Journal of Hydroscience and Hydraulic Engineering
Vol. 20, No.2 November, 2002, 11-26

LONG-TERM TRENDS OF ANNUAL TEMPERATURE AND PRECIPITATION TIME SERIES IN JAPAN

By

Z. X. Xu
K. Takeuchi
H. Ishidaira

Institute of Material and Environmental Technology, Graduate School, Yamanashi University, Kofu, Japan

SYNOPSIS

The availability of using hypothesis test techniques to identify the long-term trends of hydrological time series is investigated in this study. For the purpose to test the long-term trends of hydrological cycle, both parametric and nonparametric test techniques are employed to detect the jump and monotonic trends of the spatially averaged annual temperature and precipitation time series in Japan. Jumps (step trends) were detected from both time series. Although the annual temperature showed an obvious increasing monotonic trend, the precipitation record did not give significant evidence of a monotonic trend. The number of samples required for detecting jumps with a given magnitude at specified significance and power level was estimated with the power function of the t-test. If the magnitude of a jump reaches one or two times of the standard deviation of the time series, the previous 100-year record together with several years of new data will be available for detecting the possible step trend. Otherwise, at least 10 years or more of data record will be required to make significant inference for the plausible jumps.

INTRODUCTION

The increased concentration of greenhouse gases may cause changes in both temperature and precipitation. Their impact on regional hydrological processes may further affect nearly every aspect of human life from municipal and industrial water supplies to food productivity and energy use. The impact of climate change has, therefore, received great attentions, and numerous studies concerning climate change have been conducted over the past several decades. From these studies, potential impacts on the hydrological regime in various geographic areas were investigated and the areas, that could be negatively affected, were identified. For example, test results on temperature and precipitation data from 37 weather stations showed that the Canadian Prairies have become warmer and somewhat drier over the last several decades (see Gan (6)). Zhang et al. (21) found that the annual precipitation totals have changed –10% to 35% in Canada, with the strongest increase in northern Canada, and significant decreasing trends in winter precipitation in southeastern Canada. Another study showed that a jump of precipitation in New South Wales occurred around 1945. A dramatic change in flood risk corresponding to 1945 was apparent: of 41 long-term flood gauges, over 90% show increases in flood frequency, 30% at greater than 99% significance (see Franks (5)). In Japan, two jumps in annual temperature were found around 1920s and 1950s by Yamamoto et al. (19), a jump around 1935 and an increasing trend of the annual maximum daily precipitation over the past 100 years were detected with statistical significance by Yamamoto and Sakurai (20) and Iwashima et al. (10). By using weather pattern simulations, Terakawa et al. (16) found that annual precipitation in Kanto, Chubu, and Seinan inner belt of Japan exhibited a decreasing trend, while annual precipitation in other regions including Hokkaido and Tohoku did not show any long-term trend. In spite of these useful studies focusing on climate change impacts, there still remain considerable unsolved problems concerning both the magnitude and the timing of the impacts on hydrological processes that have occurred, or will occur.

During the past several years the severity and frequency of flooding and prolonged droughts seemed to have increased globally. Therefore, the following questions are posed: Has the increase of hydrological extremes had anything to do with climate change? Although it is quite difficult to understand the causal relationship between hydrological processes and climate change, statistical tests to find a possible causal relationship should be of interest to hydrologists. In previous studies on climate change, the global atmospheric general circulation models (GCMs) have been used to simulate the climate changes (see e.g. Xu (18)). GCM applications showed that the accumulation of greenhouse gases in the atmosphere has resulted in the long-term trend of the temperature and precipitation records (see Kite (11)). The globe exhibited a significant warming trend during the past century, and if climate change is true, what are the likely impacts for the hydrological circulation, and how will the hydrological frequency analysis be affected? Traditionally, the probabilities of occurrence for large floods are estimated by hydrological observations. A probability distribution is usually selected and the distribution is then fitted to historical data using statistical techniques. An important assumption made in this approach is that extreme values are independent and identically distributed with some unchanged underlying probability distribution. In other words, the assumption for hydrological extremes implies that natural climate change does not affect the occurrence of hydrological extremes (see e.g. Bradley (3)). As a preliminary study, recent trends in precipitation and streamflow time series in Japan were investigated by Takeuchi (15). The 30-year records of streamflow and precipitation in the selected 31 major river basins showed a decreasing trend in the streamflow series and an increasing trend in the precipitation series. In the second stage of this research, one of the important objectives of this study is to detect long-term trends in hydrological time series, which may further be used to evaluate the possible effects of climate change on hydrological frequency analysis and water resource systems.

During the past decades, many parametric and nonparametric techniques for the detection of long-term trends in time series were developed and applied (see e.g. Hirsch et al. (8)). Several of these techniques will be used to explore the hypothesis that climate change has occurred gradually or suddenly in Japan during the past century. Next section will first outline the techniques of trends detection in time series. It is followed by a presentation of results for two kinds of data sets in Japan. The number of samples required detecting jumps in these time series is then examined in the next section of the paper. And the paper concludes with a summary and discussion of the findings from this study.

MODEL DESCRIPTION

The hypothesis testing problem may be stated as follows: A null hypothesis H_0 assumes that an event E has not occurred with an alternative hypothesis H_1 that E has occurred. The test statistic T, a function of the data set $X=(x_1, x_2, \ldots, x_n)$, is used to choose the hypothesis between H_0 and H_1. In a trend test, the null hypothesis H_0 means that there is no trend in the population from which the sample X is drawn. In the hypothesis testing for long-term trends, two types of trends are usually considered. One is the monotonic trend that the time series changes monotonically over time. The other is the jump or step trend (see e.g. Hirsch et al. (8)).

Depending on the characteristics of the data being studied, either parametric or nonparametric may be employed. The advantage of nonparametric tests is that few assumptions are made on the data, and it is thereby flexible to handle problems such as seasonality more easily (see e.g. Belle and Hughes, (1)). Furthermore, nonparametric tests generally have higher power than parametric procedures if there is a substantial departure from normality and the sample size is large (see Hipel et al., (7)). Mann-Kendall test for monotonic trends and Mann-Whitney's test for jumps particularly perform well in comparison to the parametric t-tests (see e.g. Berryman et al. (2)). These techniques may be two of the best choices for trend detection.

Parametric t-test for Jump and Monotonic Trends

The two-sided t-test for a step trend in the data vector X is a technique to test whether the means for two partitions of the data set differ significantly. The hypothesis is as follows (see e.g. Lettenmaier, (12)); H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$. On the basis of estimates $\bar{x}_1$ and $\bar{x}_2$ and significance level α, the test statistic is given by

$$T_c = \frac{|\bar{x}_2 - \bar{x}_1|}{T\sqrt{1/n_1 + 1/n_2}} \tag{1}$$

in which

$$T = \sqrt{\frac{(n_1 - 1)s_1^{\ 2} + (n_2 - 1)s_2^{\ 2}}{n - 2}} \tag{2}$$

where n = n_1+n_2, and n_1, n_2 are the numbers of the samples before and after the jump occurs. Hypothesis H_0 or H_1 is accepted if $T_c \le T_{1-\alpha/2,v}$ or $T_c > T_{1-\alpha/2,v}$, respectively. Where $T_{1-\alpha/2,v}$ = quantile of the Student's t distribution with degrees of freedom v = n-2.

The time series X_t, including a linear trend may be expressed by a linear regression model as follows:

$$x_t = \beta_0 + \beta_1 t + \varepsilon_t, \qquad t = 1, 2, ..., n \tag{3}$$

where $\varepsilon_t \ \alpha \ N(0, \sigma_\varepsilon^2)$, denoting a normal distribution with mean zero and variance σ_ε^2, β_1 = trend magnitude, and β_0 = base level. Although the F-test is used to choose a regression model, t-test is usually employed to test the null hypothesis that x_t are independent and identically distributed normal random variables, not dependent on the t. The t-test statistic for monotonic trend is defined as (see Maidment (13))

$$T_c = \left| \frac{r\sqrt{n-2}}{\sqrt{1 - r^2}} \right| \tag{4}$$

where r is the correlation coefficient between X_t and t for t=1, 2, ...,n. H_0 is accepted or rejected if $T_c \le T_{1-\alpha/2.v}$ or $T_c > T_{1-\alpha/2.v}$, respectively. In which $T_{1-\alpha/2,v}$ = the quantile of the Student's t distribution with degrees of freedom v = n-2. Significance level α is selected to be 0.05 (5%) in this study.

Mann-Whitney Test for Jumps

A useful nonparametric approach for jump test is the Mann-Whitney test (see e.g. Mood and Graybill (14)). Given the data vector $X = (x_1, x_2, ..., x_n)$, partition X such that $Y = (x_1, x_2, ..., x_{n_1})$, $Z = (x_{n_1+1}, x_{n_1+2}, ..., x_{n_2+n_2})$, and n=$n_1+n_2$. The n observations may be rearranged in an ascending order as the smallest observation by 1, the second by 2, ..., and the largest one by n. These integers are known as the ranks of the observations. If the sum of the ranks for the n_1 *X* values is denoted by Ty, it can be proven that the mean and variance of Ty are

$$E(Ty) = \frac{n_1(n_1 + n_2 + 1)}{2} \tag{5}$$

$$Var(Ty) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} \tag{6}$$

The exact distribution of Ty is not easily given for large n_1 and n_2. However, Mann and Whitney have proven that Ty is approximately normally distributed for large n_1 and n_2, and the normal approximation may give quite accurate results when n_1 and n_2 are larger than 7. Therefore, one can use the normal approximation with mean and variance given by equations (5) and (6) to find a critical region for testing H_0 under large n_1 and n_2 values (see Mood and Graybill (14)). In other words, following Z_c may be used for testing the jumps,

$$Z_c = \frac{\sum_{t=1}^{n_1} r(x_t) - n_1(n_1 + n_2 + 1)/2}{\left[n_1 n_2 (n_1 + n_2 + 1)/12 \right]^{1/2}} \tag{7}$$

where $r(x_t)$ is the rank of the observation x_t. The null hypothesis H_0 is accepted if $-Z_{1-\alpha/2} \le Z_c \le Z_{1-\alpha/2}$, where $\pm Z_{1-\alpha/2}$ are the standard normal deviates, and α is the significance level for the test.

Mann-Kendall Test for Monotonic Trend

Mann-Kendall test has been used in different fields, and has been found to be an effective tool for identifying monotonic trends in hydrological time series (e.g. Burn (4), Westmacott and Burn (17)). The test statistic of the Mann-Kendall test is given as

$$S = \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} \operatorname{sgn}(x_k - x_i) \tag{8}$$

in which the x_k, x_i are the sequential data values, n is the length of the data set, and

$$\operatorname{sgn}(\theta) = \begin{cases} 1, & \theta > 0 \\ 0, & \theta = 0 \\ -1, & \theta < 0 \end{cases} \tag{9}$$

The theoretical mean and variance of the test statistic for the time series with ties are given as follows:

$$E(S) = 0 \tag{10}$$

$$\operatorname{var}[S] = \frac{n(n-1)(2n+5) - \sum_i e_i(e_i - 1)(2e_i + 5)}{18} \tag{11}$$

in which Σ denotes the summation over all ties and e_i is the number of ties of extent i. For example, in the data set 5, 5, 6, 7, 8, 8, 8, 10, 10, 11, 12, 12, the e_i values are as follows: e_1=3 [three untied values (6, 7, 11)], e_2=3 [three ties of extent two (5, 10, 12)], e_3=1 [one tie of extent three (8)], and for all larger values of i, e_i=0 (see Maidment (13)). For sample sizes larger than ten, the statistic is nearly normally distributed. In other words, the statistic

$$Z_c = \begin{cases} \dfrac{S-1}{\sqrt{\operatorname{var}(S)}}, & S > 0 \\ 0, & S = 0 \\ \dfrac{S+1}{\sqrt{\operatorname{var}(S)}}, & S < 0 \end{cases} \tag{12}$$

is a standard normal variable. In a two-sided test for trend, the null hypothesis H_0 is accepted if $-Z_{1-\alpha/2} \le Z_c \le Z_{1-\alpha/2}$, where $\pm Z_{1-\alpha/2}$ are the standard normal deviates, and α is the significance level for the test.

In addition to testing the trend of time series, it is usually necessary to estimate the magnitude of the trend, which may be defined as a slope, i.e. the change per unit time. The Kendall slope is an unbiased estimator of trend magnitude, and it has higher precision than a regression estimator where data are skewed (see Hirsch et al. (9)). It is defined as follows,

$$\beta = Median\left(\frac{x_i - x_j}{i - j}\right), \qquad \forall j < i \tag{13}$$

in which $1 \le j < i \le n$, and β is a robust estimate of the trend magnitude. In other words, the estimator β is the median over all combination of record pairs for the whole data set and is thereby resistant to the effect of extreme values in the observations (see Hirsch et al. (9)). A positive value of β indicates an increasing trend, and a negative value of β indicates a decreasing trend.

RESULT ANALYSIS AND DISCUSSION

Two types of long-term trends are tested in this study. One is the jump, which consists of a step change in the data record; and the other is a monotonic trend, which consists of a process with mean level varying gradually over the data record. These two types of trends in hydrological processes may be attributed to abrupt and gradual variations due to climate change or other reasons.

Data Analysis

Hydrological observations such as streamflow and water levels are possible choices to detect climate changes, but these time series have several problems: (1) streamflow or water level at a gauge are only representative of a limited area; (2) these series may be susceptible to external influences from urbanization or other changes in the environment; and (3) their fluctuations are usually very large (see e.g. Kite (11)). On the contrary, time series such as spatial averages of temperature and precipitation measure areally various integrated effects and are relatively uninfluenced by human's activities. The temperature and precipitation time series, therefore, are selected for trend tests in this study.

Both temperature and precipitation time series are usually measured continuously at small time intervals and a choice of intervals must be made before long-term trends are tested. The number of data available, the information contained in the data, and the test approaches affect the choice of time intervals. Daily or monthly data are cumbersome when used for periods of over 100 yr, and are also usually unavailable for areal average. For the purpose of investigating the evidence of climate change, spatial averages of the annual data are employed in this study. In order to enhance the spatial coverage of the study area, the stations and gauges are taken from different climatic zones in Japan. Therefore, the hydrological time series selected may reflect a diversity of climatic and physiographic conditions. A minimum of 50 years of record for one data set is desirable to have a reasonable record length for the statistical tests. On the basis of the data condition available, average precipitation over 46 main rain gauges, estimated by Japan Ministry of Land, Infrastructure and Transport on the basis of the meteorological data provided by Japan Meteorological Agency, is used in this study. As for annual temperature, due to its less variability, averaging over main meteorological stations may be sufficient. According to a study made by Terakawa et al. (16), the area investigated can be divided into six climatic zones. The averaging temperature time series, selected over these six different climatic zones, may reflect the diversity of climatic and physiographic conditions. Because several of the stations are located in metropolitan areas, urban effects on trend may be greater than climatic change. Trend detection with better representative data set will be carried out once new data is available at the next stage of this study. Figure 1 shows the study area and the meteorological stations from which temperature data were collected. Both data sets were evaluated qualitatively on the basis of natural conditions to ensure that only stations with good quality data were selected.

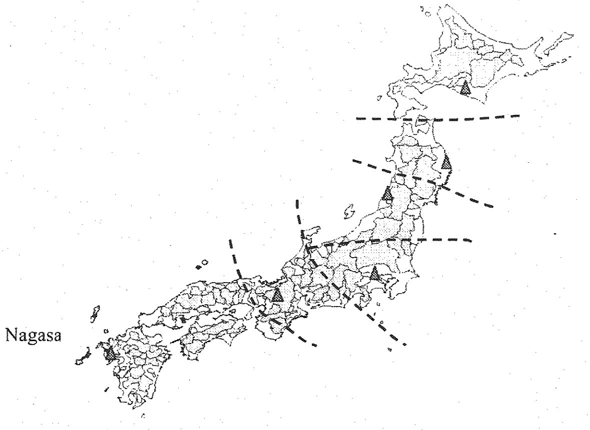

Fig. 1 Map of the study area showing 6 main meteorological stations

The statistics for those two time series are given in Table 1. Figures 2 and 3 show plots of the mean annual observations for those two records. In order to evaluate the non-randomness over a range of timescales, the 10-year moving averages (MA) are also given in those graphs. Although different t-year moving averages are employed, the 10-year period performs better than others. Scrutiny of the time series in Figure 3 reveals that precipitation peaked in the early 1920s and has been declining since 1923. Several patterns emerged in annual precipitation with two distinct regimes; a low precipitation regime from 1924 to 1940 and a significantly higher precipitation regime from 1941 to 1959. The period from 1973 to 1997 was also one of the driest periods during the period of study. Three years: 1978, 1984, and 1994, were the three driest on record; and 16 of the 25 years had below-average precipitation.

Table 1 Statistics for two time series

Statistics	Temperature	Precipitation
Mean	12.87	1634.44
Variance	0.40	24950.40
Standard deviation	0.63	157.96
Coefficient of variation	0.05	0.10
Coefficient of skewness	0.59	-0.27
Maximum value	14.74	1995.00
Minimum value	11.71	1170.00
Range	3.03	825.00

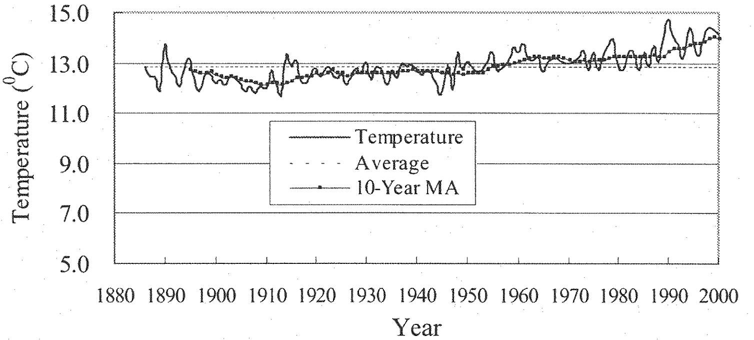

Fig. 2 Annual average temperature over 6 main meteorological stations in Japan

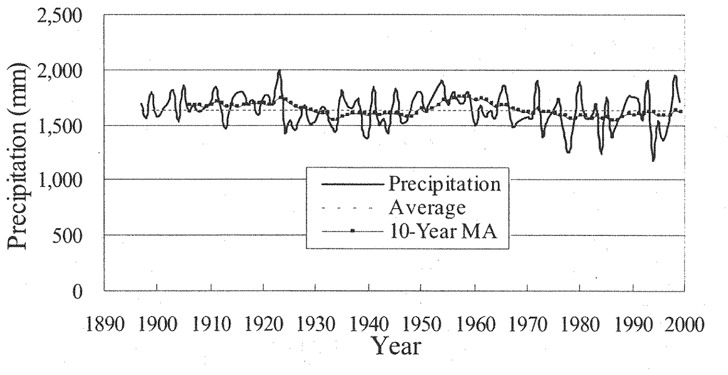

Fig. 3 Annual average precipitation over 46 main rain gauges in Japan

As mentioned above, parametric tests for long-term trends require the error terms ε_t in equation (3): (1) to be independent, (2) to have constant variance, and (3) to come from normal population. The null hypothesis H_0, i.e., the error terms ε_t are independent and identically distributed random variables without dependence on time t, is rejected if $|T_c| > T_{1-\alpha/2,\nu}$, in which $T_{1-\alpha/2,\nu}$ is the point on the Student's t distribution with ν=n-2 degrees of freedom that has the probability of non-exceedance of $1-\alpha/2$. And T_c is estimated by using equation (4) with the combination of the correlation coefficient between noise series ε_t and the corresponding time series t. The estimated T_c statistics are 0.24, and -0.04, and the values $T_{1-\alpha/2,v}$ at significance level $\alpha = 0.05$ are 0.184 and 0.194 for temperature and precipitation time series, respectively. Precipitation time series accept the hypothesis H_0, i.e. the error terms ε_t of the precipitation time series are independent and identically distributed random variables. However, the temperature time series rejects this basic hypothesis.

The test on constant variance is somewhat complex. One alternative is to test for every possible pair of sample series. In other words, the original time series is divided into any two groups and test the variance for those series. However, it is time-consuming and seems unnecessary. In this study, all possible partition points are first identified by vision for each time series. In this visual identification, two principles need to be noted. One is that there is a large difference between the mean values of the two subdivided time series; the other is that any one of the two subdivided series has enough length of record. From this viewpoint of statistics, the dramatic changes of only one or two samples in the record cannot provide sufficient evidence to confirm the occurrence of the jump or monotonic trends. For example, the temperatures in both 1945 and 1947 had small values compared with the neighboring years. However, the values in the years before 1945 or after 1947 have ordinary values. Only those two isolated two-year changes could not provide effective evidence of trends, and this period, therefore, will be excluded from the possible periods to be tested. According to these principles, each one of the two time series are further divided into four sub-series, as given in Tables 2 and 3. In these tables, the statistical features for the subdivided time series are also presented. It is interesting to note that while no large difference was found for the averages of the four subdivided precipitation time series, the standard deviation and the coefficient of variation increased drastically for the period from 1973 to 1999. Although further investigation is required, it is possible that the extreme precipitation occurred more frequently over the past several decades. A test of H_0: $\sigma_1^2 = \sigma_2^2$ versus H_1: $\sigma_1^2 \neq \sigma_2^2$ is two tailed; that is, H_0 should be rejected if $F \geq F_{n_2-1}^{n_1-1}\left(\alpha/2\right)$, in which $F = s_1^2 / s_2^2$. The estimated results for two time series are given in Table 4. It can be inferred from Table 4 that the hypothesis of constant variance for the error terms of both the temperature and the annual precipitation time series is accepted.

Table 2 Partitions of the temperature time series

No.	Time series	Length of record	Mean value	Standard dev.	Coe. of variation
1	1886—1913	28	12.36	0.45	0.04
2	1914—1947	34	12.59	0.34	0.03
3	1948—1988	41	13.13	0.36	0.03
4	1989—2000	12	14.01	0.47	0.03

Table 3 Partitions of the precipitation time series

No.	Time series	Length of record	Mean value	Standard dev.	Coe. of variation
1	1897—1923	27	1696.22	113.48	0.07
2	1924—1940	17	1559.35	123.52	0.08
3	1941—1972	32	1667.22	138.71	0.08
4	1973—1999	27	1581.07	201.86	0.13

Table 4 F Statistics for two time series

Temperature					Precipitation				
No.	Series		Test		No.	Series		Test	
	n_1	N_2	F	H_0		n_1	N_2	F	H_0
1	28	34	1.67	**A**	**1**	27	17	0.86	**A**
2	34	41	1.27	**A**	**2**	17	32	0.80	**A**
3	41	12	0.57	**A**	**3**	32	27	0.47	**A**

Note: R-rejected, A-accepted. Significance level α =0.05.

In order to test the normality of two error term series, the chi-square and Kolmogorov-Smirnov goodness-of-fit tests are employed in this study. Results are given in Table 5. In chi-square goodness-of-fit test, the number of cells into which the observations are to be tallied is set to 6. Obviously the hypothesis H_0 that the error terms are normal distribution is accepted for both temperature and precipitation time series, too. In short, precipitation time series accept three assumptions, and the assumption of independence is rejected by temperature time series. In other words, although trend tests will be implemented for two time series in the following section, the corresponding result is only significant for precipitation time series, while the results for temperature time series may be used for reference.

Table 5 Goodness-of-fit test for two time series detected

	χ^2 test				Smirnov-Kolmogorov test		
	χ^2	ϕ	$\chi^2_{0.05}$	H_0	D_n	$D_n^{(0.05)}$	H_0
Temperature	4.009	3	7.81	**A**	0.069	0.127	**A**
Precipitation	0.631	3	7.81	**A**	0.048	0.134	**A**

Result Analysis

Before a monotonic trend is tested, one has to first identify whether there is any jump in the time series. Table 6 gives the t-test results for two time series. It is clear that jumps occurred in temperature time series around 1914, 1948, and 1989. For the precipitation record, two jumps may have occurred around 1924 and 1941. Because of the limitations from the assumptions in the parametric test approach, nonparametric test techniques should be further used to confirm the existence of the jumps. Table 7 shows the test result of the Mann-Whitney for two time series. Obviously, both nonparametric and parametric approaches give the same results. In other words, three years in temperature time series; 1914, 1948, and 1989 may be regarded as a time around that jumps occurred, as given in Fig. 4. This conclusion is somewhat consistent with the one obtained by Yamamoto et al. (19). The changes of the precipitation time series that occurred in 1924 and 1941 may be regarded as possible years, around that jumps may have occurred, and evidence of this is shown in Fig. 5.

Table 6 t-test results of jumps for two time series

Temperature					Precipitation				
No.	Series		Test		No.	Series		Test	
	n_1	n_2	T_c	H_0		n_1	n_2	T_c	H_0
1	28	34	2.23	**R**	**1**	27	17	3.77	**R**
2	34	41	6.64	**R**	**2**	17	32	2.69	**R**
3	41	12	6.94	**R**	**3**	32	27	1.93	**A**

Table 7 Mann-Whitney test results of jumps for two time series

Temperature					Precipitation				
No.	Series		Test		No.	Series		Test	
	n_1	n_2	Z_c	H_0		n_1	n_2	Z_c	H_0
1	28	34	2.67	**R**	**1**	27	17	3.24	**R**
2	34	41	5.30	**R**	**2**	17	32	2.50	**R**
3	41	12	4.46	**R**	**3**	32	27	1.49	**A**

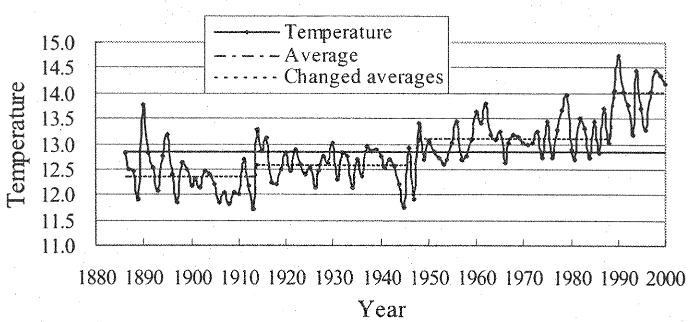

Fig. 4 Average temperature time series showing the jumps

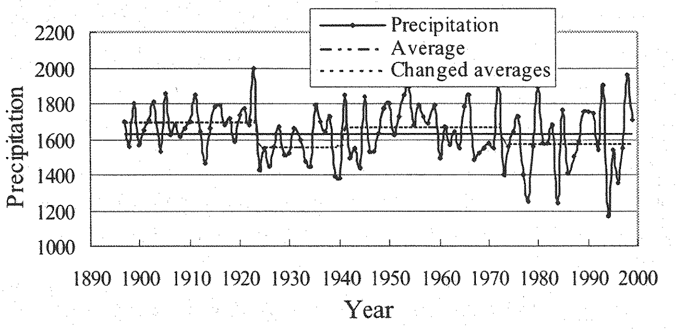

Fig. 5 Average annual precipitation showing the jumps

It should be noted that the hypothesis for the precipitation time series must be re-tested, because the length of the time series n_2 changed due to the absence of the jump for the time series around 1972. The re-tested results obtained by using both t-test and Mann-Whitney techniques are given in Table 8. It is interesting to note that the jump around 1941 disappeared from the plausible jumps detected previously. Due to this change, further detection is still needed. The result is listed in Table 9. Although the hypothesis that there was no jump around 1941 was accepted, the hypothesis that the jump had not occurred around 1923 was rejected. In other words, a jump most likely occurred around 1923 at a significance level 5%. In spite of the fact that the climate background needs to be further investigated, it seems clear that a jump exists in both temperature and precipitation time series from the viewpoint of statistics.

Table 8 Further detection of jumps for the precipitation time series

t-test					Mann-Whitney test				
No.	Series		Test		No.	Series		Test	
	n_1	n_2	T_c	H_0		n_1	n_2	Z_c	H_0
1	27	17	3.77	**R**	**1**	27	17	3.24	**R**
2	17	59	1.51	**A**	**2**	17	59	1.80	**A**

Table 9 Third detection of jumps for the precipitation time series

t-test					Mann-Whitney test				
No.	Series		Test		No.	Series		Test	
	n_1	n_2	T_c	H_0		n_1	n_2	Z_c	H_0
1	27	76	2.42	**R**	**1**	27	76	2.45	**R**

Before the monotonic trend is tested, the linear trend should be first estimated and detected. As given in equation (3), the linear regression equations for two time series are given as follows,

$$T = 12.107 + 0.013t \quad (14)$$

$$P = 1677.386 - 0.826t \quad (15)$$

The results of both t-test and Mann-Kendall test for two time series are summarized in Table 10. The results from two approaches strongly suggest that the null hypothesis, which states that there is no monotonic increasing trend for temperature record, should be rejected. In other words, the temperature time series shows a clear monotonic trend and both parametric and nonparametric tests revealed the same result. An analysis of the linear trend in this time series indicates an increase of 1.32^0C from 1886 to 2000 and the values have several peaks in 1962, 1979, and 1990, although oscillations exist from visual inspection. On the contrary, the long-term monotonic trend in annual precipitations is weak over time. For example, the average annual precipitation decreased 8.3 mm per decade, or 0.5% per decade, corresponding to a decrease of 5-6% throughout the past century. Both t-test and the Mann-Kendall test results show that the hypothesis H_0 is accepted for precipitation time series, and that there is no significant monotonic trend for the precipitation record.

Table 10 Monotonic trend test for two time series

	t-test				Mann-Kendall test		
	β_0	β_1	Tc	H_0	Z_c	β	H_0
Temperature	12.107	0.0132	10.42	**R**	8.247	0.0133	**R**
Precipitation	1677.39	-0.826	1.590	**A**	-1.271	-0.7218	**A**

Estimations of trend magnitudes are also given in Table 10. It reveals that the temperature trend magnitude is positive with a value of 0.013. This means that an increase of almost 1.3°C has taken place in the past century. On the other hand, β value for the precipitation time series is negative and shows the trend of decreasing with time, although this trend could not be identified at significance level 5%. In conclusion, both parametric and nonparametric tests confirm the presence of an increasing trend in temperature time series. However, for the precipitation time series, a significant monotonic trend is not detected by both t-test and Mann-Kendall test. In order to further confirm the plausible monotonic trend occurred in both precipitation and temperature records during the past several decades (1950-2000), which was commonly regarded as an important period with greatest climate change over the several centuries, detections were also performed for the last 50-year records of the precipitation and temperature. Part of the results is given in Table 11. It shows that the temperature shows a significant monotonic trend, but any monotonic trends at the significance level 5% is not detected for the precipitation time series. Findings also show that both time series exhibit stronger monotonic trends during the last 50 years in comparison to that over the last century. The magnitude of the trend for temperature during the last 50 years is twice of that over the last century, and the magnitude of the trend reaches four times for the precipitation time series. In other words, the hydrological cycle has changed more significantly recently. This may have resulted from the greenhouse effect and from the rapid urbanization over the past several decades.

Table 11 Monotonic trend test for the last 50-year time series

	t-test				Mann-Kendall test		
	β_0	β_1	Tc	H_0	Z_c	β	H_0
Temperature	12.754	0.0226	5.609	**R**	4.419	0.0228	**R**
Precipitation	1706.49	-3.070	1.786	**A**	-1.690	-2.9783	**A**

It should be pointed out that, although the temperature time series is an average from six main meteorological stations, these stations extend from the Southern to Northern Japan. The representation of the data records should be sufficient for one to conclude that the increasing trend of the temperature in the last century is not due to noise, but signals, and is most likely an indication of a broad scale response to a systematic mechanism, such as the greenhouse effect and urbanization. However, unlike temperature records, both parametric and nonparametric tests did not provide any positive results of the long-term monotonic trend for the annual precipitation over 46 gauges from 1887 to 1999 and the last 50-year time series. It is unclear whether it represents noise or signals, and it is still too soon to conclude whether the decrease in precipitation is the result of a climate change or partially the result of natural climatic fluctuations.

REQUIRED NUMBER OF SAMPLES TO DETECT TRENDS

When data are sampled from a population, one may occasionally get a sample which is not representative of the original population under study. For example, one may come to the conclusion from a sample that there is a trend in the population but, in fact, there is none. The rejection to a true hypothesis is called a type I error; α is usually used to express the probability to make this type of error. Likewise, one may come to the conclusion that there is not a trend from a sample taken from a population with a trend. Accepting a false hypothesis will cause a type II error and ρ is usually used to express the probability to make this kind of error. The confidence level is defined as the probability ($1-\alpha$) to get the right conclusion when there is no trend in the population. The power of the test, γ, is defined as the probability ($1-\rho$) to get the right conclusion when there is indeed a trend in the population. The number of samples needed to detect the trend with a given magnitude at the specified α and ρ values may be estimated by the power function of the test. The power function for the classical t-test was given by Lettenmaier (12) as follows:

$$\gamma = 1-\rho = F\left(N_t - T_{1-\alpha/2,\nu}\right) \tag{16}$$

where F() = the cumulative distribution function of the standard Student's distribution with ν = n-2 degrees of freedom; $T_{1-\alpha/2,\nu}$ = the quantile of the Student's t distribution at significance level α; N_t = a measure of trend magnitude and is defined as (Berryman et al. (2))

$$N_t \approx \frac{|\Delta\mu|\sqrt{n}}{\sqrt{12}\sigma_\varepsilon} \quad \text{(monotonic trends)} \tag{17}$$

or

$$N_t = \frac{|\mu_1 - \mu_2|\sqrt{n}}{2\sigma_\varepsilon} \quad \text{(jumps)} \tag{18}$$

where $\Delta\mu$ = the change in mean level over the entire length of series; μ_1 and μ_2 = the means before and after the jump occurs; n = the number of samples; and σ_ε = the standard deviation of the noise term with the zero of mean value.

The power of the t-test against either a linear or a step trend may be estimated by equations (16) or the number of samples required for detecting linear or step trends may be calculated by using specified power γ and $N_t - T_{1-\alpha/2,\nu}$. The Student's quantile $T_{1-\alpha/2,\nu}$ is a function of only the sample size and the significance level. As the sample size becomes larger, the Student's distribution approaches a normal distribution and equation (16) may be further simplified as

$$\gamma = F\left(N_t - Z_{(1-\alpha/2)}\right) \tag{19}$$

where $Z_{(1-\alpha/2)}$ = the quantile of the standard normal distribution at non-exceedance probability $(1-\alpha/2)$.

In order to investigate the required sample size to detect trends, the relationship among different parameters in equations (16), (17), and (18) should be examined. This may be quite complex due to too many unspecified parameters. Only several special cases will, therefore, be studied in this paper. Let θ denote the ratio of the change in the mean level over the entire period of series with a monotonic trend or the difference of the means before and after jumps to the standard deviation of error terms ($|\Delta\mu|/\sigma_\varepsilon$ or $|\mu_1 - \mu_2|/\sigma_\varepsilon$), equations (17) and (18) may then be simplified as follows

$$N_t \approx \sqrt{n/12}\cdot\theta \quad \text{(monotonic trends)} \tag{20}$$

or

$$N_t = \sqrt{n}\cdot\theta/2 \quad \text{(jumps)} \tag{21}$$

From equations (19), (20), and (21), some special cases can be easily examined. As an example, one would detect trends where the change in the mean level over the entire length of series with monotonic trend or the difference of the means before and after the jump equals to the standard deviation of error terms (θ =1). The number of samples may then be estimated from equations (19), (20) and (21). For example, it was found that n=153 and n=55 for monotonic trend and jump when α and ρ equal to 0.05. The results for other cases are given in Fig. 6. The graph shows that the required sample size for detecting the monotonic trend is greater than that for detecting jumps under the same power and confidence level. If the power and confidence levels are set to be 0.9 and θ =1.2, a monotonic trend can not be detected within 72 years of sampling data set; however, with only 25 yeas of samples the jump may be detected successfully.

With the combination of the analysis on the error items of both temperature and precipitation time series and the subdivided times series from them, it is found that: θ =0.55~0.85 for temperature series, and θ =0.55~0.88 for precipitation series. For both time series, step and monotonic trends could be detected with 50 and 130 years of record (n_1+n_2) under the power and confidence level of 0.90. Because some approximations were made in equation (19) and the actual time series did not satisfy the conditions of t-test approach, Fig. 6 may be used only as a reference for detecting the plausible jumps or monotonic trends in the time series. These conclusions may be useful for evaluating the impacts of climate change. Unfortunately, the number of samples for other test approaches could not be estimated due to the lack of power functions involved.

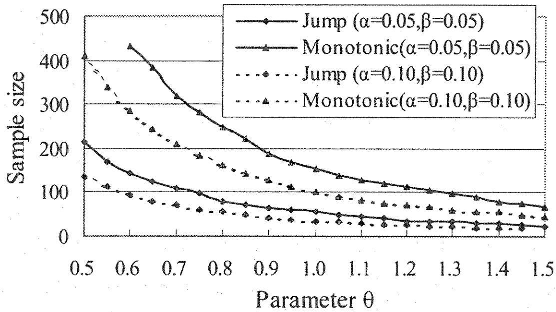

Fig. 6 Changes of sample size with parameter θ

In order to further investigate the relationship between n_1 and n_2, the sample sizes before and after a jump occurs, let $\lambda = \bar{x}_2/\bar{x}_1$, $\eta = \bar{s}_2^{\,2}/\bar{s}_1^{\,2}$, and $\bar{x}_1 = \bar{x}$, $s_1^{\,2} = s^2$, equation (1) may be re-written as follows for a given power of test,

$$\frac{|\lambda - 1|\bar{x}}{s\sqrt{\dfrac{(n_1-1)+\eta(n_2-1)}{n-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} > T_{1-\alpha/2,\upsilon} \qquad (22)$$

Equation (22) may also be further changed to

$$\frac{[(n_1-1)+\eta(n_2-1)](n_1+n_2)}{n_1 n_2 (n-2)} < \varphi \qquad (23)$$

in which

$$\varphi = \left(|\lambda - 1| \Big/ C_V T_{1-\alpha/2,\upsilon} \right)^2 \qquad (24)$$

where Cv is the coefficient of variation for the former half of the time series. As the sample size becomes larger, the Student's distribution approaches a normal distribution, $T_{1-\alpha/2,\upsilon}$ will tend to a constant, which is independent of the degree of freedom ν, and equation (23) may be further simplified as

$$n_2 = f(\lambda, \eta, C_V, n_1) \qquad (25)$$

For any given λ, η, C_v, and n_1, n_2 may be easily estimated. In order to analyze the relationship among λ, η, C_v, n_1, and n_2, the estimated parameters λ, η, and C_v for two time series are presented in Table 12.

Table 12 Parameters λ, η, and C_v for two time series

Parameter	Temperature	Precipitation	Alternatives
λ	1.02—1.07	0.92—1.07	0.90—1.10
η	0.57—1.70	1.19—2.12	0.50—2.00
C_v	0.03—0.04	0.07—0.13	0.03—0.20

The alternative values simulated for parameters λ, η, C_v in this study, resulting from the combination of both values from temperature and precipitation time series, is also given in Table 12, which are determined on the basis of the estimated values of those parameters in the same table. As several of these special cases, Fig. 7

shows the relationship between n_1 and n_2 for different C_v at specified values of λ=0.9 and η=1.0. Obviously the required sample size for detecting jumps will increase with the increase of C_v value. When the previous sample size n_1 is 50, the required size n_2 for detecting jumps will be less than 5 if the C_v of the previous time series is less than 0.1. Otherwise n_2 would be 23 if the C_v increases to 0.2. Large values of C_v imply a greater noise or larger fluctuation for the previous time series and the number of samples required will thereby be large. Fig. 8 presents the graph showing the relationship among n_1, n_2, and λ for specified values of η=1.0 and C_v=0.2. With the decreasing of λ values from 1.0, i.e., the decrease of the magnitude for jumps, the required sample size n_2 will increase dramatically. For example, when the mean value after jump occurs is 0.85 or 1.15 times of the mean value before trend occurs (λ=0.85 or 1.15) and the previous length of the time series is 50 years, the jump will be easily detected once 8 years of new data is recorded. Otherwise, more than 20 years of data will be needed if the change of the jump magnitude becomes smaller and λ value is 0.9 or 1.1. Actually this conclusion corresponds to the results in Fig. 7.

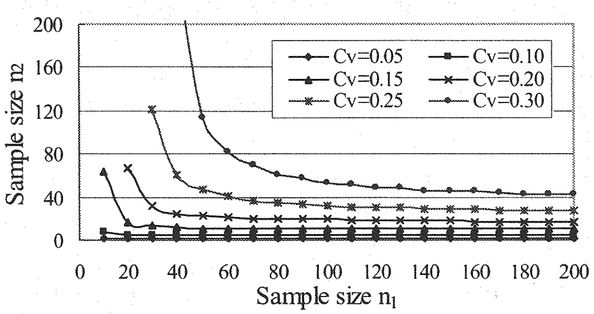

Fig. 7 Changes of the required sample sizes with C_v for λ=0.9 and η=1.0

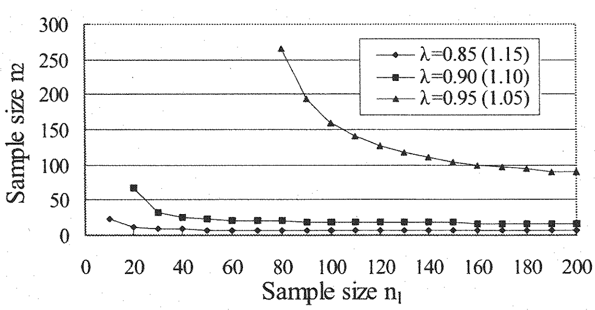

Fig. 8 Changes of the required sample sizes with λ for η=1.0 and C_v=0.2

Fig. 9 shows a more interesting finding. This graph shows clearly that the required number of samples is quite insensitive to the changes of η for specified λ=0.9 and C_v=0.2. In other words, the change of the ratio of variances for the time series does not help one to detect the jump. In conclusion, the required sample size for detecting jumps will be more dependent on the changes of the mean values of the time series as well as on the coefficient of variation of the previous time series, i.e. the magnitude of jump and the statistical features of the time series, but it is much less dependent on the change of the ratio of variances of the time series.

Part of the results represented in Fig. 7 through 9 is further summarized in Table 13. It shows the sample sizes n_2 required to detect jumps for $\bar{x}_2 = \bar{x}_1 \pm k \cdot s_1$ with n_1=50 years. For the temperature time series examined in this paper, parameters λ is 1.02~1.07, η 0.57~1.7, and Cv 0.03~0.04 for the several possible jumps detected, as shown in Table 12. For the precipitation time series, parameters λ, η, and C_v are also presented in the same Table. With the combination of these parameters and the above results, it can be deduced that on the basis of the last 50 years of record another 10 years or more longer of records will be required for detecting new jumps for the temperature and precipitation time series.

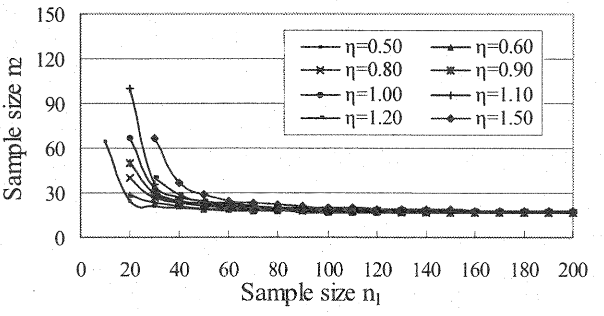

Fig. 9 Changes of the required sample size with η for η =0.9 and C_v=0.2

Table 13 Required sample size n_2 for n_1=50 when $\bar{x}_2 = \bar{x}_1 \pm k \cdot s_1$

s_2^2/s_1^2	K				
	2	1	1/2	2/5	1/3
0.5	≧2	≧5	≧19	≧32	≧53
0.6	≧2	≧5	≧19	≧34	≧60
0.8	≧2	≧5	≧21	≧40	≧78
0.9	≧2	≧5	≧22	≧43	≧92
1.0	≧2	≧5	≧23	≧47	≧113
1.1	≧2	≧5	≧24	≧51	≧145
1.2	≧2	≧5	≧25	≧57	≧204
1.5	≧2	≧5	≧29	≧87	/
1.8	≧2	≧5	≧35	≧179	/
2.0	≧2	≧5	≧40	/	/

It should be pointed out that the conclusions made in this section may only be used for reference due to the lack of strong mathematic basis on the assumption of t-test approach for the hydrological time series, as the results obtained in the third section of this paper. Although these conclusions need to be further validated from the point view of statistics, there is no doubt that some of these results may be helpful for the detecting of jumps or monotonic trends in hydrological processes due to climatic change.

CONCLUSIONS

The feasibility of using hypothesis test techniques to identify the long-term changes of hydrological time series due to climatic changes were investigated in this study. Parametric tests were limited by the assumptions such as the normality and constant variance of the error terms. Nonparametric tests did not have these additional assumptions and were better adapted to the trend test for hydrological time series. Powerful nonparametric tests include the Mann-Kendall test for monotonic trends and the Mann-Whitney test for step trends. Application of these techniques to two kinds of time series in Japan showed that jumps were detected in both temperature and precipitation time series, and that the monotonic trend found in the annual temperature might be regarded as the impact from climate changes and/or urbanization. On the contrary, the changes in precipitation may be within the range of normal fluctuation and could not be definitively ascribed to climate change. While the temperature revealed a strong increase over the last century, there seems to be no significant monotonic trend for precipitation time series in Japan. More samples will be required before a definitive conclusion can be reached as to the extent of the impact of climate change on precipitation. In other words, Japan seems to have become warmer during the past century, but the evidence is insufficient to conclude that climate change has led to much or less precipitation, as found elsewhere in the world. As the number of observations increases, it is necessary to have more effective techniques for detecting jumps or monotonic trends over a wide range of situations. Mann-Kendall and Mann-Whitney tests could be two of these choices. The sample size required for detecting step trends as well as its dependence on different statistical features of the hydrological time series was also investigated. It is believed that the approaches presented in this paper as well as some preliminary conclusions reached in this study could be useful tools for further examining the impacts of climate change on hydrological processes. Finally, it should be pointed out that it is not easy to distinguish between jumps and monotonic trends. It is not clear whether both kinds of trends or only one of them have occurred, and further studies are required to identify these trends more precisely.

REFERENCES

1. Belle, G.V. and J.P. Hughes : Nonparametric tests for trend in water quality, Water Resources Research, Vol.20, No.1, pp.127-136, 1984.
2. Berryman, D.B., Bobée, B., Cluis, D. and J. Haemmerli : Nonparametric tests for trend detection in water quality time series, Water Resources Bulletin, Vol.24, No.3, pp.545-556, 1988.
3. Bradley, A.A. : Regional frequency analysis methods for evaluating changes in hydrologic extremes, Water Resources Research, Vol.34, No.4, pp.741-750, 1988.
4. Burn, D.H. : Hydrologic effects of climate change in west-central Canada, Journal of Hydrology, Vol.160, pp.53-70, 1994.
5. Frank, S.W. : Assessing hydrological change: deterministic general circulation models or spurious solar correlation? Hydrological Processes, Vol.16, pp.559-564, 2002.
6. Gan, T.Y. : Hydroclimatic trends and possible climatic warming in the Canadian Prairies, Water Resources Research, Vol.34, No.11, pp.3009-3015, 1998.
7. Hipel, K.W., Mcleod, A.I., and R.R. Weiler : Data analysis of water quality time series in Lake Erie, Water Resources Bulletin, Vol.24, pp.533-545, 1988.
8. Hirsch, R.M., Alexander, R.B., and R.A. Smith : Selection of methods for the detection and estimation of trends in water quality, Water Resources Research, Vol.27, No.5, pp.803-813, 1991.
9. Hirsch, R.M., Slack, J.R., and R.A. Smith : Techniques of trend analysis for monthly water quality, Water Resources Research, Vol.18, No.1, pp.107-121, 1982.
10. Iwashima, T., Yamamoto, R., and Y. Sakurai, Y., 2002. Long-term trend of extremely heavy precipitation intensity in Japan in recent 100 years. Recent Research Development in Meteorology, No.1, pp.1-9, 2002.
11. Kite, G. : Use of time series analysis to detect climate change, Journal of Hydrology, Vol.111, pp.259-279, 1989.
12. Lettenmaier, D.P. : Detection of trends in water quality data from records with dependent observations, Water Resources Research, Vol.12, No.5, pp.1037-1046, 1976.
13. Maidment, D.R. : Handbook of Hydrology, McGRAW-HILL, Inc., New York, 1992.
14. Mood, M.A., Graybill, F.A., and D.C. Boes : Introduction to the Theory of Statistics, McGRAW-HILL KOGAKUSHA, LTD., 1976.
15. Takeuchi, K. : On the recent trend of river runoff in Japan. In: Takahashi, Y. (ed.), Researches on Integrated Water Management in Urban Areas in Japan: II, Tokyo, pp.23-36, 1993.
16. Terakawa, A., Watanabe, A., and M. Fujikane : Estimation of the change in precipitation over Japan due to global climate change by the weather pattern analysis. Proceedings of the International Conference on Water Resources & Environmental Research, Kyoto, Japan, Vol.II, pp.555-562, 1996.
17. Westmacott, J.R., and D. H. Burn : Climate change effects on the hydrologic regime within the Churchill-Nelson River Basin, Journal of Hydrology, Vol.202, No.1-4, pp.263-279, 1997.
18. Xu, C.Y. : Climate change and hydrologic models: A review of existing gaps and recent research developments, Water Resources Development, Vol.13, pp.369-382, 1999.
19. Yamamoto, R., Iwashima, T., Kazadi, S.N., and M. Hoshiai : Climate jump, Annuals, Disaster Prevention Research Institute, Kyoto University, Vol.29, No.B-1 (in Japanese), pp.297-313, 1986.
20. Yamamoto, R., and Y. Sakurai : Long-term intensification of extremely heavy rainfall intensity in recent 100 years. World Resources Review, Vol.11, No.2, pp.271-281, 1999.
21. Zhang, X.B., Harvey, K.D., Hogg, W.D., and T.R. Yuzyk : Trends in Canadian streamflow. Water Resources Research, Vol.37, No.4, pp.987-998, 2001.

APPENDIX – NOTATION

The following symbols are used in this paper :

C_v	=	Coefficient of variation;
e_i	=	Extent of tie i;
F()	=	Cumulative distribution function of the standard Student's distribution;
H_0	=	Null hypothesis that trend has not occurred;
H_1	=	Alternative hypothesis that trend has occurred;
n	=	Number of observations;
n_1, n_2	=	Sample sizes before and after a jump occurs;
N_t	=	A measure of trend magnitude;
r	=	Correlation coefficient;
S	=	Mann-Kendall statistics;
T_c	=	t-test statistic;
Ty	=	Sum of the ranks of the n_1 X values;
$T_{1-\alpha/2,\nu}$	=	Quantile of the Student's t distribution;
$\bar{x}_1$, $\bar{x}_2$	=	Estimate of means before and after a jump has occurred;
$Z_{1-\alpha/2}$	=	Standard normal deviates;
$\Delta\mu$	=	Change in mean level over the entire length of series;
$1-\alpha$	=	Confidence level;
β	=	Trend magnitude;
β_0	=	Base level for linear trend;
β_1	=	Linear trend magnitude;
γ	=	Power of the test;
η	=	Ratio of variance before and after a jump occurs;
θ	=	Ratio of the change in mean level;
λ	=	Ratio of averages before and after a jump occurs;
μ_1, μ_2	=	Means before and after a jump occurs;
ν	=	Degrees of freedom; and
σ_ε	=	Standard deviation of the noise term.

(Received May 9, 2002 ; revised September 10, 2002)