

Change-Point Detection in Meteorological Measurement

DANIELA JARUŠKOVÁ

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

(Manuscript received 31 October 1994, in final form 4 January 1996)

ABSTRACT

Statistical methods of change-point detection can be useful for discovering inhomogeneities in precipitation, air pressure, or temperature time series caused by a change in the measurement process such as a relocation of a gauge. The method is based on a second correlated series that can be relied on to be correct, following the approach suggested by Potter. A summary of the latest methods is given, and the necessary tests and their critical values are provided. An application to air pressure series measured at three Swiss meteorological stations is presented.

1. Introduction

In the last decade several papers concerning the detection of change in time series have appeared in the statistical literature. Unfortunately, the new ideas and results in these papers are often known only within the mathematical community and have not been used in applications. These methods can be very useful in meteorology to find systematic changes in the mean of measured quantities, such as precipitation, air pressure, or temperature. Such changes might be caused by an alteration in the measurement process such as the relocation of a gauge, a change in the gauge exposure, or a change in the time at which the measurement is taken. The detection of such a change in the series of interest $\{Y_i\}$ is made possible by the existence of some other series $\{X_i\}$ that can be relied on to be correct and that is correlated with the series $\{Y_i\}$. The series $\{X_i\}$ is called the reference series. The series $\{X_i\}$ and $\{Y_i\}$ can be compared by statistical methods and possible change in the series $\{Y_i\}$ may be detected. This approach was recommended by Potter (1981) for detecting the shift in the mean of precipitation series. The same idea of finding inhomogeneities in the series $\{Y_i\}$ using the reference series $\{X_i\}$ was later applied by Alexandersson (1986), Hanssen-Bauer and Førland (1994), and Rhoades and Salinger (1993). Other authors who applied statistical methods for change-point detection in meteorological series included Buishand (1982, 1984), Lombard (1994), and Vannitsem and Nicolis (1991).

Potter uses the results of Maronna and Yohai (1978), who studied a model with a linear regression relationship between the series of interest and the reference series, that is,

model I

$$Y_i = a + bX_i + e_i, \quad i = 1, \dots, n,$$

where $\{e_i, i = 1, \dots, n\}$ are random errors.

In our paper we study the model described above as well as another simpler model where we suppose that our series of interest oscillates with approximately the same variance as the reference series and the only difference is in the mean, that is,

model II

$$Y_i = a + X_i + e_i, \quad i = 1, \dots, n,$$

where $\{e_i, i = 1, \dots, n\}$ are again random errors.

For better understanding of the topic of our paper, it is important to realize that in practice we can meet two situations that are substantially different from the statistical point of view.

(A) The time of possible change is known.

In the first case we know from the history of our series that at a certain point of time an event occurred that might influence our measurement, for example, the time of measurement was shifted or automatic measurement started. The question is whether this event introduced a systematic change into our series.

(B) The time of possible change is unknown.

It can sometimes happen that we do not know the history of our series or the history is not complete. Thus, it is not known whether any change in the measurement process occurred or not. In this case the goal of our investigation is not only to detect a systematic

Corresponding author address: Dr. Daniela Jarušková, Department of Mathematics, Faculty of Civil Engineering, Czech Technical University, Thakurova 7, 166 29 Prague 6, Czech Republic.
E-mail: jarus@earn.cvut.cz

TABLE 1. Critical values for model II obtained by the Bonferroni inequality.

<i>n</i>	5% critical value	1% critical value
10	3.75	4.96
20	3.49	4.21
30	3.47	4.07
40	3.48	4.03
50	3.50	4.02
70	3.54	4.03
100	3.60	4.05

change in our measurement but also to find where this change occurred.

2. Mathematical formulation

To solve these problems the hypothesis testing can be applied. The log-likelihood ratio will be used to provide test statistics.

a. Model II

We start with the simple model II and introduce the new variables $Z_i = Y_i - X_i, i = 1, \dots, n$, supposing that these differences are normally distributed.

(A) Suppose first that *the time of change is known* and equal to k . The problem can be described using the following null and alternative hypotheses.

$$H_0: Z_i = a + e_i, \quad i = 1, \dots, n,$$

$$A_k: Z_i = a + e_i, \quad i = 1, \dots, k,$$

$$Z_i = a + d + e_i, \quad i = k + 1, \dots, n, \quad d \neq 0.$$

Here and in the following the variables $\{e_i\}$ are independent, identically normally distributed with zero mean and unknown variance σ_e^2 .

The problem described above can be treated by the two-sample t test with the test statistic T_k distributed under the null hypothesis according to the Student's t -distribution with $(n - 2)$ degrees of freedom:

$$T_k = \left[\frac{(n - k)k}{n} \right]^{1/2} (\bar{Z}_k - \bar{Z}_k^*) \frac{1}{S_k}$$

$$= \frac{\frac{1}{n^{1/2}} \left(S_k - \frac{k}{n} S_n \right)}{\left[\frac{k}{n} \left(1 - \frac{k}{n} \right) \right]^{1/2}} \frac{1}{S_k},$$

where

$$S_k = \sum_{i=1}^k Z_i, S_k^* = \sum_{i=k+1}^n Z_i, \bar{Z}_k = \frac{S_k}{k}, \bar{Z}_k^* = \frac{S_k^*}{n - k},$$

$$s_k^2 = \frac{1}{n - 2} \left[\sum_{i=1}^k (Z_i - \bar{Z}_k)^2 + \sum_{i=k+1}^n (Z_i - \bar{Z}_k^*)^2 \right].$$

Denote by $t_p(m)$ the 100 $p\%$ quantile of a t distribution with m degrees of freedom. The null hypothesis is rejected if

$$|T_k| > t_{1-\alpha/2}(n - 2).$$

For n large the t distribution can be approximated by the standard normal distribution and quantiles of the t distribution by quantiles of the standard normal distribution.

(B) The more complicated situation occurs if *the time of change is unknown*. Then we can set the null hypothesis and the alternative in the following way:

$$H_0: Z_i = a + e_i, \quad i = 1, \dots, n,$$

$$A: \exists k \in \{1, \dots, n - 1\} \text{ such that}$$

$$Z_i = a + e_i, \quad i = 1, \dots, k,$$

$$Z_i = a + d + e_i, \quad i = k + 1, \dots, n, \quad d \neq 0.$$

The null hypothesis can be rejected if at least one of the statistics $\{|T_k|\}$ is greater than the critical value, that is, the test statistic has the form $T(n) = \max_{k=1, \dots, n-1} |T_k|$. The exact distribution of $T(n)$ was derived by Worsley (1979). However, the distribution is so complex that Worsley was able to calculate the critical values only for the number of observations n less than 10.

Approximate critical values can be obtained by several different methods, namely

- the Bonferroni inequality,
- simulation,
- the asymptotic distribution.

1) BONFERRONI INEQUALITY

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

$$P\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n P(A_k).$$

Applying this inequality to the events $A_k = \{|T_k| > c\}, k = 1, \dots, n - 1$, we obtain

TABLE 2. Critical values for model II obtained by simulation.

<i>n</i>	5% critical value	1% critical value
10	3.66	4.90
20	3.28	4.06
30	3.19	3.87
40	3.16	3.80
50	3.15	3.76
70	3.15	3.73
100	3.16	3.71
200	3.19	3.71
400	3.23	3.73
1000	3.28	3.79

TABLE 3. Critical values for model II obtained from the asymptotic distribution.

n	5% critical value	1% critical value
10	3.62	4.88
20	3.60	4.70
30	3.61	4.65
40	3.61	4.62
50	3.62	4.60
70	3.63	4.59
100	3.64	4.57
200	3.66	4.55
400	3.68	4.54
1000	3.71	4.54

$$P[T(n) > c]$$

$$= P\left(\max_{k=1, \dots, n-1} |T_k| > c\right) \leq \sum_{k=1}^{n-1} P(|T_k| > c).$$

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

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

Table 1 presents the critical values obtained in this way for several values of n .

2) SIMULATION

To estimate the critical values we performed a simulation study using the program MATLAB. For $n = 10, 20, \dots, 100, 200, \dots, 1000$ we got 100 000 realizations of the statistic $T(n)$ and calculated the empirical distribution function and its quantiles. The procedure was repeated ten times and the critical values were estimated by the average of the corresponding empirical quantiles. The estimation of the both 5% as well as 1% critical values are given in Table 2.

3) ASYMPTOTIC DISTRIBUTION

Yao and Davis (1986) derived the limit distribution of the statistic $T(n)$ normalized by the appropriate constants a_n, b_n :

$$\lim_{n \rightarrow \infty} P\left[\frac{T(n) - b_n}{a_n} < x\right] = \exp(-2e^{-x\pi^{-1/2}}), \quad (1)$$

where $a_n = (2 \log \log n)^{-1/2}$ and $b_n = a_n^{-1} + (a_n/2) \log \log n$. Some chosen critical values for the significance level $\alpha = 0.05$ and $\alpha = 0.01$ are presented in Table 3.

It is well known that the convergence of (1) is slow. From our simulation study it follows that for $100 \leq n \leq 1000$ the 50%, 51%, \dots , 99% quantiles obtained from the asymptotic distribution are greater than the corresponding critical values obtained by simulation. This means that the test based on the asymptotic critical values is more conservative than that based on simu-

lation. Moreover, the maximal difference between the distribution function of $T(n)$ estimated by simulation and the approximate distribution function obtained by (1) is about 0.08 and the 95% quantile of limit distribution (1) corresponds to the 98%–99% quantile obtained by simulation.

If the time of change is known we can use the test statistic $B_k = T_k^2(n-2+T_k^2)^{-1}$ instead of T_k . Under the null hypothesis B_k is distributed according to the beta distribution with parameters $1/2$ and $(n-2)/2$. For the unknown time of change we can use the statistic $B(n) = \max_{k=1, \dots, n-1} B_k$ instead of the statistic $T(n)$. In the paper of Maronna and Yohai (1978), the statistic $nB(n)$ is denoted by T_1 .

b. Model I

(A) First we suppose the time of change is known and equal to k . The null hypothesis and the alternative can be set as follows:

$$H_0: Y_i = a + bX_i + e_i, \quad i = 1, \dots, n,$$

$$A: Y_i = a + bX_i + e_i, \quad i = 1, \dots, k,$$

$$Y_i = a + d + bX_i + e_i, \quad i = k + 1, \dots, n, \quad d \neq 0.$$

The test statistic \tilde{B}_k has the following form (see Worsley 1983 and Maronna and Yohai 1978):

$$\begin{aligned} \tilde{B}_k &= \frac{C_k^2}{\left[1 - \frac{(\bar{x}_k - \bar{x}_n)^2}{\sum (x_i - \bar{x}_n)^2} \frac{nk}{n-k}\right] Q} \frac{k(n-k)}{n} \\ &= \frac{\frac{nk}{n-k} [\bar{Y}_k - \bar{Y}_n - \hat{b}(\bar{x}_k - \bar{x}_n)]^2}{\left[1 - \frac{(\bar{x}_k - \bar{x}_n)^2}{\sum (x_i - \bar{x}_n)^2} \frac{nk}{n-k}\right] Q}, \end{aligned}$$

where x_1, \dots, x_n are realizations of the random variables X_1, \dots, X_n ; $r_i = Y_i - \hat{a} - \hat{b}x_i, i = 1, \dots, n$ are the residuals under the null hypothesis; $C_k = \sum_{i=1}^k r_i, Q = \sum_{i=1}^n r_i^2; \bar{x}_k = \sum_{i=1}^k x_i/k; \bar{Y}_k = \sum_{i=1}^k Y_i/k$. Under the null hypothesis \tilde{B}_k is distributed according to the beta distribution with parameters $1/2, (n-3)/2$. We reject the null hypothesis if

TABLE 4. Critical values for model I obtained by the Bonferroni inequality.

n	5% critical value	1% critical value
10	3.95	5.31
20	3.52	4.26
30	3.48	4.09
40	3.49	4.04
50	3.50	4.03
70	3.54	4.03
100	3.60	4.06

TABLE 5. Critical values for model I obtained by simulation.

n	5% critical value	1% critical value
10	3.86	5.13
20	3.30	4.04
30	3.26	3.87
40	3.21	3.88

$$\tilde{B}_k > \beta_{1-\alpha} \left[\frac{1}{2}, \frac{(n-3)}{2} \right],$$

where $\beta_{1-\alpha}(p, q)$ is the $100(1 - \alpha)\%$ quantile of the beta distribution with parameters p, q . There is a more frequently used statistic \tilde{T}_k that is equivalent to the statistic \tilde{B}_k . They are related as follows:

$$\tilde{B}_k = \frac{\tilde{T}_k^2}{n - 3 + \tilde{T}_k^2}.$$

Under the null hypothesis the statistic \tilde{T}_k is distributed according to the t distribution with $(n - 3)$ degrees of freedom. Thus, we reject the null hypothesis at the level α if

$$|\tilde{T}_k| > t_{1-\alpha/2}(n - 3).$$

(B) If the time of change is unknown, the null hypothesis and the alternative can be set as follows:

$$H_0: Y_i = a + bX_i + e_i, \quad i = 1, \dots, n,$$

$$A: \exists k \in \{1, \dots, n - 1\} \quad \text{such that}$$

$$Y_i = a + bX_i + e_i, \quad i = 1, \dots, k,$$

$$Y_i = a + d + bX_i + e_i, \quad i = k + 1, \dots, n, \quad d \neq 0.$$

For testing the null hypothesis we can use the statistic $\tilde{B}(n) = \max_{k=1, \dots, n-1} \tilde{B}_k$ or equivalently $\tilde{T}(n) = \max_{k=1, \dots, n-1} |\tilde{T}_k|$. Note that in the paper of Maronna and Yohai (1978) the statistic $n\tilde{B}(n)$ is denoted by T_0 . Approximate critical values can be again obtained by

- the Bonferroni inequality,
- simulation,
- the asymptotic distribution.

1) BONFERRONI INEQUALITY

Using the Bonferroni inequality in the same way as in the model II we obtain conservative critical values being $\beta_{1-\alpha/(n-1)}[1/2, (n - 3)/2]$ for the statistic $\tilde{B}(n)$ and $t_{1-\alpha/2(n-1)}(n - 3)$ for the statistic $\tilde{T}(n)$. The conservative critical values of $\tilde{T}(n)$ are given in Table 4.

If we are interested in computation of p values, the procedure suggested by Worsley (1983) gives upper bounds that are less conservative than those obtained by the application of the Bonferroni inequality. Nevertheless, a limitation of this approach is that the approximation might be poor for sample sizes $n > 50$, which are often encountered in climatology.

2) SIMULATION

The distribution of $\tilde{B}(n)$ and $\tilde{T}(n)$ depends on the values of x_1, \dots, x_n . This means that if we wish to simulate $\tilde{B}(n)$ or $\tilde{T}(n)$, we have to use our measured values x_1, \dots, x_n . Maronna and Yohai (1978) supposed that x_1, \dots, x_n are realizations of independent identically distributed random variables X_1, \dots, X_n with the normal distribution and obtained, with the help of a Monte Carlo study, the critical values of $T_0 = n\tilde{B}(n)$. Their critical values of $\tilde{T}(n)$ are given in Table 5.

3) ASYMPTOTIC DISTRIBUTION

Maronna and Yohai (1978) showed that if x_1, \dots, x_n are realizations of the i.i.d. random variables X_1, \dots, X_n with $0 < EX_i^2 < \infty$ and if, moreover, $\{(Y_1, X_1), (Y_2, X_2), \dots\}$ form an independent sequence and for every fixed i the variables X_i and e_i are independent, then the limit distribution of $\tilde{T}(n)$ and $T(n)$ is the same. Hušková (1994, personal communication) showed that the limit distribution of $\tilde{T}(n)$ and $T(n)$ is the same under less restrictive conditions assuming the vectors $\{X_1, \dots, X_n\}$ and $\{e_1, \dots, e_n\}$ are independent and $\{X_i\}$ is a stationary ARMA sequence. This result enables us to use the critical values from Table 3 also for the model I if the number of observations n is large.

3. Discussion about the application of the suggested methods to the meteorological series

The suggested methods can be applied to annual as well as to monthly averages. However, especially in the application to monthly averages, we may meet some problems because the assumptions of the suggested methods are not fulfilled. In the following discussion we deal with the problem of how seriously the violation of the assumptions can affect the conclusions of our inference.

TABLE 6. The 5% critical values of the statistic $T(n)$ for an AR(1) sequence.

n	ρ						
	0.1	0.2	0.3	0.4	0.5	0.6	0.7
10	3.90	4.21	4.56	4.94	5.39	5.86	6.38
20	3.52	3.82	4.18	4.64	5.19	5.89	6.82
30	3.45	3.74	4.12	4.56	5.14	5.89	6.96
40	3.43	3.73	4.10	4.55	5.12	5.89	6.99
50	3.42	3.72	4.09	4.54	5.11	5.89	6.96
70	3.43	3.74	4.10	4.56	5.12	5.89	6.96
100	3.44	3.75	4.12	4.58	5.14	5.90	6.99
200	3.48	3.80	4.19	4.65	5.23	5.99	7.06
400	3.52	3.85	4.25	4.72	5.31	6.09	7.16
1000	3.59	3.94	4.36	4.84	5.45	6.25	7.34

TABLE 7. The 1% critical values of the statistic $T(n)$ for an AR(1) sequence.

n	ρ						
	0.1	0.2	0.3	0.4	0.5	0.6	0.7
10	5.22	5.68	6.12	6.65	7.20	7.80	8.29
20	4.38	4.81	5.27	5.87	6.61	7.49	8.61
30	4.23	4.58	5.06	5.64	6.44	7.36	8.74
40	4.14	4.52	5.00	5.58	6.32	7.34	8.74
50	4.09	4.46	4.97	5.51	6.25	7.25	8.66
70	4.07	4.43	4.90	5.47	6.19	7.17	8.56
100	4.05	4.42	4.91	5.46	6.20	7.14	8.56
200	4.06	4.44	4.92	5.47	6.17	7.10	8.44
400	4.08	4.48	4.95	5.52	6.21	7.12	8.44
1000	4.14	4.56	5.04	5.62	6.34	7.25	8.60

a. Problem of the distribution

In the methods described above we assumed a normal distribution of the random errors. However, the limit distribution of the statistic

$$T(n) = \max_{k=1, \dots, n-1} \frac{\frac{1}{n^{1/2}} \left| S_k - \frac{k}{n} S_n \right|}{\left[\frac{k}{n} \left(1 - \frac{k}{n} \right) \right]^{1/2} s_k}$$

does not depend on the distribution of the variables $\{Z_i\}$ under the assumption that they are iid, satisfying $E|Z_i|^{2+\delta} < \infty$ for some $\delta > 0$ (see Csörgő and Horváth 1988). So, for every reasonable distribution, the same table of critical values can be used provided the number of observations n is moderately large. For series where the change is more likely to occur in the quotient Y_i/X_i than in the difference $Y_i - X_i$, it is possible to use the logarithm of the original series. Supposing that after this transformation the random errors are distributed according to the normal law, we can again use the suggested procedure. This approach is recommended by "Deutsche Wetterdienst" for the monthly averages of precipitation and was applied by Buishand (1982) and Rhoades and Salinger (1993).

b. Dependence

It is well known that there exists a certain persistence in the behavior of nature. This persistence is manifested by a dependence between the meteorological observations that are close in time. Such time series are very often modeled by ARMA series. The typical behavior of an autoregressive sequence of the first-order AR(1), with a large positive autoregressive coefficient, is a conspicuous slow oscillation around the mean. If we suppose that our variables are independent, this oscillation might be misinterpreted as a change in the mean. From this fact it is clear that for such series the critical value for rejecting the null hypothesis of no change

TABLE 8. Comparison of p values obtained by the asymptotic distribution (2) and by simulation for $\rho = 0.3$ and $100 \leq n \leq 1000$.

p value from Eq. (2)	0.4	0.3	0.2	0.1	0.05
p value by simulation	0.25	0.17	0.09	0.03	0.01

must be larger than for the independent variables. It was shown by Hušková (1994, personal communication) that if $\{Z_t\}$ is a stationary ARMA (p, q) sequence satisfying

$$Z_t - \Phi_1 Z_{t-1} - \dots - \Phi_p Z_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q},$$

where $\{\epsilon_t\}$ are i.i.d. such that $E\epsilon_1 = 0, E\epsilon_1^2 > 0, E\epsilon_1^4 < \infty$ and the polynomial $\Phi(z) = 1 - \Phi_1 z - \dots - \Phi_p z^p$ satisfies the condition that $\Phi(z) \neq 0$ for all complex z such that $|z| \leq 1$, then the limit behavior of the statistic $T(n)$ is

$$\lim_{n \rightarrow \infty} P \left\{ \frac{[2\pi f(0)/\gamma]^{-1/2} T(n) - b_n}{a_n} < x \right\} = \exp(-2e^{-x\pi^{-1/2}}), \quad (2)$$

where $\gamma = \text{var}Z_t$ and $f(\cdot)$ denotes the spectral density of the corresponding ARMA process. Thus, the critical values obtained from Table 3 have to be multiplied by $[2\pi f(0)/\gamma]^{1/2}$. Especially, for an AR(1) sequence that is one of the most frequently used in meteorology, the critical values should be multiplied by $[(1 + \rho)(1 - \rho)^{-1}]^{1/2}$, where ρ is the first autoregressive coefficient. We remark that the same standardization was recommended by Lombard (1994) for detecting change points by Fourier analysis if the observations are correlated.

To get a feeling for how much the distribution of $T(n)$ differs from the asymptotic distribution if n is small and the variables $\{Z_t\}$ are dependent, we simulated 100 000 realizations of AR(1) sequence satisfying the following equation

$$Z_i = \rho Z_{i-1} + \epsilon_i, \quad i = 1, \dots, n,$$

where $Z_0, \epsilon_1, \dots, \epsilon_n$ were independent and Z_0 was distributed according to the distribution $N[0, (1 - \rho^2)^{-1}]$ and $\epsilon_1, \dots, \epsilon_n$ according to the distribution $N(0, 1)$. The 5% critical values for selected values of ρ are given in Table 6; the 1% critical values in Table 7.

Tables 8 and 9 compare for $\rho = 0.3$ and $\rho = 0.5, p$ -values obtained by the asymptotic distribution (2) and

TABLE 9. Comparison of p values obtained by the asymptotic distribution (2) and by simulation for $\rho = 0.5$ and $100 \leq n \leq 1000$.

p value from Eq. (2)	0.4	0.3	0.2	0.1	0.05
p value by simulation	0.22	0.15	0.08	0.03	0.01

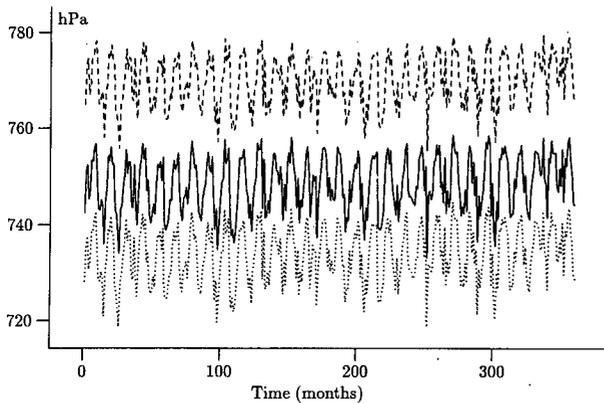


FIG. 1. Monthly averages of the air pressure (hPa) measured at Saentis (solid line), Guetsch (dashed line), and Weissfluhjoch (dotted line) from 1961 to 1990.

p values obtained by simulation. The difference between the asymptotic distribution function and the real distribution function estimated by simulation decreases very slowly. Due to the slow speed of convergence, the difference between p values obtained by (2) and p values obtained by simulation (supposing these are less than 0.4) is the same up to two decimal places for all n in the range 100–1000. We also notice that the test based on the critical values obtained by (2) is more conservative.

The coefficients of the ARMA series or the value of the spectral density function at zero should be known from the researcher's experience. If we are sure that in a certain interval the series is not subjected to any change, these coefficients can be obtained by estimation using this part of the series.

c. Removing seasonality

The monthly averages are almost never identically distributed random variables because of their seasonal character. This property might be inherited also by their differences. Therefore, we suggest to remove the seasonality by subtracting the corresponding means from the series.

4. Applications

Figure 1 presents three series of monthly averages of air pressure given in hectopascals measured at three Swiss meteorological stations—Saentis, Guetsch, and Weissfluhjoch—in the years 1961–90. The positions of the barometers in these stations are the following: Saentis—47°15'N, 9°21'E, height 2500.1 m; Guetsch—46°39'N, 8°37'E, height 2284.0 m; Weissfluhjoch—46°50'N, 9°49'E, height 2669.2 m. Every series consists of 360 observations. The series measured at Saentis was homogenized and serves as the reference series. The problem was to detect the inho-

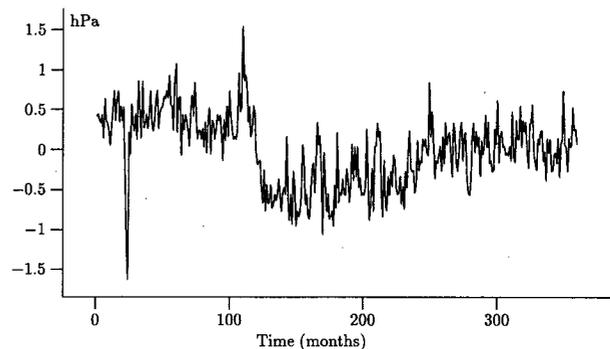


FIG. 2. Differences between the Guetsch and the Saentis seasonal adjusted series.

mogeneities in the Guetsch and Weissfluhjoch series caused by changes in the measurement process. Later on we obtained the history of these series so that we had the possibility to compare our results with reality and to see if the suggested procedures worked properly.

At the beginning of the statistical inference we removed the seasonality in our series subtracting from January's data the overall January average, from the February's data the overall February average, etc.

a. Guetsch–Saentis series

Supposing that the air pressure data measured at Saentis and Guetsch differ only in the mean we started by applying model II. The difference between the Guetsch and Saentis seasonal adjusted series is plotted in Fig. 2.

For every $k = 1, \dots, n - 1, n = 360$, the series of the differences was split into two parts so that the first part consisted of k observations and the second one of the remaining $(n - k)$ observations. For each split the statistic T_k was calculated. The values of statistics $\{T_k\}$ are shown in Fig. 3. The maximum of the statistics $\{T_k\}$

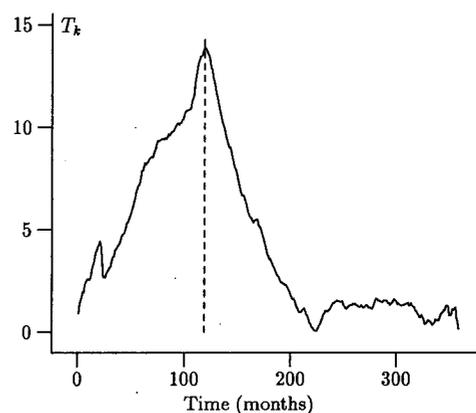


FIG. 3. Statistics $\{T_k\}$ corresponding to model II for the Guetsch–Saentis difference series.

TABLE 10. Results for the Guetsch–Saentis series.

	Time interval	<i>n</i>	Time of change	Test statistic
Step 1	1–360	360	119	13.92
Step 2	1–119	119	106	2.87
	120–360	241	240	12.82
Step 3	120–240	121	191	2.78
	241–360	120	356	1.83

occurred for $k = 119$ and was equal to 13.92. Comparing this value with the interpolated 1% critical value from Table 2, the null hypothesis is rejected at the 0.01 significance level. Applying model I to the same data gave similar results since the variances of both series are approximately the same and therefore the estimate of the regression coefficient b equals to 0.987, which is very close to the value 1.

Having detected the change after observation 119 the question may arise if there is some other change in addition to the one detected. The correct way to test the hypothesis that two changes in the series exist would be to find the maximum of test statistics corresponding to all possible division into three parts. This procedure would be time consuming and the properties of the test statistic derived from the log-likelihood principle have not been studied. Another possibility is to study separately two parts of our original series, that is, the part before the change and the part after the change. If the intervals between the changes are large, then Vostrikova (1981) showed that this procedure applied repeatedly discovers all the inhomogeneities in the series.

Proceeding in this way we did not detect any change in the first part of the series: the statistic $T(n)$ for $n = 119$ was equal to 2.87. Obviously, this value is smaller than the 5% critical value obtained from Table 2 by interpolation. In the second part we discovered a shift in location after observation $k = 240$, the statistic $T(n)$ for $n = 241$ being equal to 12.816. Continuing in the same way we did not find any other shift in location. The results are given in Table 10.

The means, medians, and standard deviations for all three parts of the series of the differences are given in Table 11.

Adjusting each part of the series by subtracting the corresponding mean and joining all three parts together, we got the series in Fig. 4.

TABLE 11. Descriptive statistics for the Guetsch–Saentis series.

Time interval	<i>n</i>	Mean (hPa)	Median (hPa)	Standard deviation (hPa)
1–119	119	0.385	0.365	0.358
120–240	121	-0.427	-0.468	0.317
241–360	120	0.049	0.047	0.255

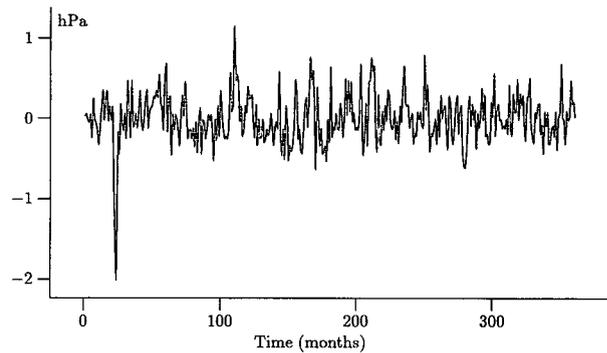


FIG. 4. Adjusted Guetsch–Saentis difference series.

The normal plot is given in Fig. 5. Both figures show a small number of outliers. Their appearance should be of interest for the meteorologists but cannot invalidate our results about the inhomogeneities in the series.

The more important problem is the problem of dependence between the observations. The autocorrelation function corresponding to the adjusted series presents a positive correlation between the neighboring observations (see Fig. 6).

Thus, our assumption about the independence was false. Assuming the differences between the seasonal adjusted series form an autoregressive sequence of first order, the estimate of the autoregressive coefficient ρ equals $\hat{\rho} = 0.33$. The comparison of the test statistics with corresponding interpolated critical values from Tables 6 and 7 gives the same results about the change points as the previous ones obtained under the assumption of independence.

Finally, we would like to show how our results coincide with the history of the Guetsch series. In the records three important events appear:

- 1) a change of gauge (16 December 1970) after observation $k = 120$;

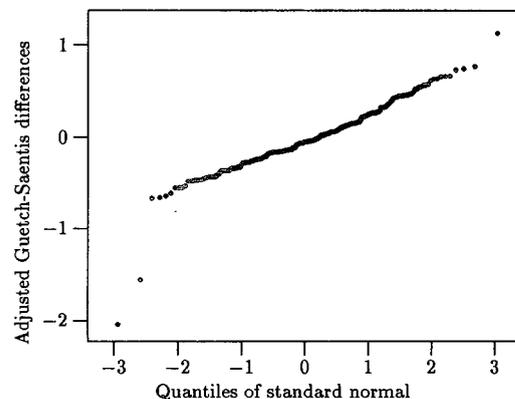


FIG. 5. Normal plot for the adjusted Guetsch–Saentis difference series.

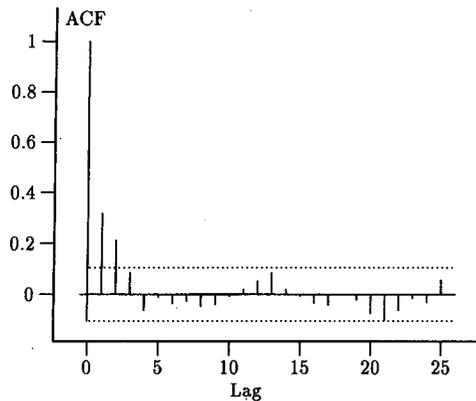


FIG. 6. Autocorrelation function of the adjusted Guetsch-Saentis difference series. Dotted lines denote approximate 95% confidence limits about zero.

2) another change of gauge (23 August 1979) after observation $k = 224$;

3) the beginning of automatic measurement (1 January 1981) after observation $k = 240$.

Thus, our procedure correctly discovered the change in the first and third case but did not detect any inhomogeneity after observation $k = 224$.

b. Weissfluhjoch-Saentis series

Using model II for the differences between Weissfluhjoch and Saentis (see Fig. 7), the test statistic $T(n)$ for $n = 360$ had the value 3.422 and the maximum occurred for $k = 189$. The behavior of $\{T_k\}$ is plotted in Fig. 8.

Supposing that the variables are independent, we compare the value 3.422 with the interpolated critical values from Table 2. We see that the null hypothesis can be rejected at the 0.05 significance level but it is not rejected at the 0.01 level. Supposing the variables form an AR(1) sequence and estimating ρ by $\hat{\rho} = 0.36$, the value $T(n) = 3.422$ is smaller than the interpolated 5% critical value from Table 6. This means that the change after the observation $k = 189$ was not confirmed in spite of the fact that the difference between the mean of the first and second part of the series was 0.129 hPa. We obtained this result because we supposed that we did not know anything about the history of the series so that we applied (B). However, later we were informed that in the history of the series there was a record about the change of the gauge on 24 August 1976, that is, after the observation $k = 188$, which might influence the series. Using (A) for the independent random variables, the null hypothesis is rejected at the significance level $\alpha = 0.05$ because

$$|T_k| > 1.96.$$

If we suppose that the variables form an AR(1) with

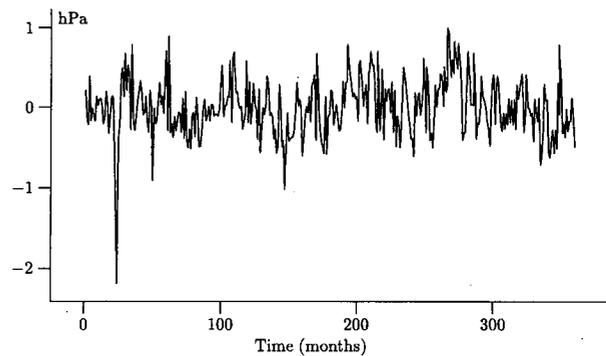


FIG. 7. Differences between the Weissfluhjoch and the Saentis seasonal adjusted series.

$\rho = 0.36$, then under the null hypothesis the statistic T_k is asymptotically distributed according to the normal distribution $N\{0, [(1 + \rho)(1 - \rho)^{-1}]^{1/2}\}$. The null hypothesis is rejected at the significance level $\alpha = 0.05$ because

$$|T_k| > 1.96 \left(\frac{1.36}{0.64} \right)^{1/2} = 2.85.$$

In the above examples the decision of whether there is a change point was based on the comparison of test statistics with the interpolated critical values from Tables 2, 6, and 7. Some readers might prefer to conduct tests using asymptotic distributions (1) and (2). Application of asymptotic distributions enables one to calculate p values and to gain in this way more information about the result of the test. However, the real p values are smaller than those obtained from (1) and (2). For some situations the difference can be roughly estimated, see the discussion about the speed of convergence of asymptotic distributions.

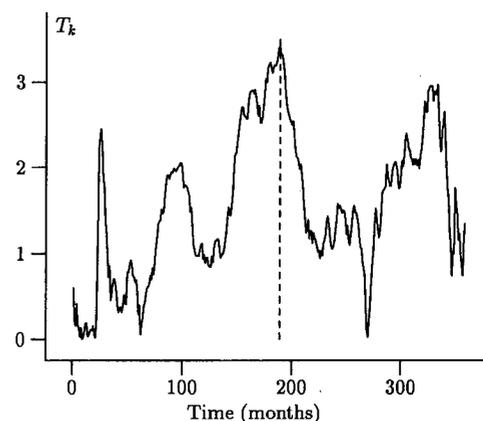


FIG. 8. Statistics $\{T_k\}$ corresponding to model II for the Weissfluhjoch-Saentis series.

5. Conclusions

For rejection of the null hypothesis in the case of unknown change, the difference between the mean of the first and second part of the series must be rather large, much larger than for the case where we had a record about particular event and we were interested as to whether this event inserted the inhomogeneity into the series. This procedure protects the users from altering the data without substantial reasons.

The method was suggested for only one sudden change in the series. If there are more changes, as in the Guetsch example, the procedure can be applied as well, but the user has to bear in mind that the power of the tests decreases.

Acknowledgments. I would like to thank A. Aschwanden and her colleagues of the Swiss Meteorological Institute for introducing me to this interesting problem, as well as for providing me with their meteorological data. I am greatly indebted to W. Stahel from ETH Zürich, Switzerland, and J. Antoch and M. Hušková from Charles University, Prague, Czech Republic, for their help in the preparation of the manuscript. The idea of this paper began during my study stay at ETH Zürich sponsored by the Swiss government.

REFERENCES

- Alexandersson, H., 1986: A homogeneity test applied to precipitation data. *J. Climatol.*, **6**, 661–675.
- Buishand, T. A., 1982: Some methods for testing the homogeneity of rainfall records. *J. Hydrol.*, **58**, 11–27.
- , 1984: Tests for detecting a shift in the mean of hydrological records. *J. Hydrol.*, **73**, 51–69.
- Csörgő, M., and L. Horváth, 1988: Nonparametric methods for change-point problems. *Handbook of Statistics*, P. R. Krishnaiah and C. R. Rao, Eds., North Holland Publishing House, 403–425.
- Hanssen-Bauer, I., and E. J. Førland, 1994: Homogenizing long Norwegian precipitation series. *J. Climate*, **7**, 1001–1013.
- Lombard, F., 1994: On a Fourier method of detecting climatic transitions. *Beitr. Phys. Atmosph.*, **67**, 201–208.
- Maronna, R., and V. J. Yohai, 1978: A bivariate test for the detection of a systematic change in mean. *J. Amer. Stat. Assoc.*, **73**, 640–645.
- Potter, K. W., 1981: Illustration of a new test for detecting a shift in mean in precipitation series. *Mon. Wea. Rev.*, **109**, 2040–2045.
- Rhoades, D. A., and M. J. Salinger, 1993: Adjustment of temperature and rainfall records for site changes. *J. Climatol.*, **13**, 899–913.
- Vannitsem, S., and C. Nicolis, 1991: Detecting climatic transitions: Statistical and dynamical aspects. *Beitr. Phys. Atmosph.*, **64**, 245–254.
- Vostrikova, L. J., 1981: Detecting “disorder” in multidimensional random processes. *Sov. Math. Dokl.*, **24**, 55–59.
- Worsley, K. J., 1979: On the likelihood ratio test for a shift in location of normal populations. *J. Amer. Stat. Assoc.*, **74**, 365–367.
- , 1983: Testing for a two-phase multiple regression. *Technometrics*, **25**, 35–42.
- Yao, Y.-C., and R. A. Davis, 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.