ENVIRONMETRICS, VOL. 8, 469–483 (1997)

SOME PROBLEMS WITH APPLICATION OF CHANGE-POINT DETECTION METHODS TO ENVIRONMENTAL DATA

DANIELA JARUŠKOVÁ

Department of Mathematics, Faculty of Civil Engineering, Czech Technical University, Prague, Czech Republic

SUMMARY

This paper summarizes the author's experience in researching methods of discovering a change in the behaviour of meteorological and hydrological series. Basic statistical tests applying 'maximum' type statistics to detect a sudden or gradual change in location are given. The author stresses that the characteristic properties of the meteorological and hydrological data, especially the dependence between neighbouring observations, have to be considered by performing statistcal tests for change-point detection.

Environmetrics, **8**, 469–483 (1997)
No. of Figures: 10. No. of Tables: 3. No. of References: 38.

KEY WORDS hydrological and meterological data; non-stationarity; change-point problem; sudden and gradual type of change

1. INTRODUCTION

Scientists are afraid that the impact of human activity on nature may cause a change in nature. This apprehension precipitated the launching of the world project to study important hydrological and meteorological series. The contribution of the Czech Republic to the *World Climatological Program – Water* consisted of two projects:

(i) Analysis of long hydrometeorological series.
(ii) Analysis of hydrological observations in the Czech Republic.

These projects were directed by the researchers of the Czech Hydrometeorological Institute. The goal was to create a statistical software package for detection of non-stationarity in time series and to apply this software to some hydrological and meteorological series. The desire was to create a program that would not be difficult to understand and to use even for non-statisticians. As I was known to be interested in problems of change-point detection, I was invited to suggest appropriate statistical methods and to check whether all procedures used were mathematically

* Correspondence to: D. Jarušková, Department of Mathematics, Faculty of Civil Engineering, Czech Technical University, Prague, Czech Republic

CCC 1180–4009/97/050469–15$17·50

Received 20 August 1996
Revised 28 February 1997

correct. In the course of my co-operation with the Czech Hydrometeorological Institute I had the opportunity to study several hydrometeorological series:

(i) the air temperature series;
(ii) the water temperature series;
(iii) the precipitation series;
(iv) the water discharges series.

Later on, researchers from different fields within the Czech Republic, as well as from abroad, were sending me their data for analysing whether their series could be considered stationary or not. For example, I analysed, among others, the series of total ozone amount measured by the Czech Meteorological Institute in Hradec Králové, the influence of woodcutting on the series of water discharges that was originally studied by researchers of the Institute of Forest Management in Frýdek – Místek, and the air pressure series measured by the Swiss Meteorological Institute in Zurich. In this way I gained a lot of experience with the performance of different statistical methods and tests for change-point detection. I also had the opportunity to observe some features typical of the behaviour of hydrometeorological series as well as other series arising from the measurement of some natural processes in the environment. I realized that despite many new results in the study of change-point detection, practical problems can often be encountered that have not yet been solved. In this paper, I would like to illustrate some of these problems and to warn against mistakes that might be made.

Before creating the statistical software for change-point detection we have to answer several questions. First, what statistical properties do the series under study possess? Second, what kind of inhomogeneities or non-stationarities are we looking for? Third, which statistical methods shall we apply? Of course, the answers are not unique. In what follows I present my answers to these questions and I am aware that the answers of many other researchers to the same problems may differ substantially.

2. PROPERTIES OF HYDROMETEOROLOGICAL DATA

2.1. Seasonality

As far as I know, the basic hydrometeorological data are monthly observations. Unfortunately, they usually have a strong seasonal character. The easiest way to handle the seasonality is to subtract from January's data the overall January average, from the February's data the overall February average etc. One could object that the deseasoning can influence the detection of a change in the global mean. For the series I was dealing with the practical impact of the deseasoning on the detection of a global change was small. However, my strong opinion is that for most problems of searching for inhomogeneities in hydrometeorological series, the annual averages are more appropriate.

2.2. Skewed distribution

The monthly water discharges are often skewed. The statisticians analysing the hydrometeorological data often transform the observations, usually using the logarithmic transformation, remove the seasonality and then use tests for normally distributed random variables. Sometimes the interpretation of a change in parameters causes problems as the mean and the variance of

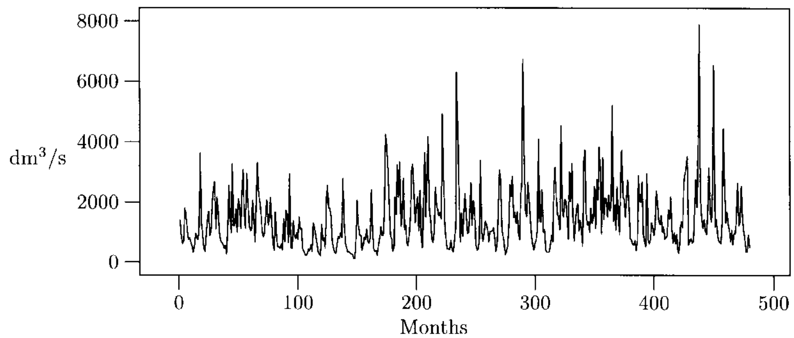

Figure 1. Monthly water discharges of Načetínský Creek, years 1951–1990

log-normal distribution are functions of both parameters. Some time ago we studied the monthly averages of water discharges of a small creek called Načetínský in the Erzgebirge mountains (see Figure 1).

The forest in the Erzgebirge mountains was heavily damaged by acid rain. In this situation one might expect change in variation of runoffs as soil loses the capability for water retention and in rainy periods the water discharges are large and in dry periods they are small. The transformed deseasonalized series is shown in Figure 2. The tests detected the change in the mean but not in the variance of the transformed data. Therefore, we concluded that the scale characteristic changed but the shape characteristic of the original series remained the same.

It is well known that averaging reduces the skewness of the data. If we deal with the annual averages instead of monthly averages the problem of skewness is usually not so serious. Let us take the example of the water discharges of the river Labe measured in Děčín (Czech Republic). The basic descriptive statistics are given in Table I.

2.3. Dependence

It is well known that there exists a certain persistance in the behaviour of nature. This persistance is manifested by a dependence between the meteorological observations which are close in time. The series with quickly decreasing correlograms are often modelled by ARMA sequences.

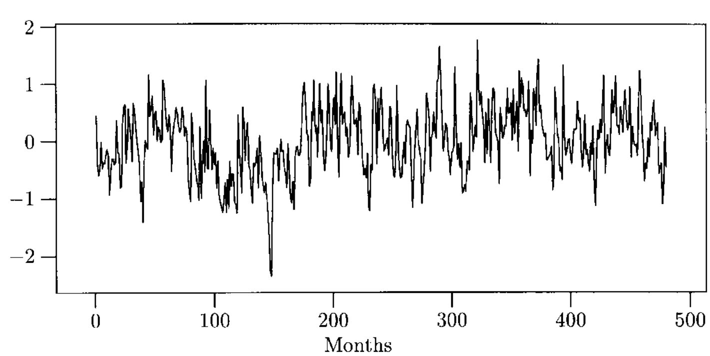

Figure 2. Transformed deseasonalized series of monthly water discharges of Načetínský Creek, years 1951–1990

Table I. Descriptive statistics for monthly and annual water discharges of Labe in Děčín measured in 1851–1989

	Average	Standard deviation	Coefficient of skewness
November	222·658	139·995	1·881
December	277·628	189·330	1·736
January	311·940	210·630	1·870
February	386·968	237·220	1·100
March	519·315	267·562	1·317
April	501·754	259·748	1·444
May	344·584	164·657	1·468
June	260·901	170·763	3·222
July	222·674	147·234	2·150
August	199·247	127·236	1·944
September	191·594	148·636	3·301
October	205·346	121·595	2·091
year	303·742	94·003	0·719

Consider the example of *global world temperature anomalies* compiled by Jones *et al.* (1994) (see Figure 3).

If we suppose that the first part of the series (years 1854–1903) is stationary (for in the 19th century emission of gases that contribute to the greenhouse effect was low), then the values of the sample autocorrelation function of this series calculated for lags $t = 1, 2, 3, 4, 5$ are the following: $ar(1) = 0{\cdot}237$; $ar(2) = 0{\cdot}0227$; $ar(3) = -0{\cdot}026$; $ar(4) = 0{\cdot}005$; $ar(5) = -0{\cdot}006$. Hence, the first part of series of global world temperature anomalies is one of the examples where a short-memory model could be applied.

On the other hand, Lawrance and Kottegoda (1977) mentioned that hydrologists, who study rainfall and riverflow series, sometimes observe long periods of very low or very high flows, the so-called Joseph effect. This effect causes failure of the correlogram to die out. Then, the long-memory models are usually applied.

The behaviour of a process with strong dependent observations (where the value of the spectral density at zero is large) is 'lazy' in the sense that the process can stay for a certain period on one level and then slowly change to another level, therefore the change must be more apparent to be

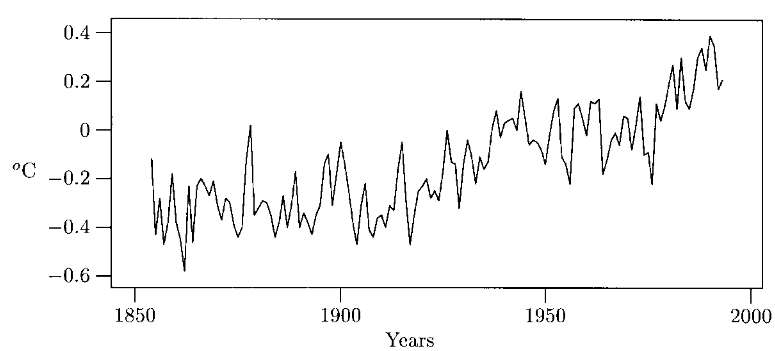

Figure 3. Global annual temperature anomalies (ENSO – uncorrected), years 1854–1993

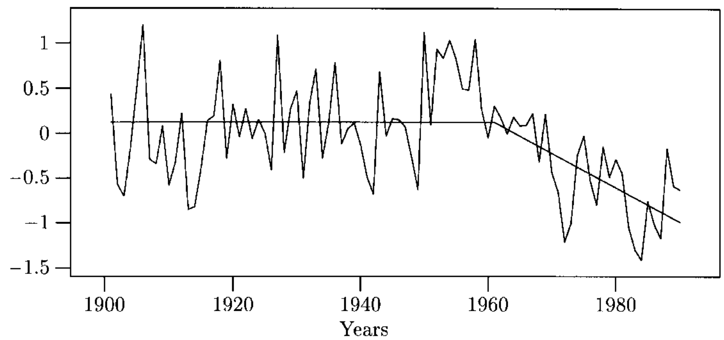

Figure 4. Standardized annual rainfall departures in Sahel, 1901–1990

discovered. I would like to stress that the type of assumed dependence largely influences the decision about the existence of non-stationarities in the series.

2.4. How to establish properties of the studied series?

Statisticians usually provide the researchers studying real data with a battery of tests for change-point detection, i.e. with parametric and non-parametric tests, as well as tests for the independent and dependent observations. It would be easy to decide which test to use if the researchers knew that a certain part of a series is definitely stationary. Unfortunately, such a situation is very rare. The procedure to discover the properties from the observed data is misleading because if there are inhomogeneities in the series, the behaviour of basic statistical characteristics such as the autocorrelation function is 'weird'; see Figure 4 presenting the annual rainfall departures in Sahel constructed by Nicholson (1994) and the corresponding autocorrelation function in Figure 5.

The effect of a change in parameters on the sample autocorrelation function as well as on other sample characteristics as the coefficient of skewness or kurtosis is well known to all statisticians. The non-statisticians may easily make the mistake of using the original measurements for the model setting.

Sometimes, the researchers can use their knowledge about properties of some series that are similar to the series under study due to the similar environmental conditions. Of course, this approach requires a long experience with the hydrometeorological data.

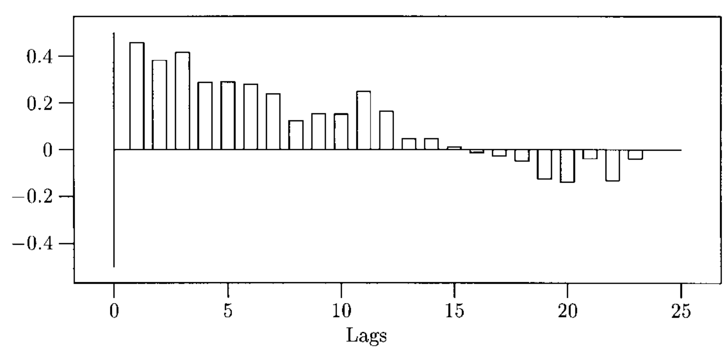

Figure 5. Autocorrelation function of standardized annual rainfall departures in Sahel, years 1901–1990

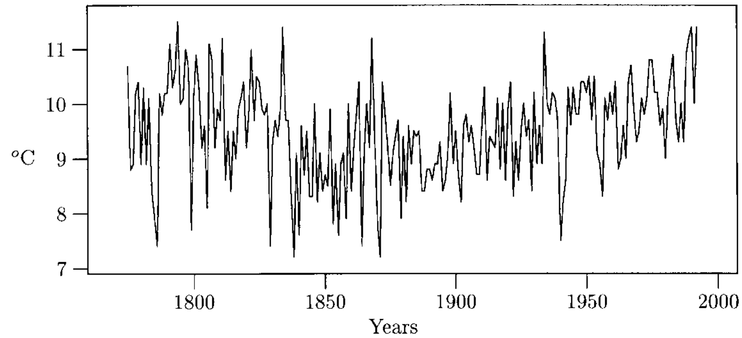

Figure 6. Annual temperature series measured in Klementinum, Prague, 1775–1992

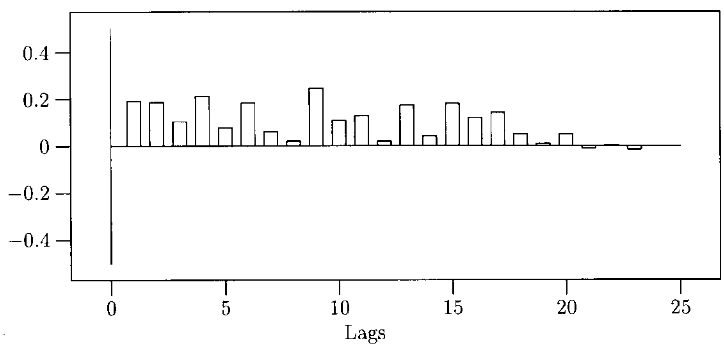

Figure 7. Autocorrelation function of the first part of the annual temperature series measured in Klementinum, Prague, years 1775–1900

Finally, I would like to add a reminder that if the finite part of a time series is observed, it is impossible to distinguish between a stationary series with the positive dependence between the neighbouring observations and a sequence of independent variables with the slowly changing mean. Consider the long temperature series measured in Klementinum, Prague (see Figure 6).

If we suppose (for the same reason as we gave for the series of *global world temperature anomalies*) that the first part of the series (years 1775–1900) is stationary and calculate its correlogram (see Figure 7), then the long-memory model might be applied. On the other hand, some statisticians might prefer the model with a change occurring between years 1835–1840. We conclude that the choice of the model is always subjective. On the other hand, some statisticians might prefer the model with a change occurring between years 1835–1840. We conclude that the choice of the model is always subjective.

3. NON-STATIONARITY

What does it mean that a process is non-stationary? The question is not easy because the stationarity of a process can be violated in many ways. However, hydrometeorologists usually expect that the change occurs either in location or in variation or in both. The increase of temperature due to global warming is one example of a change in location. The increase of variance of water discharges caused by the diminishing capability for water retention of the soil

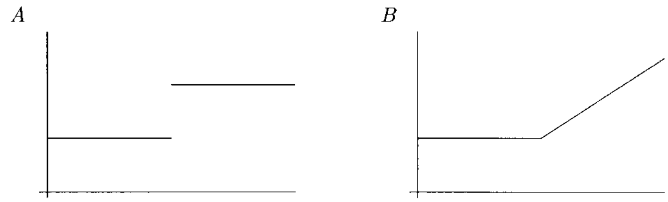

Figure 8. Sudden change of type A, and continuous change of type B

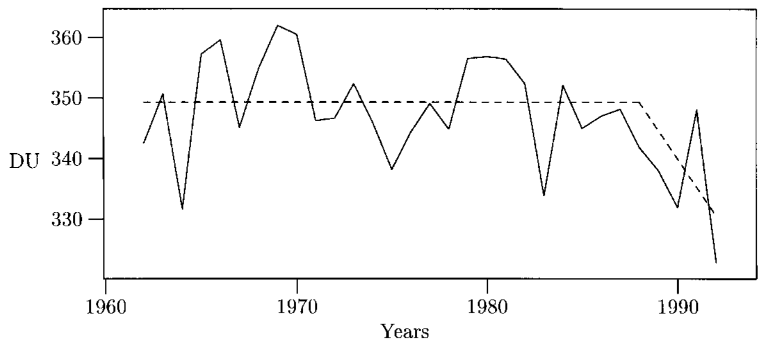

Figure 9. Total ozone amount measured in Hradec Králové, 1962–1992

due to deforestation serves as an example of a change in variation. In the change-point literature the sudden change of type A is usually studied. The scientists object that they more likely expect the continuous change of type B (see Figure 8). It seems that such a change may be observed in the series of *total ozone amount* measured in the solar observatory of the Czech Hydrometeorological Institute in Hradec Králové (see Figure 9) or in the series of *global world temperature anomalies* (see Figure 3) or in the series of *rainfall departures in Sahel* (see Figure 5).

From the statistical point of view, the search for a change of type B means that the change in linear regression is to be detected where time plays the role of the independent variable.

In our study we used methods where a non-stationarity of a certain type was supposed. Some statisticians prefer techniques which were not suggested to discover a particular type of departure from stationarity. These methods are based on cumulative sums (or moving sums) of recursive residuals and were first introduced by Brown *et al.* (1975).

The problem of more than one change can often be encountered in applications. Suppose one change was already detected. Then there exists the possibility of studying separately two parts of the original series, i.e. the part before the change and the part after the change, and to try to decide whether these two parts can be considered stationary or whether another change can be discovered. In the case where another change-point is detected, one can proceed in the same way until all change-points are found. For the detection of a sudden change in the mean of type A, Vostrikova (1981) showed that if the intervals between the changes are large then the procedure of sequential division discovers all the inhomogeneities in the series. This follows from the fact that if a step-function with several jumps is approximated in the L^2 sense by a two value function with only one jump, then the 'jump-point' of the best approximation coincides with one of the 'jump-points' of the approximated function. On the other hand, if a function with several changes of type B is approximated by a function with only one change of the same type, the change-point may not coincide with any change of the approximated function. Therefore, the sequential approach yields completely incorrect results for the detection of continuous alternations of the process.

4. STATISTICAL METHODS

The most frequently applied procedures for the detection of a sudden change in a series of observations are based on statistical hypothesis testing. In our study we used test statistics that are of so called *maximum type*. The idea behind this is described below. Suppose for a while that we do not know whether a series changed or not, but we know for sure that if the series changed then it occurred at a certain point k. In the framework of mathematical statistics, the problem can be solved by testing the null hypothesis that claims that there is no change in the distribution of the series against the alternative hypothesis that claims that the distribution of the series changed at time k. Suppose that the test statistic for this problem is known. In the case of testing whether a change of parameter of a known distribution occurs, the usually applied test statistic is equivalent to the log-likelihood ratio. Now, consider the situation where we do not know whether a change occurred nor do we have any idea where the change point could be. We then test the null hypothesis of no change against the alternative that there exists a time when the distribution of the series changed. In this case it is natural to use for testing the maximum of preceding test statistics, where the maximum is taken over all possible time points where the change might occur.

For illustration, consider the following simple example of the detection of a change of type A in the mean of independent normally distributed variables $X_1, \ldots, X_n$. The null hypothesis claims that $X_1, \ldots, X_n$ are distributed according to the same $\mathrm{N}(\mu, \sigma^2)$. The alternative claims that there exists a time point $k \in \{1, \ldots, n-1\}$ such that $X_1, \ldots, X_k$ are distributed according to $\mathrm{N}(\mu_1, \sigma^2)$ and $X_{k+1}, \ldots, X_n$ are distributed according to $\mathrm{N}(\mu_2, \sigma^2)$ with $\mu_1 \neq \mu_2$. Supposing σ^2 is unknown, then the test statistic $T(n)$ is the maximum of the absolute values of two-sample t-test statistics

$$T(n) = \max_{1 \leqslant k<n} |T_k| = \max_{1 \leqslant k<n} \sqrt{\left(\frac{(n-k)k}{n}\right)} |\bar{X}_k - \bar{X}_k^*| \frac{1}{s_k}$$

$$\bar{X}_k = \frac{\sum_{j=1}^{k} X_j}{k}, \quad \bar{X}_k^* = \frac{\sum_{j=k+1}^{n} X_j}{n-k}, \quad s_k = \sqrt{\{(\Sigma(X_i - \bar{X}_k)^2 + \Sigma(X_i - \bar{X}_k^*)^2)/(n-2)\}}.$$

The null hypothesis is rejected if the statistic $T(n)$ is larger than a corresponding critical value. The calculation of exact critical values is complicated because the test statistics (that we take a maximum of) are dependent and their covariance function $R(k, l)$ depends on both k and l and not only on the difference $k - l$. Moreover, if the number of observations increases, the maximum type test statistic tends to infinity w.p.1. The intuitive explanation for erratic behaviour of the sequence $\{|T_k|\}$ near the edges (k near 0 or near n) is the following. When k is near the edge (say near 0) we compare an estimate calculated from the large number of observations – the last $n-k$ ones – with an estimate provided by a very small number of observations – the k first ones. Some authors (see James *et al.* 1987) cope with this problem so that they use for testing the maximal statistic where the maximum is taken only over a certain portion of time points $\{k; [t_0 n] < k < [(1-t_0)n]\}$ for some $0 < t_0 < 0{\cdot}5$. Some other statisticians (see Deshayes and Picard 1986) suggest weighting the statistic corresponding to the point k by the weight $w(k/n)$, i.e. instead of the statistic $T(n) = \max_{1 \leqslant k<n} |T_k|$ they apply the statistic

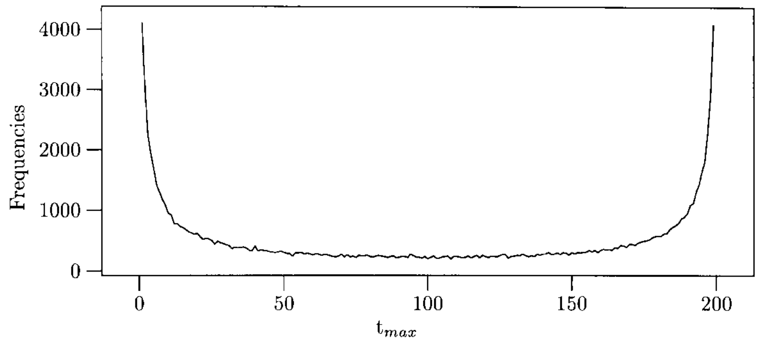

Figure 10. Polygon of frequencies of time t_{max} where the statistic $T(n)$ takes its maximum, $n = 200$, H_0 holds

$\max_{1 \leqslant k<n} w(k/n) \mid T_k \mid$. The weights suppress influence of the statistics $\mid T_k \mid$ near the beginning and near the end of the series (see Gombay and Horváth 1988). The most frequently used weights are $\{w(k/n) = \sqrt{\{k/n(1 - k/n)\}}, k = 1, \ldots, n-1\}$. We conclude that the three types of test statistics are usually applied to solve the same problem:

(i) global maximum of statistics over all time points;
(ii) maximum of statistics over a trimmed portion of time points;
(iii) maximum of weighted statistics.

The supporters of the first approach believe that it is more natural to consider the global maximum than the trimmed maximum because it is our decision how to choose t_0 and the rejection-acceptance decision may depend on this choice (see Gombay and Horváth 1994b). The supporters of trimmed maximal statistics emphasize that there is a large probability that under the null hypothesis the maximal value occurs for time point k near 0 or n. This effect is illustrated in Figure 10 which shows the polygon of frequencies of the time k_0, where the statistic $T(n)$ takes its maximum, provided $n = 200$ and H_0 holds. The polygon was based on 100,000 simulations of sequences of 200 independent variables with N(0,1) distribution. It is clear that if we leave out the values $\mid T_k \mid$ near both edges, the critical values of $T(n, t_0) = \max_{[t_0 n]<k<[(1-t_0)n]} \mid T_k \mid$ are smaller. If the researchers are sure that the change cannot occur at the very beginning or at the very end of the series, then detection of a change by applying $T(n, t_0)$ is easier. Table II shows the difference between critical values of the global maximal statistic $T(n)$ and the statistic $T(n, 0{\cdot}05)$ when 5 per cent of the values $\{\mid T_k \mid\}$ in the beginning and 5 per cent of $\mid T_k \mid$ in the end were trimmed. (Both critical values were estimated by simulation.)

Table II. Several examples of 5 per cent and 1 per cent critical values of the global maximal statistic $T(n)$ and of the trimmed maximal statistic $T(n, 0{\cdot}05)$

n	5% critical values of:		1% critical values of:	
	$T(n)$	$T(n, 0{\cdot}05)$	$T(n)$	$T(n, 0{\cdot}05)$
50	3·15	3·08	3·76	3·69
100	3·16	3·06	3·71	3·62
200	3·19	3·07	3·72	3·61
300	3·21	3·08	3·73	3·62
500	3·24	3·09	3·73	3·62

The maximum of weighted statistics detects more easily a change in the middle of the series than at the edges (see the intuitive explanation and numerical results given by James *et al.* (1987)). Notice that the maximum of weighted statistics $TP(n)$ in the case of a search for a change in the mean of normally distributed random variables is the Studentized CUSUM statistic as it holds

$$TP(n) = \max_{1 \leqslant k<n} \mid T_k \mid \sqrt{\left\{\frac{k}{n}\left(1-\frac{k}{n}\right)\right\}} = \max_{1 \leqslant k<n} \frac{1}{\sqrt{n}}\left|\sum_{i=1}^{k}(X_i - \bar{X})\right|\frac{1}{s_k}.$$

Personally, I prefer the procedures based on the maximum of non-weighted statistics, for their estimate of the change-point is better.

For all testing statistics described above, approximate critical values can be obtained by several different methods, namely:

(i) the Bonferroni inequality and its improvements;
(ii) the asymptotic distribution;
(iii) simulation.

The Bonferroni inequality for the sequence of random events $\{\mathcal{A}_k\}$ can be expressed in the form

$$P\left(\bigcup_{k=1}^{n} \mathcal{A}_k\right) \leqslant \sum_{k=1}^{n} P(\mathcal{A}_k).$$

Applying this inequality, for example, to the events $\mathcal{A}_k = \{\mid T_k \mid > c\}$, $k = 1, \ldots, n-1$, we obtain

$$P(T(n) > c) = P\left(\max_{1 \leqslant k \leqslant n-1} \mid T_k \mid > c\right) \leqslant \sum_{k=1}^{n-1} P(\mid T_k \mid > c).$$

The value $t_{1-\alpha/2(n-1)}(n-2)$ can serve as a conservative critical value because

$$P(T(n) > t_{1-\alpha/2(n-1)}(n-2)) \leqslant \alpha.$$

For small significance levels and moderate lengths of the series, the Bonferroni inequality gives rather good estimates of critical values.

The asymptotic distribution of the global maximal statistic for a change in the mean as well as the variance of normally distributed variables was studied by Yao and Davis (1986), Horváth (1993) and Gombay and Horváth (1994a). General results for the global maximum of the log-likelihood statistics were obtained by Gombay and Horváth (1996). Consider again the example of the statistic $T(n)$. Under the null hypothesis the limit distribution $T(n)$ normalized by the appropriate constants a_n and b_n is of the extreme type, i.e.

$$\lim_{n\to\infty} P\left(\frac{T(n) - b_n}{a_n} > x\right) = 1 - \exp\left(\frac{-2\mathrm{e}^{-x}}{\sqrt{\pi}}\right) \tag{1}$$

where $a_n = (2 \log \log n)^{-1/2}$ and $b_n = a_n^{-1} + (a_n/2)\log \log \log n$. Unfortunately, the convergence to the limit distribution is slow. From our simulation study it follows that for $100 \leqslant n \leqslant 1000$ the 50, 51, ..., 99 per cent quantiles obtained from the asymptotic distribution are greater than the corresponding critical values obtained by simulations. This means that the test based on the asymptotic critical values is more conservative than that based on simulations. Moreover, the maximal difference between the distribution function of $T(n)$ estimated by simulations and the approximate distribution function obtained by (1) is about 0·08 and the 95 per cent quantile of limit distribution (1) corresponds to the 98–99 per cent quantile obtained by simulations.

I would like to mention that the limit distribution (1) was obtained for any independent zero mean random variables such that $E\,|\,X_i\,|^{2+\delta} < \infty, (\delta > 0)$. This enables us to use (1) even for non-normally distributed (e.g. skewed) data (see Csörgö and Horváth 1988).

Antoch *et al.* (1996) show that if the variables $X_1, \ldots, X_n$ are not independent but form an ARMA sequence then the asymptotic critical values have to be multiplied by $\surd\{2\pi f(0)/\gamma\}$ where $\gamma = \text{var}\ X_t$ and $f(\cdot)$ denotes the special density of the corresponding ARMA process. For AR sequences of the first order the difference between asymptotic critical values and critical values obtained by simulations is even greater than for the independent variables (see Jarušková 1996).

The critical values for detection of a change in the mean of independent normally distributed variables as well as in the mean of an autoregressive sequence of the first order obtained by simulation were tabulated by Jarušková (1996). The critical values for a change in the variance of independent normally distributed variables were tabulated by Jarušková and Antoch (1993).

The distribution of a trimmed maximal statistic can be approximated by the distribution of the maximum of a limit process on the interval $[t_0, 1 - t_0]$. Gombay and Horváth (1996) give conditions under which the limit process for statistics derived from the log-likelihood principle for detection of a change in a d-dimensional parameter θ is $\|\,B(t)\,\|\,/\surd\{t(1-t)\}$, where $\|\,B(t)\,\|$ denotes the Euclidean norm of the d-dimensional Brownian bridge. Siegmund (1985) derived the approximation of the exceedence probability that gives good estimates of critical values for trimmed maximal log-likelihood statistics

$$P\left(\max_{t_0 \leqslant t \leqslant 1-t_0} \frac{\|\,B(t)\,\|}{\surd\{t(1-t)\}} > x\right) \simeq \frac{x^d\, \mathrm{e}^{-x^2/2}}{2^{(d-2)/2}\Gamma(d/2)}\left(\left(1 - \frac{d}{x^2}\right)\log\frac{1-t_0}{t_0} + \frac{2}{x^2}\right). \tag{2}$$

The approximation (2) was derived for x large but it gives surprisingly good estimates even for moderate values of x. Using (2), the 5 per cent asymptotic critical value of $T(n, 0{\cdot}05)$ is 3·15 and 1 per cent critical value is 3·67.

Now, we shall consider the change of type B. In the case of a change in the mean of type B of normally distributed random variables we test the null hypothesis which claims that $X_1, \ldots, X_n$ are distributed according to $\mathrm{N}(a, \sigma^2)$ against the alternative hypothesis that there exists a time point $k \in \{1, \ldots, n-1\}$ such that for $i = 1, \ldots, k$, the variable X_i is distributed according to $\mathrm{N}(a, \sigma^2)$, and for $i = k+1, \ldots, n$ the variable X_i is distributed according to $\mathrm{N}(a + b(i-k)/n, \sigma^2)$, where $b \neq 0$. For the known variance σ^2 (we suppose for simplicity $\sigma^2 = 1$) the test statistic of the 'maximum type' $U(n) = \max_{1 \leqslant k<n} |\,U_k\,|$ can be applied, where the statistics

$$U_k = \frac{\sum_{i=k+1}^{n} (X_i - \bar{X})(i-k)}{\sqrt{\left\{\frac{(n-k)(n-k+1)(2n-2k+1)}{6} - \frac{(n-k)^2(n-k+1)^2}{4n}\right\}}}, \quad k = 1, \ldots, n-1$$

are distributed according to N(0,1), but are not independent. For unknown σ^2 the test statistic has the form

$$\tilde{T}(n) = \max_{1 \leqslant k<n} |\tilde{T}_k| = \max_{1 \leqslant k<n} \frac{|U_k|}{\sqrt{(\mathrm{RSS} - U_k^2)}} \sqrt{(n-2)}$$

where RSS is the residual sum of squares under the null hypothesis. Under the null hypothesis the statistics $\tilde{T}_k$ have t-distribution with $n-2$ degrees of freedom. In addition to the global maximum of statistics over all time points we can again introduce the maximum of statistics over a trimmed portion of time points $\tilde{T}(n, t_0) = \max_{1 \leqslant k \leqslant [(1-t_0)n]} |\tilde{T}_k|$. Critical values of $\tilde{T}(n)$ and $\tilde{T}(n, t_0)$ can be obtained by the Bonferroni or Worsley inequality if the number of observations is small.

The asymptotic behaviour of statistics $\tilde{T}(n)$ and $\tilde{T}(n, t_0)$ is given by the limit process

$$X(t) = \frac{\displaystyle\int_t^1 (s-t)\,\mathrm{d}W(s) - W(1)\frac{(1-t)^2}{2}}{\sqrt{\left\{\dfrac{(1-t)^3}{3} - \dfrac{(1-t)^4}{4}\right\}}}.$$

Critical values of $\tilde{T}(n)$ can be calculated from the asymptotic distribution

$$\lim_{n\to\infty} P\left(\frac{\tilde{T}(n) - \tilde{b}_n}{\tilde{a}_n} > x\right) = 1 - \exp\left(\frac{-\sqrt{3}\mathrm{e}^{-x}}{2\pi}\right) \tag{3}$$

where $\tilde{a}_n = (2 \log \log n)^{-1/2}$ and $\tilde{b}_n = (2 \log \log n)^{1/2}$.

Conservative critical values of $\tilde{T}(n, t_0)$ can be obtained by the inequality

$$\begin{aligned} \lim_{n\to\infty} P(\tilde{T}(n, t_0) > x) &= P\left(\max_{0 \leqslant t \leqslant 1-t_0} |X(t)| > x\right) \\ &\leqslant 2(1-\Phi(x)) + \sqrt{3}\left(\log\frac{(1+\sqrt{(1-t_0)})}{(1-\sqrt{(1-t_0)})} - \frac{2}{\sqrt{3}}\arctan\sqrt{3(1-t_0)}\right)\frac{\mathrm{e}^{-x^2/2}}{2\pi}. \end{aligned} \tag{4}$$

Using (4) the 5 per cent conservative critical value of $\tilde{T}(n, 0{\cdot}05)$ is 2·51 and the 1 per cent critical value is 3·07.

For several values of n the 5 per cent and 1 per cent critical values of $\tilde{T}(n)$ and $\tilde{T}(n, 0{\cdot}05)$ obtained by simulations are listed in Table III.

Similarly, as in the case of (1), the convergence in (3) is slow. The 5 per cent and 1 per cent critical values calculated from (3) are smaller than those obtained by simulations. On the other hand, the critical values obtained by (4) are close to the critical values obtained by simulations.

At the end of this section I would like to mention that besides the statistics of the 'maximum type' one can also apply (for the same kind of problems) the statistics derived from the Bayesian principle. The Bayesian approach was introduced by Chernoff and Zacks (1964), Kander and Zacks (1966) and Gardner (1969) and further developed by MacNeill (1974), and Jandhyala and MacNeill (1991).

Table III. Several examples of 5 per cent and 1 per cent critical values of the global maximal statistic $\tilde{T}(n)$ and of the trimmed maximal statistic $\tilde{T}(n, 0{\cdot}05)$

n	5% critical values of:		1% critical values of:	
	$\tilde{T}(n)$	$\tilde{T}(n, 0{\cdot}05)$	$\tilde{T}(n)$	$\tilde{T}(n, 0{\cdot}05)$
50	2·62	2·54	3·27	3·20
100	2·63	2·50	3·21	3·09
200	2·65	2·49	3·22	3·09
300	2·65	2·48	3·22	3·07
500	2·68	2·47	3·22	3·05

The distribution theory for the statistics derived from this principle is in many cases easier than for the 'maximum type' statistics. For the positive (negative) change in the mean of type A the testing statistic has the form

$$TB = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X})(i - 1)$$

(see Chernoff and Zacks 1964) and for the positive (negative) change in the mean of type B the statistic is

$$TB = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X}) i (i - 1)$$

(see Farley and Hinich 1970). Supposing σ^2 is known, both of them have normal distribution with a zero mean and an easily calculable variance. On the other hand, the 'Bayesian' statistics do not provide a user with the estimate of the change-point.

REFERENCES

Alexandersson, H. (1986). 'A homogeneity test applied to precipitation data', *Journal of Climatology*, **6**, 661–675.

Antoch, J., Hušková, M. and Prášková, Z. (1996). 'Effect of dependence on statistics for determination of change', *Journal of Statistical Planning and Inference* (to appear).

Bernier, J. (1994). 'Statistical detection of changes in geophysical series', in Duckstein, L. and Parent, E. (eds), *Engineering Risk and Reliability in a Changing Physical Environment*, NATO ASI, series E: Applied Sciences, **275**, Kluwer, Academic Publisher, the Netherlands, pp. 153–165.

Brown, R. L., Durbin, J. and Evans, J. M. (1975). 'Techniques for testing the constancy of regression relationship over time (with discussion)', *Journal of the Royal Stastical Society, Series B*, **37**, 149–192.

Buishand, T. A. (1982). 'Some methods for testing the homogeneity of rainfall records', *Journal of Hydrology*, **58**, 11–27.

Buishand, T. A. (1984). 'Tests for detecting a shift in the mean of hydrological records', *Journal of Hydrology*, **73**, 51–69.

Cobb, G. W. (1978). 'The problem of the Nile: Condition solution to a changepoint problem', *Biometrika*, **65**, 243–251.

Chernoff, H. and Zacks, S. (1964). 'Estimating the current mean of a normal distribution which is subject to changes in time', *Annals of Mathematical Statistics*, **35**, 999–1018.

Cšorgö, M. and Horváth, L. (1988). 'Nonparametric methods for changepoint problems', *in* Krishnaiah, P. R. and Rao, C. R. (eds), *Handbook of Statistics* **7**, 403–425.

Deshayes, J. and Picard, D. (1986). 'Off-line statistical analysis of change-point models, Detection of abrupt changes in signals and dynamic systems', *Lecture Notes in Control and Information Sciences*, **77**, 103–168.
Farley, J. U. and Hinich, M. J. (1970). 'A test for shifting slope coefficient in a linear model', *Journal of the American Statistical Association*, **65**, 1320–1329.
Gardner, L. A. (1969). 'On detecting changes in the mean of normal variates', *Annals of Mathematical Statistics*, **40**, 116–126.
Gombay, E. and Horváth, L. (1994a). 'Asymptotic distributions of maximum likelihood tests for change in the mean', *Biometrika*, **77**, 411–414.
Gombay, E. and Horváth, L. (1994b). 'Limit theorems for change in linear regression', *Journal of Multivariate Analysis'*, **48**, 411–414.
Gombay, E. and Horváth, L. (1996). 'On the rate of approximations for maximum likelihood tests in change-point models', *Journal of Multivariate Analysis*, **56**, 120–152.
Hanssen-Bauer, I. and Førland, E. J. (1994). 'Homogenizing long Norwegian precipitation series,' *Journal of Climate*, **7**, 1001–1013.
Horváth, L. (1993). 'The maximum likelihood method for testing changes in the parameters of normal observations', *Annals of Statistics*, **21**, 671–680.
Hubert, P., Carbonnel, P. J. and Chaouche, A. (1989). 'Segmentation des séries hydrométéorologiques – application à des séries de précipitations et de débits de l'Afrique de l'Ouest', *Journal of Hydrology*, **110**, 349–367.
James, B., James, K. L. and Siegmund, D. (1987). 'Tests for a change-point', *Biometrika*, **74**, 71–83.
Jandhyala, V. K. and MacNeill, I. B. (1991). 'Tests for parameter changes at unknown times in linear regression models', *Journal of Statistical Planning and Inference* **27**, 291–319.
Jarušková, D. (1990). 'Detection of change-point in series of river discharges' (in Czech), *Vodohospospodářský časopis*, **38**, 501–515.
Jarušková, D. (1993). 'Statistical method for change-point detection using rainfalls–runoffs relationship' (in Czech), *Vodohospospodářský časopis*, **41**, 85–99.
Jarušková, D. and Antoch, J. (1993). 'Detection of change in variance'. In Mandl, P. and Hušková, M. (eds), *Proceedings of the Fifth Prague Symposium on Asymptotic Statistics*, Physica Verlag, Springer Verlag, pp. 297–301.
Jarušková, D. (1996). 'Change-point detection in meteorological measurement', *Monthly Weather Review*, **124**, 1535–1543.
Jones, P. D., Wigley, M. L. and Briffa, K. R. (1994). 'Global and hemispheric temperature anomalies – land and marine instrumental records'. *In* Boden, T. A., Kaiser, D. P., Sepanski, R. J. and Stoss, F. W. (eds), *Trends '93: A Compendium of Data on Global Change*, ORNL/CDIAC – 65, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn., USA, pp. 603–608.
Kander, Z. and Zacks, S. (1966). 'Test procedures for possible changes in parameters of statistical distributions occurring at unknown time points', *Annals of Mathematical Statistics*, **37**, 1196–1210.
Kim, H. J. and Siegmund, D. (1989). 'The likelihood ratio test for a change-point in simple linear regression', *Biometrika*, **76**, 409–423.
Lawrance, A. J. and Kottegoda, N. T. (1977). 'Stochastic modelling of riverflow time series (with discussion)', *Journal of the Royal Statistical Society, Series A*, **140**, 1–46.
Lombard, F. (1994). 'On a Fourier method of detecting climatic transitions', *Beitraege zur Physik der Atmosphaere*, **67**, 201–208.
MacNeill, I. B. (1974). 'Tests for change of parameter at unknown times and distributions of some related functionals on Brownian motion', *Annals of Statistics*, **2**, 950–962.
Nicholson, S. E. (1994). 'Century-scale series of standardized annual departure of African rainfall'. *In* Boden, T. A., Kaiser, D. P., Sepanski, R. J. and Stoss, F. W. (eds), *Trends '93: A Compendium of Data on Global Change*, ORNL/CDIAC – 65, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn., USA, pp. 952–962.
Potter, K. W. (1981). 'Illustration of a new test for detecting a shift in mean in precipitation series', *Monthly Weather Review*, **109**, 2040–2045.
Rhoades, D. A. and Salinger, M. J. (1993). 'Adjustment of temperature and rainfall records for site changes', *Journal of Climatology*, **13**, 899–913.
Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*, Springer-Verlag, New York, Heidelberg, Berlin.

Vannitsem, S. and Nicolis, C. (1991). 'Detecting climatic transitions: statistical and dynamical aspects', *Beitraege zur Physik der Atmosphaere*, **64**, 245–254.
Vostrikova, L. Ju. (1981). 'Detecting "disorder" in multidimensional random processes', *Soviet Mathematics-Doklady*, **24**, 55–59.
Worsley, K. J. (1982). 'An improved Bonferroni inequality and application', *Biometrika*, **69**, 297–302.
Yao, Yi-Ching and Davis, R. A. (1986). 'The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates', *Sankhya*, **48**, 339–353.

