change-points

Identifying the production of a series of observations with the sampling of elements from a given population leads to relate statistical homogeneity to simple randomness, that is when each element of the series is randomly and identically distributed. It follows that testing statistical homogeneity reduces to the verification that each element of the series is independent of (uncorrelated with) all the other ones and that they all have an identical distribution function (i.i.d.) (*Sneyers*, 1963).

For this verification, two kinds of tests are available, the parametric ones, for which the type of the underlying distribution, generally the normal one, is ascertained, and distribution free ones. It is obvious that, in absence of an ascertained normal distribution, testing randomness with the use of parametric tests would, in a certain manner, be begging the question. On the other hand, for the rank $r(x_i)$ of the element x_i of the series, when elements are arranged in increasing order, for the corresponding value of the distribution function $F(x_i)$ and its expectance $E[F(x_i)]$, the relation

$$E[F(x_i)] = r(x_i)/(n+1), \tag{1}$$

is true whatever the distribution function $F(x)$ is (*Gumbel*, 1958). It follows that with applying distribution free tests on the series where elements have been replaced by their corresponding rank, exact significance levels are found accounting for all the available information.

3.1 Testing independence of the elements

For verifying the independence of each element with every other ones of a series, the alternative of serial correlation of all orders should be tested.

However, serial correlation of first order involving all alternatives to independence induced by consecutive values persistently either above or below the mean, testing randomness against correlation between elements may be made in using the first order serial covariance or correlation test statistic.

The rank serial correlation test statistic r may be defined as the serial correlation coefficient between the consecutive ranks of each elements of the series, given by the relation

$$r = r[r(x_i), r(x_{i+1})], \tag{2}$$

with $i = 1, 2,\dots, (n - 1)$ or n, with $(n+1) = 1$, according that the non-circular or the circular form is given to the statistic (2). The circular form if the one used in the *Wald-Wolfowitz* (1943) test.

When the alternative to independence is the linear autoregression of first order, the use for the test statistic of the non-circular form instead of the circular form, avoids an obvious loss of power (*Kendall* and *Stuart*, 1966). However, this is less true for the more general alternative of persistence and for long series.

Under the hypothesis of simple randomness, the rank version of the Wald-Wolwovitz test statistic r has asymptotically a normal distribution, with exact means $m_r = -(n - 1)^{-1}$ or 0, whether the test statistic is circular or not, and with variances approximated by var $r = n^{-1}$. When testing persistence, large values of r have to be considered as significant.

Compared to the exact values given for var r by *Bartels* (1982), the error involved in its approximated value by n^{-1}, as well as the bias due to the use of the circular form, is of order n^{-2}.

For short series, the exact probabilities may be derived from the tables given by *Bartels* (1982) for the non circular form of the Von Neumann test statistic.

The asymptotic relative efficiency of the rank test with respect to the most powerful parametric test is .91 (*Knoke*, 1977).

3.2 Testing the stability of distribution

Any statistical distribution may be characterized by a central reference value, such as the mean, the median or any other one, and by the dispersion (scatter) of the deviations of each element of the series around the reference value. It follows that testing the stability of the distribution needs testing the stability of both the reference value and the dispersion.

The trend test used is the one of *Mann* (1945) for which the test statistic t is defined by the summation

$$t = \Sigma\, n_i, \text{ with } i = 1, 2,\dots, n, \tag{3}$$

where n_i is the number of inequalities $x_i > x_j$ for $j = 1, 2,\dots, (i - 1)$, and where each equality $x_i = x_j$ increases n_i by .5.

Under the hypothesis of randomness, the exact mean and variance of t are

$$E(t) = n(n - 1)/4 \quad \text{and} \quad \text{var } t = n(n - 1)(2n+5)/72. \tag{4}$$

Randomness meaning equal probabilities for the inequality $x_i > x_j$ and its reverse, whatever are i and j, the alternatives to the null hypothesis are here all groupings of high or low values at the beginning or at the end of the series, that is instability in the median.

The statistic t has asymptotically a normal distribution giving acceptable critical values for $n > 10$. For $n < 10$, such values may still be used when corrected for continuity. However, for very small sizes, exact probabilities for change-points have to be derived through combinatorial analysis.

The considered alternatives being an increasing or decreasing trend, the trend test has to be applied in the one-sided case.

An alternative trend test is given by Spearman, where the test statistic is the correlation coefficient of the ranks $r(x_i)$ with the corresponding order number i in the time series. The tested property is thus the linearity of the trend. Having asymptotically a normal distribution, the Spearman trend test statistic has 0 mean and approximated variance equal to n^{-1}.

Both tests have an asymptotic relative efficiency of .98 with respect to the corresponding most powerful parametric test (*Kendall* and *Stuart*, 1966). Though apparently equivalent, the tested property in the Mann test makes this last one more appropriate than the first one, when considering the definition of randomness.

3.3 The determination of change-points

When a time series is divided into random sequences having different distributions, the points separating these sequences are change-points. It follows that, at these points, the distribution change may result from a change either in the mean or in the dispersion, or in both mean and dispersion.

In the case of one change-point, its determination may be made by scanning the series in testing for randomness the two partial series into which the complete series is divided by each element x_i, for $i = 2, 3.., (n - 1)$. The estimate of the position of the point is the one for which the probability of stability is simultaneously highest for each of the two partial series defined in this way (*Sneyers*, 1958). With the standardized trend statistics $u(t_i)$ for the series (x_1, x_i) and $u(t'_{i+1})$ for the series (x_{i+1}, x_n), this determination may be made by searching the minimum value of the test statistic (*Sneyers*, 1995)

$$X_i(t) = u^2(t_i) + u^2(t'_{i+1}), \tag{5}$$

which has an approximated Chi-square distribution with two degrees of freedom.

When the partial series defined by the change-point are random, the significance level for the change-point is then derived from the value $u(t_n)$ for the complete series.

In the case of several change-points, using the same method, first estimations may be made in removing, from the beginning or from the end of the series, successively sequences appearing from the sequential u(t) or u(t') values as stable before a systematic increase or decrease of the successive u(t) or u(t') values. Re-estimations are then made in repeating, for each change-point, its one-point determination in the series joining the defining contiguous sequences, onward and backward up to stabilized estimations.

An alternative solution used is to calculate extremes of cumulated sums of deviations from the general mean.

In particular, the statistic derived by *Pettitt* (1979) from the Mann-Whitney two-sample homogeneity test (*Siegel*, 1956), reduces to this cumulated sums applied to ranks. However, the test statistic neglecting a part of information relative to each of the partial series, a lowered power efficiency has to be expected (*Sneyers*, 1997).

It is worth mentioning here, that using the trend method where the Mann statistic has been replaced by the Spearman one, *Easterling* and *Peterson* (1995) have evaluated the power of this method against other techniques. In particular, compared by simulations to the parametric one of cumulative sums of *Alexandersson* (1986), the trend technique appears to have better performances.

As expected, this result may be assigned to the separate account in the test statistic of particular statistical properties for each of the two partial series. However, due to an incomplete account of the internal properties of the series (homogeneity of residual dispersion inside series) and to non-verification of the underlying normality assumption, it is not clear which property is in disagreement with the underlying hypotheses.

3.4 Final significance levels

In the case of several change-points, for making comparisons possible, when having computed the significance level α for each change-point from the trend involved inside the two contiguous i.i.d. series defining it, reduction of α to the size of the complete series has to be made. For the size n of the complete series and the sum of sizes n' of the contiguous partial series defining the change-point, this reduced level α_0 is given by the relation (*Sneyers*, 1975)

$$1 - \alpha_0 = (1 - \alpha)^{n/n'}. \tag{6}$$

On the other hand, the significance of a change-point depends also on the number of repetitions with which its value appears inside the complete series.

To solve the difficulty raised, the solution consists in re-organizing the partial series inside the series with ranging the corresponding means in increasing order. With the application of the scanning method to the new arranged series, the detection of homogeneous sets of partial series and the derivation of the significance levels for the change-points separating each set of partial series becomes possible.

With the reduced significance levels α_0 derived for the differences between the means of the homogeneous groups of sequences, the significance level for the complete set of change-points is then given by the Fisher test statistic

$$X(F) = \Sigma(-\ln \alpha_0), \tag{7}$$

the summation being extended to each of the reduced α_0 values.

Under the null hypothesis of simple randomness, the statistic X(F) has a gamma distribution with shape parameter equal to the number of differences and scale parameter with value 1.

A detailed distribution free or parametric test of variance homogeneity may also be used for the determination of the significance of the different alternatives involved in the rejection of simple randomness.

4. The history of the Oulu air temperature series

The climatological station of Oulu was set up in 1846 by the Finnish Society of Sciences. With geographical coordinates of 65°02 N and 25°29 E and 13 m height above mean sea level, it is located at the border of the Gulf of Bothnia which extends the Baltic Sea at the north. At this location, earlier meteorological observations were apparently made at the request of the Palatine Meteorological Society of Mannheim from 1776 to 1805 (*Heino*, 1994).

The station was initially located in the centre of the city, in an unchanged site up to April 1919. Three changes of thermometer screens took place during the 73 years (1871, 1886, 1905). Several relocations occurred afterwards near the neighbourhood of the first location from May 1919 to December 1971. In January 1972, observations were taken up by the University of Oulu. Observations were then made successively at the old botanic garden up to August 1983 and till now at a new one, 4 km from the city centre.

Daily temperature averages were based on three observations made at 7, 14 and 21 hours up to 1926, at 7, 15 and 21 hours from 1927 to 1946 and at 8, 14 and 20 hours up to now.

The original series of monthly means was computed after adjusting for the differences resulting from the observation time changes and from averaging (weighted averages) method changes, taking the coastal character of the Oulu station into account.

Afterwards, the adjusted series were assessed for observations inhomogeneities. The search for breaks was made in a sequential way from the beginning and from the end of the series, using the *Alexandersson* parametric method (1986), breaks being only accepted when explained by the station history (1971, 1983), by differences in averaging methods (1900, 1926) or in interpolation (1922, 1923, 1956). In this case, the date of change was used for the calculation of adjustments at the monthly scale. For this purpose, each time, at least four reference series known as homogeneous have been used.

Moreover, known relocations or other data calculation changes that produced non-significant inhomogeneities have been ignored.

Evidence of changes in the variances of the differences between Oulu and reference stations were statistically undetectable (*Heino*, 1994). Moreover, testing did not find homogeneity breaks related to screen changes, although according to *Nordli et al.* (1997) changes of screen type can cause small discontinuities especially during spring and summer. During the 19th century, the homogeneity testing was less efficient due to lack of reference stations with very high correlation. At that time, the observation network was sparse. Therefore, some far away Swedish stations had to be used as reference. Also the period 1938-1947 may contain some small non-climatic disturbances caused by frequent station relocations.

At the annual and seasonal scale, the adjustments applied to the original series are summarized for year, winter, spring, summer and autumn in Table 1.

These adjustments show that, in the mean, the last location is colder than the old ones. The obvious reason is that up to 1983, all the locations were influenced by the close proximity of water expanses or flows (sea, river), in contrast with the last one, located further inland out of the city centre. The adjustments due to observation time and averaging changes dominate in summer before 1926.

Table 1. Homogenizing adjustments applied to the original series (°C).

Period	Year	Wi	Sp	Su	Au
1846-1900	-.37/-.38	-.46/-.47	-.36/-.37	-.46/-.47	-.20
1901-1922	-.32/-.33	-.43/-.44	-.30	-.36/-.37	-.20
1923	-.22	-.43	-.14	-.10	-.20
1924-1926	-.32/-.33	-.43/-.44	-.30	-.36/-.37	-.20
1927-1955	-.18/-.19	-.43/-.44	-.13/-.14	+.03/.04	-.20
1956	-.18	-.33	-.13	+.03	-.20
1957-1971	-.18/-.19	-.43/-.44	-.13/-.14	+.03/+.04	-.43/-.44
1972	-.48/-.49	-.60	-.60	-.16	-.43
1973-1982	-.48	-.73/-.74	-.60	-.16/-.17	-.43/-.44
1983	-.32	-.73	-.60	-.16	.00
1984-1995	.00	.00	.00	.00	.00

5. The statistical analysis of the series of annual averages of air temperature at Oulu (1847-1995)

In order to make clear modifications resulting from the homogenizing procedure, the analysis of the original and of the adjusted series has been made step by step simultaneously.

5.1 Testing the complete series for simple randomness

Mean, standard deviation and test statistics for verifying simple randomness are given in Table 2. These first results show a decrease in the mean m, a slight increase in the standard deviation s, for the test statistics u(r) and u(t), poorly significant values for the original series but much higher significant for the adjusted one. Moreover, the intermediate extreme values $u_x(t)$ show extreme ones more divergent from 0 for the adjusted series than for the original one. The values of $\alpha(r,t)$ show that simple randomness may be rejected for both cases, but in a more significant way after the elimination of observational inhomogeneities.

Table 2. Testing simple randomness of the annual air temperature averages series of Oulu (°C) (1847-1995). Results for the original Ori and adjusted Adj series and for the corresponding dispersion series (D).

Data	m	s	r	u(r)	u(t)	$u_x(t)$	$\alpha(r,t)$	$u_d(t)$	$u_{dx}(t)$
Ori	2.019	1.118	.1133	1.45	1.57	2.45/1.95	.028	.02	-/-
Adj	1.754	1.133	.1386	1.77	2.31	3.14/2.53	.003	.65	-/-

For each series, m and s are the mean and the standard deviation, r is the derived serial correlation coefficient, u(r) and u(t) are the standardized rank serial correlation and trend statistics, $u_d(t)$ is the trend statistic for the dispersion of the rank series, $u_x(t)$ and $u_{dx}(t)$ are the extreme values of some importance among the onward and backward sequential values of the trend statistics u(t) and $u_d(t)$, $\alpha(r,t)$ is the joint significance level for u(r) and u(t) derived from the Fischer test statistic.

In contradistinction, for the dispersion, absence of trend may be accepted.

On the other side, in Figs. 1, 2 and 3, for each series, the graphical representation has been given of the values of $E[F(x_i])$ derived from the ranks, of the trend statistics $u(t_i)$ and $u(t'_i)$ and of the Pettitt statistic $X_k(P)$. Besides the great similarity of the graphs in each case, the main difference appears in a smaller scatter of the successive points along a very similar evolution in time (Fig. 2 and Fig. 3).

The first general conclusion is the higher significance of the instability in the mean revealed by the adjusted series.

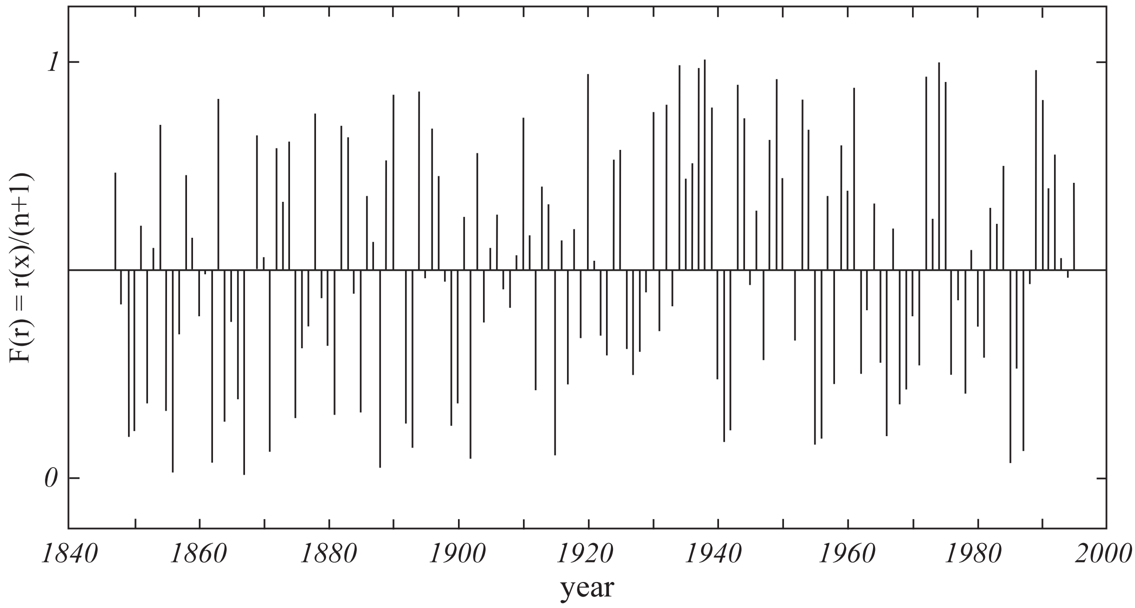
5.2 Search for change-points

The first change-point estimations have been made by removing from the end of the series successively the sets of contiguous elements having u(t') values alternating closely around the zero value. In particular, for the first point 1987, account has been taken of the evolution of u(t'), as well as of u(t) and of the rank values preceding the ones of 1995 up to 1988. The change-points P given in Table 3.1 have been estimated in the same way.

The successive values given in Table 3.1 for the averages m_P for the complete series stopped successively in i = P show that, exception made for the successive increases of the m_P values in 1872/1881 and 1933, all the other ones have alternating differences. This explains the alternating values found for the trend statistics $u(t_i)$, for the complete series stopped in i = P, and for $u(t_P)$, for the joined contiguous partial series defining the point P. On the other hand, the acceptance of the relatively low $u(t_P)$ values is justified either by the persistence of positive or negative $u(t_i)$ values for long series or, in the case of very short series, by the exact significance level determined by combinatorial considerations.

In example, for 1984, all the ranks for 1982, 1983 and 1984 are smaller than the ones for 1985, 1986 and 1987. The exact probability for $u(t_P)$ is thus $(3!3!)/6! = .05$.

Annual temperature averages in Oulu (1847-1995) - Original observations



Annual temperature averages in Oulu (1847-1995) - Adjusted observations

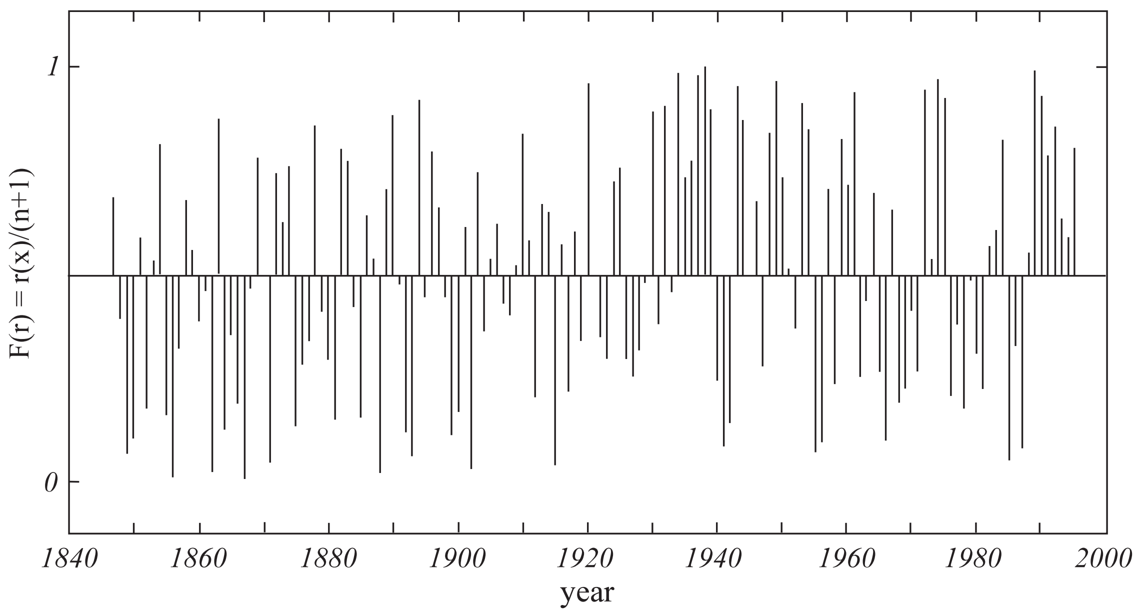
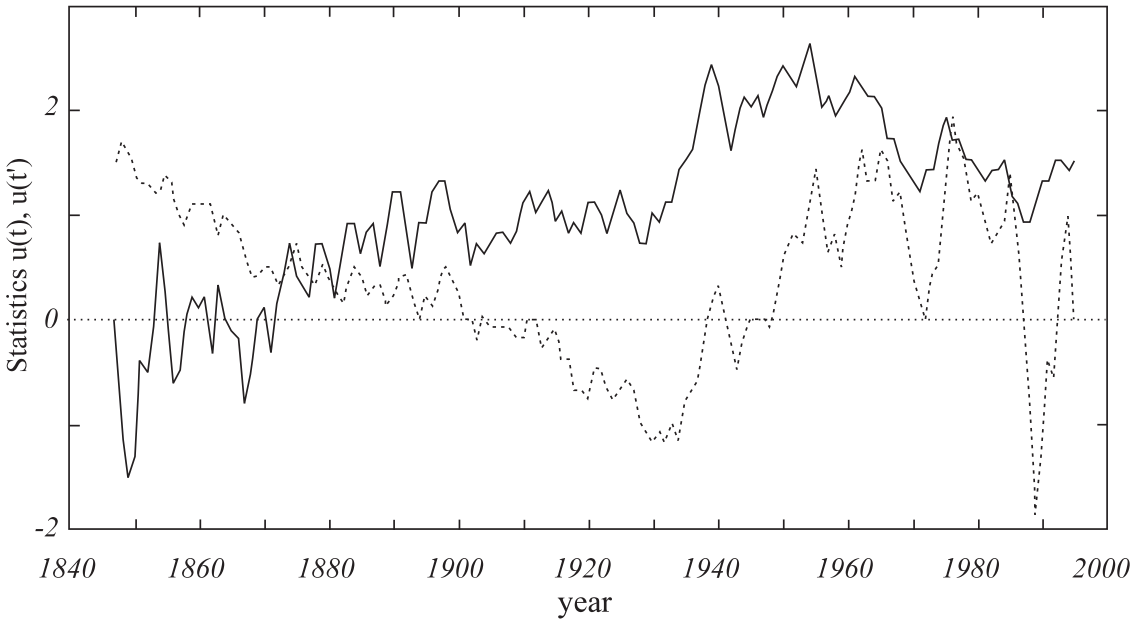


Fig. 1. Expectance F(r) derived from the rank $r(x_i)$ of the element x_i.

Annual temperature averages in Oulu (1847-1995) - Original observations



Annual temperature averages in Oulu (1847-1995) - Adjusted observations

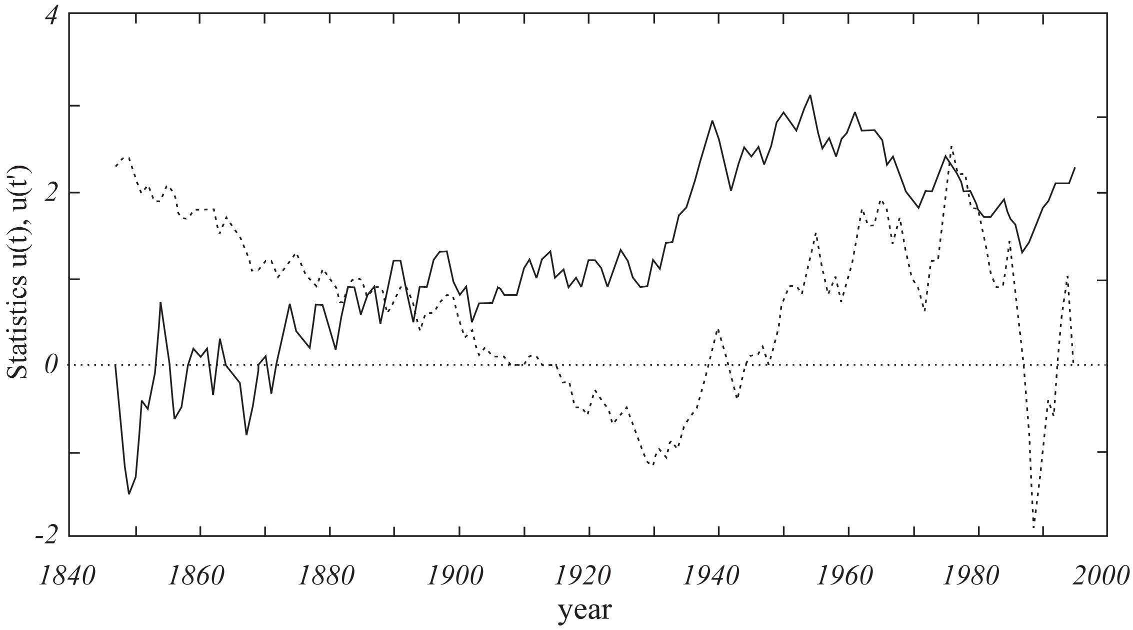
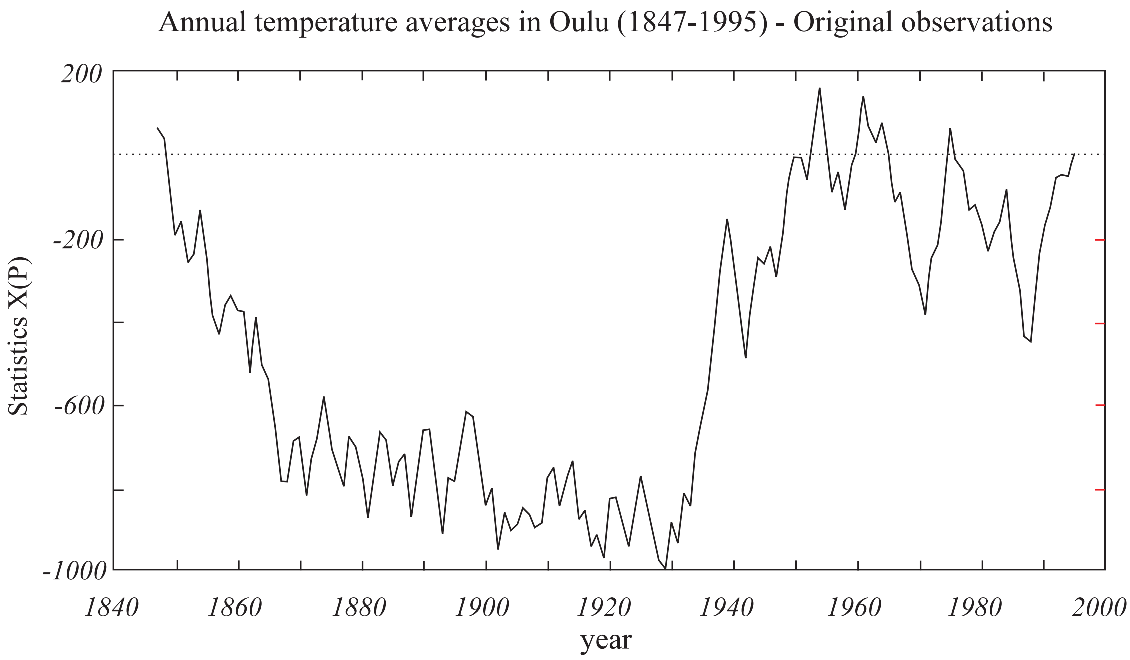


Fig. 2. Standardized trend statistics u(t) and u(t') for the series (x_1, x_2, ..., x_i) and for the series (x_i, x_{i+1}, ..., x_n).



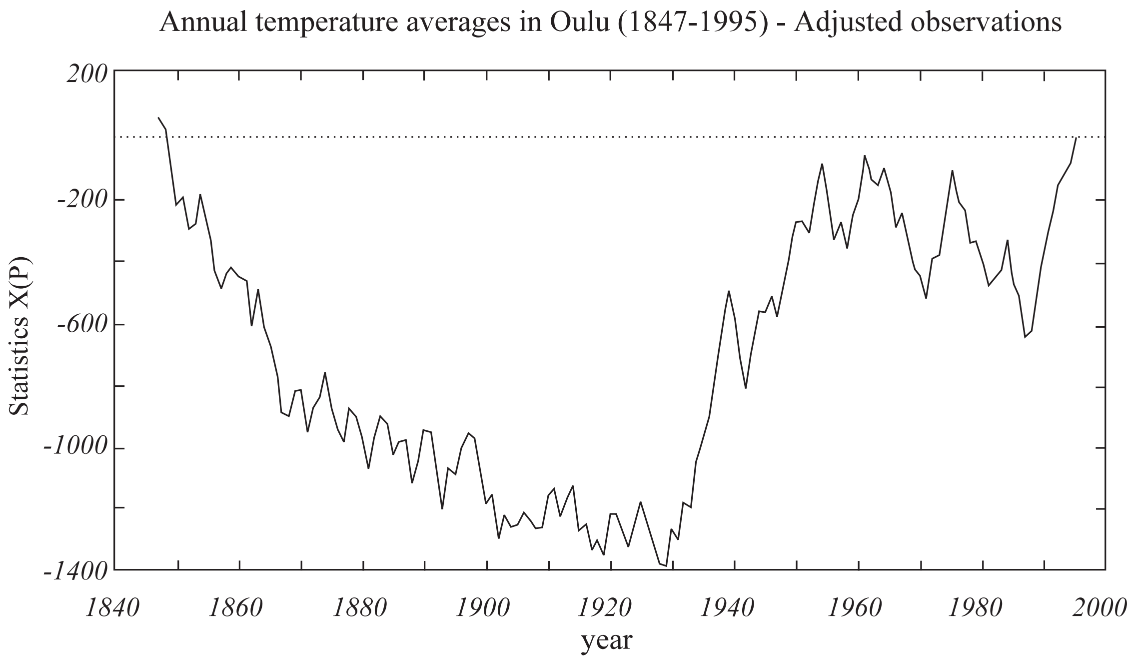


Fig. 3. Successive values of the Pettitt statistic X(P).

The final justification for the change-point estimates has been made by ranging with increasing means all the partial series defined by the change-points.

Analysing the new arranged series in the same way as made for the first ones, leads to the grouping of partial series given in Table 3.2.

Table 3.1 Change-point determination.

	ORI			ADJ		
P	$u(t_i)$	$u(t_P)$	m_P	$u(t_i)$	$u(t_P)$	m_P
1872/1881	.11	1.11	1.59	.27	1.42	1.30
1933	1.11	1.70	2.03	1.42	1.76	1.73
1939	2.44	-1.67	3.55	2.81	-1.67	3.36
1942	1.69	1.53	.64	2.08	1.53	.45
1954	2.60	-1.71	2.69	3.14	-1.71	2.51
1958	1.95	2.55	1.05	2.49	2.55	.87
1961	2.37	-2.07	3.05	2.95	-2.09	2.87
1971	1.20	1.81	1.48	1.83	1.75	1.30
1975	1.93	-2.06	3.57	2.48	-2.06	3.08
1981	1.34	2.09	1.62	1.72	2.29	1.13
1984	1.55	-1.72	2.62	1.95	-.94	2.35
1987	.92	1.48	.34	1.38	1.48	.34
1995	1.57		2.74	2.31		2.74

For the original series (ORI) and the adjusted one (ADJ), estimated change-point P, trend statistic $u(t_i)$ for the partial series beginning in 1847 and ending in i = P, trend statistic $u(t_P)$ for the joint partial series defining the change-point P, mean m_P for the partial series ending in P.

Table 3.2 Randomness for sets of partial series found as homogeneous in the median for the original (ORI) and the adjusted series (ADJ). Arrangement with increasing means.

Series	n	u(r)	u(t)	$u_d(t)$	m	s	N	s_0	t_s	α_0
ORI										
1987, 1942, 1958	10	-1.32	1.16	-1.34	.711	.873	.886	.991		
1971, 1872, 1981	42	-1.31	.25	.03	1.572	1.051	.963		2.47	.021
1933	61	-.64	.09	-1.04	2.031	.965	-.146*		2.31	.016
1984, 1954, 1995	23	.52	-.05	-.05	2.698	.666	.985	.661	3.56	4E-4
1961, 1939, 1975	13	-.33	.98	-.12	3.438	.653	.923		3.23	.006
ADJ										
1987, 1942, 1958	10	-1.04	.89	-1.07	.585	.849	.857	.993		
1981, 1881, 1971	51	-.59	.25	-.43	1.278	1.025	-.550*		2.02	.055
1933	52	-.78	-.06	-1.16	1.733	.984	-1.00*		2.33	.016
1984, 1954, 1995	23	.59	.88	-.61	2.568	.679	.989	.669	4.26	4E-5
1961, 1975, 1939	13	-.17	.61	-.61	3.162	.650	.981		3.60	.002

For each set of series, total size n, trend and serial correlation test statistic u(t) and u(r), trend statistic $u_d(t)$ for dispersion, mean m and standard deviation s, test statistic N for normality, final estimation s_0 of standard deviation. N values are given by the Shapiro-Wilk test statistic, exception made for values with *, when derived from the D'Agostino test statistic. Student test statistic t_s for the difference between contiguous m values.

The computation of the trend statistic for median and dispersion, as well as of the serial correlation statistic of the series formed by each group gives u(t) and $u_d(t)$ as well as u(r) values sufficiently near to zero for accepting the null hypothesis of simple randomness. Similarly, after reduction to the size of the complete series, the probabilities derived from the test for normality, make the normality assumption, and thus further parametric testing, acceptable.

In particular, for the original (ORI) and adjusted series (ADJ), applying the likelihood ratio test of variance homogeneity (*Mood*, 1950) for the standard deviations s of the three first set of series in Table 3.2, leads to Chi-square values having respectively the probabilities .27 and .22. Similarly, testing homogeneity for the two last standard deviations s, the F-test gives the probabilities .51 and .54. Acceptance of the tested homogeneity leads, respectively to the common standard deviations .991 and .993, for the three first standard deviations, and .661 and .669, for the two last ones. On the other side, the F-test applied on the common standard deviations for the three first and two last ones gives the probabilities .996 and .995 rejecting the null hypothesis of homogeneity for the variances of the complete set of series.

For the means, homogeneity of consecutive ones in Table 3.2 has been tested with applying a Student test.

From the significance levels reduced to the complete size of the series, the Fisher statistic $-\Sigma \ln \alpha_0$ gives for the original series and the adjusted one respectively 18.7 and 23.2. For a gamma distribution with shape parameter 4 and scale parameter 1, it means high significance against homogeneity of the temperature means for each case. This significance is notwithstanding higher for the adjusted series as for the original one.

5.3 Final characterization of the temperature evolution

Taking account of these last results gives in Table 3.3 the estimated evolution for the air temperature in each of the two series.

The computed shifts between consecutive partial series have standard deviations amounting respectively to 1.49 and 1.67 for each case.

Moreover, removing the variance of the partial series from the variance of the complete series gives the variance part explained by change-points. Due to the adjustment for inhomogeneities, this part increases from 31% to 33% of the variance of the complete series.

Finally, it appears that in this evolution, negative shifts correspond all to sequences with large variances, while to positive shifts correspond smaller ones.

The two first results show that, if an increase of 7% is found for the explained variance by change-points, for the variance of the shifts, this increase amounts to 26%. It follows that if adjustment for inhomogeneities has but a poor effect in the estimation of the change-point positions, it has increased the importance of the involved temperature instability.

The third result reveals that cold sequences are bound to weather types having a larger variability than the ones bound to warm sequences. The two first sequences are significantly longer than the ones following 1933.

Concerning the values found for the differences δ_m, a trend analysis applied to each series gives the values of Table 3.3. They show non significant values for the trend statistic u(t), but relatively high significant values for u(r), favorable to the alternative of alternance, this alternance being more significant for the adjusted series as for the original one.

From these results, it may be concluded that the climate evolution at Oulu has essentially been characterized by a drastic change in the persistence of weather types involving a non-identifiable climate warming.

Table 3.3 Final characterization of the temperature evolution.

			ORI			ADJ	
P	n	m	s	δ_m	m	s	δ_m
1872/1881	26/35	1.57	.991		1.28	.993	
1933	61/52	2.03	.991	.46	1.73	.993	.46
1939	6	3.44	.650	1.53	3.16	.666	1.52
1942	3	.71	.991	-2.86	.58	.993	-2.67
1954	12	2.70	.650	1.32	2.57	.666	2.02
1958	4	.71	.991	-1.32	.58	.993	-2.02
1961	3	3.44	.650	1.32	3.16	.666	2.02
1971	10	1.57	.991	-.46	1.28	.993	-1.32
1975	4	3.44	.650	1.99	3.16	.666	1.97
1981	6	1.57	.991	-1.99	1.28	.993	-1.97
1984	3	2.70	.650	.46	2.57	.666	1.23
1987	3	.71	.991	-1.32	.58	.993	-2.02
1995	8	2.70	.650	.86	2.57	.666	2.02
m				.00			0.10
s				1.56			1.89
s_m				.45			.55
u(t)				-.21			.27
u(r)				-2.09			-2.60

For the homogeneous partial series ORI and ADJ, end and change-point P, size n, final estimations of mean m and standard deviation s, shifts in the mean δ_m at the change-point P. Values of m, s, u(t) and u(r) defined as above. s_m is the standard deviation for the mean m.

6. Correlation with the NH global land air temperature average series (1856-1995)

Global temperature averages have been recomputed by *Jones* (1994) as well for land as for sea surface for both hemispheres (NH and SH). In the particular case of Oulu, for determining the degree of correlation between local land and global averages, the NH land series seemed us to be here the most appropriate. Such determination allows here to determine how far conclusions at the NH scale are valid for this local scale or, inversely, how much this local scale influences the considered NH one.

Note that this correlation gives also a measure of the representative value of this local station at the regional one.

Therefore, the correlation coefficients between the NH land series and the Oulu one have been computed for the complete series and for each of the i.i.d. partial series into which the NH land series has been divided by change-points (*Sneyers*, 1997). Moreover, account has here been taken of the weak significant 1884 and 1892 change-points in the

NH series, these points resulting from some change-points having occurred at the local scale.

Results of the computation of the correlation coefficients for the two stations are given in Table 4.

Table 4. Correlations of the original (ORI) and of the adjusted (ADJ) series with the complete and the partial NH series (1856-1995).

	ORI		ADJ	
Series	R	R^2	R	R^2
1856-1995	.380	.145	.428	.183
1884	.596	.355	.596	.356
1892	.861	.742	.860	.740
1918	.275	.076	.277	.076
1929	-.015	.000	-.044	.002
1949	.469	.220	.470	.221
1976	.210	.044	.221	.049
1987	.086	.007	.044	.002
1995	.148	.022	.148	.022
R average	.345	.119	.342	.117

Values for the correlation coefficients R for the complete series, the NH partial series and the weighted average of R. For R and the weighted R averages, squares R^2 of R.

Comparing the results derived respectively from the original and from the adjusted Oulu data, it appears that for the coefficients R, the single obviously modified value is the one for the complete series, which shows a correlation increase. Noting that R^2 gives in % the part of variance common to the two correlated variables, this means that this common part increases from 14.5% for the original Oulu data to 18.3% for the adjusted one.

In the same way, the weighted averages of the values of the correlation coefficient R for each partial NH series estimate the common part of variance to 11.9% and to 11.7% for each series. This common part neglecting the existence of the NH change-points, this means an increase of the total common part of variance from 14.5-11.9 = 2.6% to 18.3-11.7 = 6.6%, to be assigned to the homogenizing procedure.

The smallness of this part assigned to the existent NH change-points has to be related to the absence of a significant increase of temperature means.

7. *Concluding remarks*

The final remarks may be summarized as follows:

(1) The climate change-point positions in the Oulu temperature series seem to have not been much perturbed by observations inhomogeneities. However, as expected, their degree of significance increased in an important way. Moreover, they led to stronger correlations with the global NH temperature averages. Exception made for the first change-point, all the other ones, as well as the randomness of the partial series

remained unmodified after the homogenizing procedure. The increase of the correlation with the NH series may thus be assigned to better estimates of the shifts.

(2) Due to these improvements, a clearer view was given of the Oulu climate for what concerns temperature. For this point, it is worth mentioning that, as found for global, regional or other local series, the Oulu climate evolution is characterized by i.i.d. sequences separated by unexpected shifts (Sneyers, 1997). It follows that, the increasing or decreasing trends detected do not result from a progressive warming or cooling, but from particular arrangements of i.i.d. partial series inside the complete one or from varying persistence inside these series. In particular, part of correlation with the global NH temperature series appears to be due to particular arrangements of the partial series like successive increasing means or to special places occupied by the partial series having lowest or highest mean.

It is to be noted, that at the seasonal scale and exception made for spring, which presents but two increasing shifts, alternance of shifts is present in every of the three other seasons.

Actually, the climate summaries of the weather evolution suggest that changes in the weather type frequencies explain largely such statistical results (*Wessels et al.,* 1994, *Ashcroft*, 1994) and examples do exist showing the direct relation between special climate evolutions and persistent general atmospheric circulation characteristics (*Wallace et al.,* 1996).

On the other hand, the atmosphere and the ocean general circulations (G .C.), from which the weather and thus the climate evolution depends, are ruled by differential non-linear equations. As this non-linearity may be responsible for a chaotic instability of the weather evolution, it strongly suggests that a similar chaotic instability of the climate might be but an intrinsic property resulting from this non linearity of the G.C. differential equations (*Sneyers*, 1997).

It follows that in the case of climate evolution, the validity of its explanation or prediction, depends not only of the accuracy with which physics are integrated in the G.C. models (Quinet and Vanderborght), but also of the full account of the statistical properties generated by the mathematical part of these models.

(3) At the methodological point of view, the present analysis has given an example of efficiency for a statistical analysis leading to reliable significance levels. Statistical analysis being a mathematical procedure, as for any applied mathematical model, in all its rigour, it needs verifying all underlying hypotheses before going ahead. In particular, the asymptotic properties given by the Central limit theorem, generally used for justifying the systematic use of parametric tests, ignores the degree of error due to divergences existing in the finite case between the real distribution and the normal one (*Pearson*, 1971). Any deviation from the correct way involves thus an unmeasurable risk of uncertainties concerning the derived results.

Acknowledgements

We are indebted to Mr Folland, Head of the Climate Variability Research Group (Hadley Centre, Met. Office, UK) for the full communication of the most recent Jones global temperature averages, to Prof. Dr. A. Quinet for useful suggestions and to Ms L. Frappez for the computer realization of the graphs.

References

Alexandersson, H., 1986. A homogeneity test applied to precipitation data. *J. Climatol.* **6**, 661-675.

Ashcroft, J., 1994. The pattern of temperature change in autumn in the United Kingdom in the period from 1931 to 1990. Climate variations in Europe, *Proceedings of the European Workshop on Climate Variations*, Raino Heino (Ed.), Kirkkonummi (Majvik), Finland, 15-18 May 1994, 156-166.

Bartels, R., 1982. The rank version of von Neumann's ratio test for randomness, J. Am. Statist. *Assoc*., **77**, 40-46.

Conrad, V., 1925. Homogenitätsbestimmung meteorologischer Beobachtungen. *Met. Zeitschr*., **42**, 482-485.

Dettwiller, J., 1970. Deep soil temperature trends and urban effects at Paris. *J. Appl. Meteor*., **9**, 178-180.

Easterling, D.R. and T.C. Peterson, 1995. A new method for detecting undocumented discontinuities in climatological series. *Int. J. Climatology*, **15**, 369-377.

Gumbel, E.J., 1958. Statistics of extremes. Columbia University Press, N.Y., 374 pp.

Heino, R., 1992. Changes of climatic characteristics in Fennoscandia. 5th International Meeting on Statistical Climatology, Toronto, 22-26 June 1992, Environment Canada, 101-104.

Heino, R., 1994. Climate in Finland during the period of meteorological observations. *Finnish Meteorological Institute. Contributions*, No **12**, 212 pp.

Jones, P. D., 1994, Hemispheric surface air temperature variations. A reanalysis and an update to 1993. *J. Clim*. **7**, 1794-1802.

Karl, T.R., F.D. Henry and G. Kukla, 1988. Urbanization: Its detection and effect in the United States record. *J. Climate*, **1**, 1099-1123.

Kendall, M.G. and A. Stuart, 1966. The advanced theory of statistics, Vol. 3, Griffin, London, 552 pp.

Kendall, M.G., 1971. Studies in the history of probability and statistics, XXVI: The work of Ernst Abbe, *Biometrika*, **58**, 369.

Knoke, J.D. 1977. Testing for randomness against autocorrelation: alternative tests. *Biometrika*, **64**, 523-529.

Lewis, L. F. 1947. Variations of temperature in London 1764-1939. Meteor. Mag., 76, 135-138.

Lysgaard, L., 1948. Recent climatic variations. *Nature* **161**, 442-443.

Mann, H. B., 1945. Non parametric test against trend. *Econometrika*, **13**, 245-259.

Mood, A. M., 1950. Introduction to the theory of statistics. McGraw Hill, 434 pp.

Nordli, O., H. Alexandersson, P. Frich, E. Forland, R. Heino, T. Jonsson, H. Tuomenvirta and O.R. Tveito, 1997. The effect of radiation screens on Nordic time series of mean temperature. *Int. J. Climatol*. (in print).

Pearson, E.S., 1971. Power of parametric tests in the case of non normal distributions. (Unpublished paper presented at a meeting of the Technical Committee ISO/TC 69 "Application of Statistical Methods" of the International Organisation for Standardization (ISO).

Pettitt A.N., 1979. A non-parametric approach to the change-point problem. *Appl. Statist.* **28**, 126-135.

Quinet, A. and J. Vanderborght, 1996. Clear-sky nocturnal temperatures forecast and the greenhouse effect. *J. Appl. Meteor*. **35**, 3, 401-415.

Siegel, S. 1956. Nonparametric statistics for the behavioral siences. McGraw-Hill, New-York, 312 pp.

Sneyers, R., 1955. Sur la determination de l'homogénéité des séries climatologiques. *J. Scientifique de la Météorologie*, VII, **28**, 359-372.

Sneyers, R., 1958. Connexions thermiques entre saisons consécutives à Bruxelles-Uccle. *Institut R. Météorol. de Belgique*, Publ. B, No **23**, 23 pp.

Sneyers, R., 1963. Sur la détermination de la stabilité des séries climatologiques. Recherches sur la zone aride. Les changements de climat. *Actes du colloque de Rome organisé par l'Unesco et l'Organisation météorologique mondiale*, 37-44.

Sneyers, R., 1969. Notes sur la notion d'indépendance climatologique. *Rev. Stat. Appl*., **7**, 4, 45-53.

Sneyers, R., 1972. Sur une méthode statistique de détection des observations aberrantes du réseau thermométrique Belge. *Arch. Meteor., Geoph. u. Bioclim*., **20**, 3-4, 269-278.

Sneyers, R., 1975, Sur l'analyse statistique des séries d'observations. O.M.M., N.T. 143, Genève, Suisse, 210 pp. English version, W.M.O., 1990.

Sneyers, R., 1989. On the homogeneity of long series of observations. Proceedings. Fourth International Meeting on Statistical Climatology, 27-31 March, 1989, Rotorua, New Zealand, 21-23.

Sneyers, R., 1995. Climate Instability Determination. Discussion of Method and Examples. 6th International Meeting on Statistical Climatology, June 19-23, 1995, Galway, Ireland, 547-550.

Sneyers, R., 1997. Climate Chaotic Instability. Statistical Determination and Theoretical Backgrounds. *Environmetrics*, **8**, 517-532.

Snijders, T.A.B., 1984. Antithetic variates for Monte Carlo estimation of probabilities. *Statistica Neerlandica*, **38**, 22-73.

Tuomenvirta, H. and A. Drebs, 1994. Homogeneity testing and management of metadata in Finland. In: Heino, R. (ed), Climate Variations in Europe. Proceedings of the European Workshop held in Kirkkomuni (Majvik), Finland, 15-18 May 1994. Publications of the Academy of Finland 3/94, 321-329.

Tuomenvirta, H. and R. Heino, 1996. Climatic changes in Finland - recent findings. *Geophysica*, **32**(1-2), 61 75.

Vandenplas A., 1948. Variation séculaire de la température de l'air à Bruxelles-Uccle. *Inst. R. Météorol. de Belgique*, Misc. **35**, 16 pp.

Wald A. and J. Wolfowitz, 1943. An exact test for randomness in the non-parametric case based on serial correlation. *Ann. Math. Statist*., **14**, 378-388.

Wallace, J.M., Zhang, Y. and Bajuk, L., 1996. Interpretation of interdecadal trends in Northern Hemisphere Surface Air Temperature. *J. Climate*, **9**, 2, 249-259.

Wessels, H.R.A., A.M.G. Klein Tank and G.P. Können, 1994. Trends in the annual temperature in de Bilt: warmer atmosphere or changed advection?. Climate variations in Europe, Proceedings of the European Workshop on Climate Variations, Raino Heino (Ed.), Kirkkonummi (Majvik), Finland, 15-18 May 1994, 152- 155.