Distribution functions of a spurious trend in a finite length data set with natural variability: Statistical considerations and a numerical experiment with a global circulation model

S. Nishizawa and S. Yoden

Department of Geophysics, Kyoto University, Kyoto, Japan

Received 19 December 2004; revised 5 January 2005; accepted 8 March 2005; published 18 June 2005.

[1] A linear trend estimated from a finite-length data set with random internal variability has a spurious component which is a difference from the true trend caused by changes in external conditions or parameters. Some moments and distribution functions of the spurious trend depending on the length of data are derived theoretically under general statistical assumptions. When the internal variability has a normal distribution, the spurious trend also has a normal distribution. In general cases of nonnormal distributions, we derive the distribution function of the spurious trend by the Edgeworth expansion. A few low-order moments of the internal variability are necessary to obtain the approximate distribution function from the expansion. Population moments of the internal variability of a simple global circulation model are calculated using a 15,200-year data set generated by a numerical experiment with a purely periodic annual forcing. Dependence of the estimation error of sample moments on the length of data is computed to evaluate an appropriate sample size for each moment. An ensemble experiment with the same model is used to estimate the detectability of a cooling trend in the stratosphere from a finite length data set with internal variability. Hypothesis tests for the statistical significance of the estimated trend are made: Student's *t* test, bootstrap test, and the more accurate test using the distribution function derived by the Edgeworth expansion. In the regions and seasons in which kurtosis of the internal variability is large the assumption that the spurious trend has a normal distribution is not appropriate, and the significance derived by the *t* test is different from that by the test using the Edgeworth expansion.

Citation: Nishizawa, S., and S. Yoden (2005), Distribution functions of a spurious trend in a finite length data set with natural variability: Statistical considerations and a numerical experiment with a global circulation model, *J. Geophys. Res.*, *110*, D12105, doi:10.1029/2004JD005714.

1. Introduction

[2] A global warming trend in the troposphere and a cooling trend in the stratosphere have been estimated with various kinds of observed and analyzed data [e.g., *Intergovernmental Panel on Climate Change*, 2001; *Ramaswamy et al.*, 2001]. Increases of greenhouse gases such as carbon dioxide are considered to be the most important cause of these temperature trends, and the decrease of ozone is equally important for the stratospheric cooling trend [e.g., *Hare et al.*, 2004]. Changes in the frequency or intensity of extreme weather events have also been reported in recent years, such as heavy precipitation [*Iwashima and Yamamoto*, 1993; *Frei and Schär*, 2001; *Osborn and Hulme*, 2002; *Palmer and Rälsänen*, 2002], intense extratropical cyclones [*Graham and Diaz*, 2001], and intense hurricanes [*Landsea et al.*, 1996]. Anthropogenic effects have been proposed as major contributors to these changes or trends [e.g., *Schär and Jendritzky*, 2004; *Hegerl et al.*, 2004]. There is common difficulty to detect such a trend due to limited length of data sets.

[3] Detection of a trend due to external causes from a finite length data set may be difficult when the data contain natural internal variability. In such a case, the estimated trend may have a spurious component which is the difference from the true trend caused by the changes in external conditions or parameters, and we call the component spurious trend. Figure 1 shows a simple example of spurious trends in a model representing a linear trend superimposed on random variability. The estimated trends from some finite length data segments differ from the true linear trend that is externally imposed in the model. For example, the estimated value from the first 5- or 10-year data is positive though the true trend is negative. The difference depends on the length of the data segment which is used to estimate the trend, and on the period of the estimation (or, the time series of random

0148-0227/05/2004JD005714$09.00

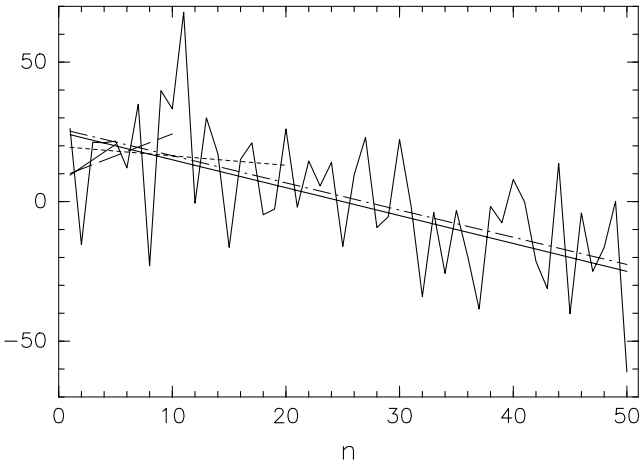

Figure 1. An example of spurious trends. The polygonal line shows a time series of data which has a linear trend represented by the straight thin solid line and random variability. The thick solid, long-dashed, short-dashed, and dash-dotted lines represent estimated trends from the first 5, 10, 20, and 50 years of data, respectively.

variations). In addition to such random internal variations, there are some other possible causes of spurious trends due to the finite length of any data set: rather periodic external forcings with a long period, such as the solar 11-year cycle; intermittent external forcings with a long interval, such as large volcanic eruptions; or, sudden level shift of data quality, such as changes in instruments for observation. Statistical considerations are necessary to interpret correctly the estimated trend that contains spurious trends due to these possible causes.

[4] *Tiao et al.* [1990] and *Weatherhead et al.* [1998] investigated statistical properties of the spurious trend due to internal variability and sudden level shift. It was shown that the detection of linear trends is affected by data length, the magnitude of internal variability, and the autocorrelation of the variations in the data. They derived the standard deviation of the estimated spurious trends; shorter data length, larger standard deviation of internal variability, and higher positive autocorrelation of the variations in the data increase the standard deviation of the spurious trend.

[5] If we can assume that the spurious trend has a normal distribution, we need only its mean and standard deviation in order to specify its distribution function. With this assumption, the Student's t test is often used to estimate the statistical significance of the estimated trend. In general, however, a spurious trend may have various distribution functions, because the internal variability has various distribution functions.

[6] *Taguchi and Yoden* [2002b] performed 1000-year integrations with a simple global circulation model under a purely periodic annual forcing; they obtained highly skewed or bimodal histograms of the internal interannual variations of the polar stratospheric temperature. Linear trend estimated in such a quantity may suffer from the effect of the spurious trend having a nonnormal distribution. The statistical considerations without the assumption of a normal distribution could be applicable to the trends in frequency or intensity of extreme weather events such as heavy precipitation and gusty wind.

[7] In this study, we consider distribution functions of spurious trends due to finite length of data with internal variability, which generally has a nonnormal distribution. In section 2 we describe the derivations of moments and distribution function of the spurious trend under general statistical assumptions. Some Monte Carlo simulations were performed to quantitatively estimate the derivations. In section 3 a 15,200-year data set which was obtained by long time integrations of the model of *Taguchi and Yoden* [2002b] is used to examine dependence of estimation error of sample moments (mean, variance, skewness, and kurtosis) on the data length. The spatial and seasonal variations of these moments of the internal variability in the simple global circulation model are examined. We also performed 96 ensemble runs with the same model in which the radiative heating had a linear cooling trend in the stratosphere. Statistics of the spurious trends are obtained and compared with the theoretical results in section 4. The statistical significance tests of the estimated trend are also argued with these results. Estimation of linear trends with the real atmospheric data and some applications of the present statistical considerations are discussed in section 5. Conclusions are given in section 6.

2. Distribution Function of Spurious Trend: Theoretical Basis

2.1. Estimation of Spurious Trend

[8] A linear trend estimated from a finite length data set contains a spurious trend if it has random internal variability as shown in Figure 1.

[9] We consider a simple linear trend model with random internal variability;

$$X(n) = an + b + \epsilon(n), \qquad n = 1, \cdots, N, \tag{1}$$

where $X(n)$ is an observed value at time n, a is true linear trend, b is a constant, and $\epsilon(n)$ is a random number at time n which gives internal variability. When we estimate a linear trend, $\hat{a}$, by the method of least squares for a finite length of data, N, then the spurious trend, a', is given by

$$a' = \hat{a} - a = S_{nn}^{-1} \sum_{n=1}^{N} \left(n - \frac{N+1}{2} \right) \epsilon(n), \tag{2}$$

where

$$S_{nn} = \sum_{n=1}^{N} \left(n - \frac{1}{N} \sum_{n=1}^{N} n \right)^2 = \frac{N(N+1)(N-1)}{12}. \tag{3}$$

The derivation is given in the Appendix A.

2.2. Moments of Spurious Trend

[10] Let $\epsilon(n)$s be independently identically distributed (i.i.d.) random variables. Then a' has various values for each trial, and we can calculate some moments of the probability of a'. The mean and skewness of the spurious

trend are 0, while the standard deviation, $\sigma_{a'}$, and kurtosis, $\beta_{2a'}$, are given by

$$\sigma_{a'} = \sqrt{\frac{12}{N(N+1)(N-1)}}\sigma_\epsilon \approx \sqrt{12}N^{-\frac{3}{2}}\sigma_\epsilon \quad \text{when N is large enough,} \tag{4}$$

$$\beta_{2a'} = \frac{3}{5}\frac{3N^2-7}{N(N+1)(N-1)}\beta_{2\epsilon} \approx \frac{9}{5}N^{-1}\beta_{2\epsilon} \quad \text{when N is large enough,} \tag{5}$$

where σ_ϵ and $\beta_{2\epsilon}$ are the standard deviation and kurtosis of $\epsilon(n)$, respectively. We can also calculate the moments of higher order. Generally, the moments of odd order are 0, because the probability of a value of a' is identical to that of $-a'$, which can be obtained from the data whose time sequence is in the reverse order. The standard deviation of the spurious trend is approximately proportional to $N^{-3/2}$ times the standard deviation of the internal variability, which result is basically the same as that of *Tiao et al.* [1990] and *Weatherhead et al.* [1998]. The kurtosis of the spurious trend is approximately proportional to N^{-1} times the kurtosis of the internal variability. As N increases, the kurtosis approaches zero at a slower rate than the standard deviation.

2.3. Distribution Function of Spurious Trend

[11] The probability density function (pdf) of a', $f_{a'}(x)$, is an even function, for the same reason that the moments of odd order are 0.

[12] First, we consider the case that $\epsilon(n)$ has a normal distribution with a standard deviation σ_ϵ. Characteristic function (ch.f.) of a', $\psi_{a'}(\omega)$, which is defined as the Fourier transform of the pdf, is derived as

$$\begin{aligned}\psi_{a'}(\omega) &= \int_{-\infty}^{\infty} \exp(i\omega x) f_{a'}(x)dx \\ &= E(\exp(i\omega a')) \\ &= \prod_{n=1}^{N} \psi_\epsilon\left(S_{nn}^{-1}\left(n-\frac{N+1}{2}\right)\omega\right) \\ &= \exp\left(-\frac{1}{2}\omega^2 S_{nn}^{-1}\sigma_\epsilon^2\right),\end{aligned} \tag{6}$$

where $E(x)$ is the expectation of x and $\psi_\epsilon(\omega)$ is the ch.f. of $\epsilon(n)$. By substituting the exact form of (4) into (6), we obtain that a' has a normal distribution with the mean 0 and the standard deviation $\sigma_{a'}$, $N(0, \sigma_{a'}^2)$.

[13] It has been often assumed that the spurious trend has a normal distribution, but the distribution function of the spurious trend depends on that of the internal variability. Here we consider general cases with any symmetric form of the pdf by calculating the Edgeworth expansion of cumulative distribution function (cdf) of the spurious trend. The Edgeworth expansion was originally used in arguments of sample mean [e.g., *Shao* 2003]. Let a_s' be standardized a'; $a_s' = a'/\sigma_{a'}$. The Edgeworth expansion of the cdf of a_s', $F_{a_s'}(x)$, is given by

$$F_{a_s'}(x) = \Phi(x) + \sum_{l=1}^{\infty} Q_l(x)\phi(x)N^{-\frac{l}{2}}, \tag{7}$$

where $\Phi(x)$ and $\phi(x)$ are the cdf and pdf of the standard normal distribution, $N(0, 1)$, respectively. The derivation is given in the Appendix B. For odd l, $Q_l(x)$ is 0, while for even l, $Q_l(x)$ is a polynomial of degree at most $2l - 1$ with coefficients depending on the first $l + 2$ cumulants of $\epsilon(n)$. The cumulants, κ_k, are statistics related to the moments; κ_1, κ_2, $\kappa_3\kappa_2^{-3/2}$, and $\kappa_4\kappa_2^{-2}$ are the mean, variance, skewness, and kurtosis, respectively. The $Q_l(x)$s for the first two even ls are

$$Q_2(x) = -\frac{3}{40}\frac{\kappa_4}{\kappa_2^2}H_3(x) = -\frac{3}{40}\beta_{2\epsilon}\left(x^3 - 3x\right), \tag{8}$$

$$Q_4(x) = -\frac{3}{560}\frac{\kappa_6}{\kappa_2^3}H_5(x) - \frac{9}{3200}\frac{\kappa_4^2}{\kappa_2^4}H_7(x), \tag{9}$$

where $H_k(x)$ is kth Hermite polynomial.

[14] Distribution of the spurious trend converges to a normal distribution asymptotically as N increases (cf. the central limit theorem). This expansion is useful for considering convergence speed, which for a normal distribution is $O(N^{-1})$, because $Q_1(x) = 0$. In general, when the convergence speed to a distribution is $O(N^{-l/2})$, the distribution has lth-order accuracy. For example, the normal distribution has second-order accuracy in the case of the spurious trend.

[15] We can obtain approximate distributions of the spurious trend from the Edgeworth expansion which are more accurate than the normal distribution. If we know the standard deviation and kurtosis of the random internal variability, we can obtain the approximate cdf with fourth-order accuracy, and also obtain the approximate pdf by differentiating the cdf.

[16] The approximate pdfs of a' with fourth-order accuracy are shown by solid curves for $N = 10$, 20, and 50 in Figure 2, in the case that $\sigma_\epsilon = 1$ and $\beta_{2\epsilon} = 6$. Histograms of a' obtained by Monte Carlo simulations are also shown in Figure 2: we replicated the calculation of a' with i.i.d. random numbers having a Weibull distribution with scale parameter of 1 and shape parameter of 1, *We*(1, 1), whose standard deviation and kurtosis are 1 and 6, respectively. The internal variability of the real atmosphere may have such a value of kurtosis, as shown in section 3. Though there are small differences in the tails of the histogram and the approximate pdf for $N = 10$, they show a good correspondence for larger N. The histogram of a' itself becomes closer to a normal distribution as N increases; the difference is discernible for $N = 50$ around the peak but is not for $N = 100$ (not shown).

2.4. Difference Between Distribution Function of Spurious Trend and Normal Distribution

[17] The difference between the cdf of the spurious trend and normal distribution, $|\Delta F(x)|$, is

$$|\Delta F(x)| = \left|F_{a_s'}(x) - \Phi(x)\right| \approx |Q_2(x)\phi(x)|N^{-1} \quad \text{when N is large enough,} \tag{10}$$

because of (7). The maximum of $|\Delta F(x)|$ is approximately $0.04|\beta_{2\epsilon}|N^{-1}$ from (8). When the data length, N, is larger than $4|\beta_{2\epsilon}|$, the difference is less than 0.01. To an accuracy of 1%, we can assume that the spurious trend has a normal distribution for that data length.

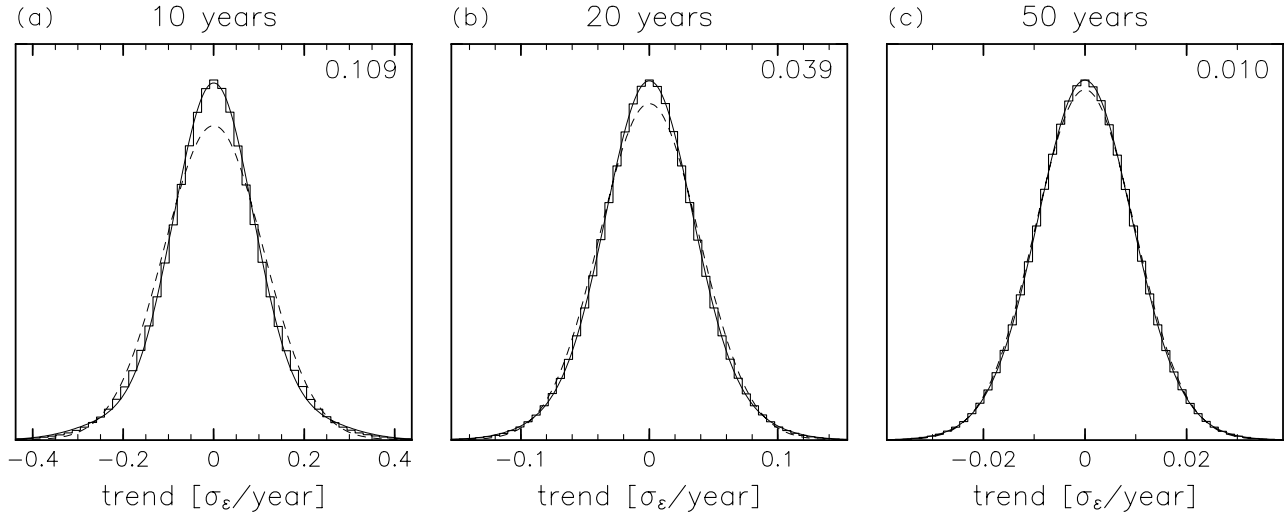

Figure 2. The pdf of the spurious trend derived by the Edgeworth expansion up to $O(N^{-1})$ represented by solid curve and the histogram derived by the Monte Carlo simulation (replication number is 10,000,000) in which $\epsilon(n)$ has *We*(1, 1) from the (a) 10-, (b) 20-, and (c) 50-year data. The dashed line represents the pdf of the normal distribution, whose standard deviation is the same as that of the histogram. The number at the top right corner is the standard deviation of the histogram.

[18] An inverse Edgeworth expansion is referred to as a Cornish-Fisher expansion [e.g., *Shao* 2003]. It is an expansion of percentile, $w_\alpha = F_{a's}^{-1}(\alpha)$, and the expansion is given by

$$w_\alpha = z_\alpha + \sum_{l=1}^{\infty} q_l(z_\alpha) N^{-\frac{l}{2}}, \tag{11}$$

where $z_\alpha = \Phi^{-1}(\alpha)$. For odd l, $q_l(x)$ is 0, while for even l, $q_l(x)$ is a polynomial depending on the first l $Q_k(x)$s. The $q_l(x)$s for the first two even ls are

$$q_2(x) = -Q_2(x), \tag{12}$$

$$q_4(x) = Q_2(x)\frac{d}{dx}Q_2(x) - \frac{1}{2}x\{Q_2(x)\}^2 - Q_4(x). \tag{13}$$

When $z_\alpha = \pm\sqrt{3}$, $q_2(x)$ is 0 because of (8) and (12), and $\alpha_- = \Phi(-\sqrt{3}) \approx 0.04$, $\alpha_+ = \Phi(\sqrt{3}) \approx 0.96$. Then $w_{0.04}$ and $w_{0.96}$ is approximately equal to $z_{0.04}$ and $z_{0.96}$, respectively. Namely, the difference between a 92% confidence interval of normal distribution, $[z_{0.04}, z_{0.96}]$, and that of the distribution of a', $[w_{0.04}, w_{0.96}]$, is approximately 0. For the confidence level larger than 92%, the difference becomes larger as the level approaches 100%.

2.5. Statistical Significance Test and Interval Estimation

[19] When a' has a normal distribution, a statistic t, which is the Studentized a', $t = a'/s_{a'}$, has a t distribution with degree of freedom $N - 2$, $t(N - 2)$, where $s_{a'}$ is the estimation of $\sigma_{a'}$ determined by the estimation of σ_ϵ, s_ϵ:

$$s_{a'}^2 = S_{nn}^{-1} s_\epsilon^2, \tag{14}$$

$$s_\epsilon^2 = \frac{1}{N-2}\sum_{n=1}^{N}\left\{X(n) - \left(\hat{a}n + \hat{b}\right)\right\}^2, \tag{15}$$

where $\hat{b}$ is the estimation of b derived by the method of least squares.

[20] If we know the cdf of a', we can obtain the interval estimation of the trend and test the statistical significance of the estimated trend. The t distribution, $t(N - 2)$, is often used for the interval estimation and hypothesis test (Student's t test), with the assumption that the spurious trend has a normal distribution. We can also use the distribution obtained by the Edgeworth expansion for the interval estimation and statistical significance test without any prerequisite for the distribution function. This result can be also used to validate the result with the t distribution.

[21] In addition to the method with the t distribution, the bootstrap method, which is a resampling method [*Efron*, 1979], is used for the interval estimation and statistical significance test [*Hall*, 1988]. This method requires many trials to obtain enough samples, which is now possible as a consequence of the advancement of computing facilities. The bootstrap method is also evaluated in section 4.

3. Numerical Experiment on Moments of Atmospheric Internal Variability

[22] The spurious trend depends on the nature of the random internal variability. Because the atmospheric internal variability often has a nonnormal distribution, we need to know some moments of the atmospheric internal variability in order to obtain distribution function of the spurious trend due to the atmospheric variations. Usually, the time series of atmospheric observations are not long enough to fully characterize the internal variability. Therefore we perform a long time integration with a global circulation model to examine such internal variability in the model atmosphere.

3.1. Model and Experiment

[23] The atmospheric model used in this study is the same as that used by *Taguchi and Yoden* [2002a]. It is based on a

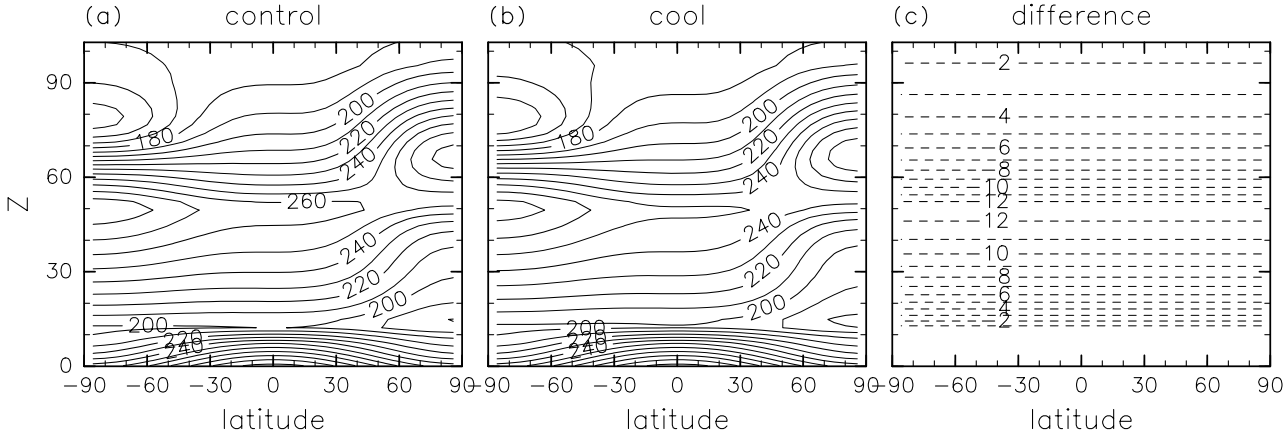

Figure 3. The radiative equilibrium temperature profile for the Newtonian heating/cooling in midwinter (a) in the control experiment and (b) at the last year in the trend experiment and (c) the difference between Figures 3a and 3b.

three dimensional primitive equation model (Swamp Project, AGCM5 (in Japanese), available at http://www.gfd-dennou.org/arch/agcm5/). The horizontal resolution is a T21 spherical harmonic truncation and it has 42 levels from the surface to the mesosphere. Some physical processes were simplified in the model; Newtonian heating/cooling, dry atmosphere with no moist processes, dry convective adjustment, and Rayleigh friction at the bottom and some upper levels.

[24] The radiative equilibrium temperature for the Newtonian heating/cooling in midwinter in the control experiment is shown in Figure 3a. A sinusoidal surface topography of zonal wave number 1 was assumed in the Northern Hemisphere, and the amplitude was set to 1000 m. Ten 1520-year model runs were made after 1-year spin up integration from each initial state of an isothermal atmosphere (250 K) at rest with small disturbances. This gives 15,200 years of data, and we analyze only the monthly averaged zonal mean temperature.

3.2. Estimation Error of Sample Moments

[25] Sample moments estimated from finite length data sets have errors from the population moments of variability. The error depends on the data length and the distribution of the internal variability. Figure 4 shows histograms of the monthly mean polar temperature at 2.6 hPa. The climatological mean of the polar temperature is high in summer and low in winter and shows a smooth seasonal variation. It is much higher than the prescribed radiative equilibrium temperature from late fall to early spring because of large wave driving of planetary waves propagated from the troposphere. The distribution also depends on season. The standard deviation is larger in winter than in summer: it is 15.5 K in January and 0.4 K in July. The polar temperature in winter has a highly skewed distribution with a large fraction of low-temperature months near the radiative equilibrium temperature. The large deviations to warmer side reflect the occurrence of stratospheric sudden warmings [*Yoden et al.*, 2002]. The warmer side tail of the distribution is elongated, so that the mean of the distribution is higher than the mode of the distribution. Therefore both of the skewness and kurtosis are positive and large. In January and February, the distribution has two modes. In the tropo-

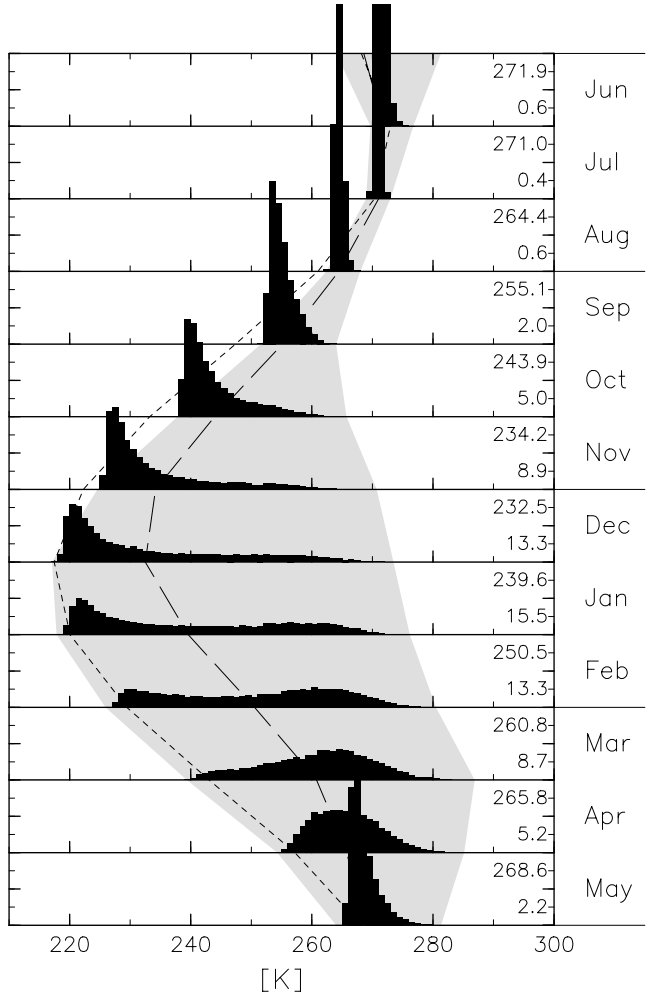

Figure 4. Dependence of the histogram of the monthly averaged zonal mean temperature at 86°N and 2.6 hPa on month. The mean at each month and the monthly mean radiative equilibrium temperature are connected by long-dashed and short-dashed lines, respectively. The region between the maximum and minimum temperature of each month is shaded. The two numbers at right-hand side of each month are the mean (top number) and the standard deviation (bottom number).

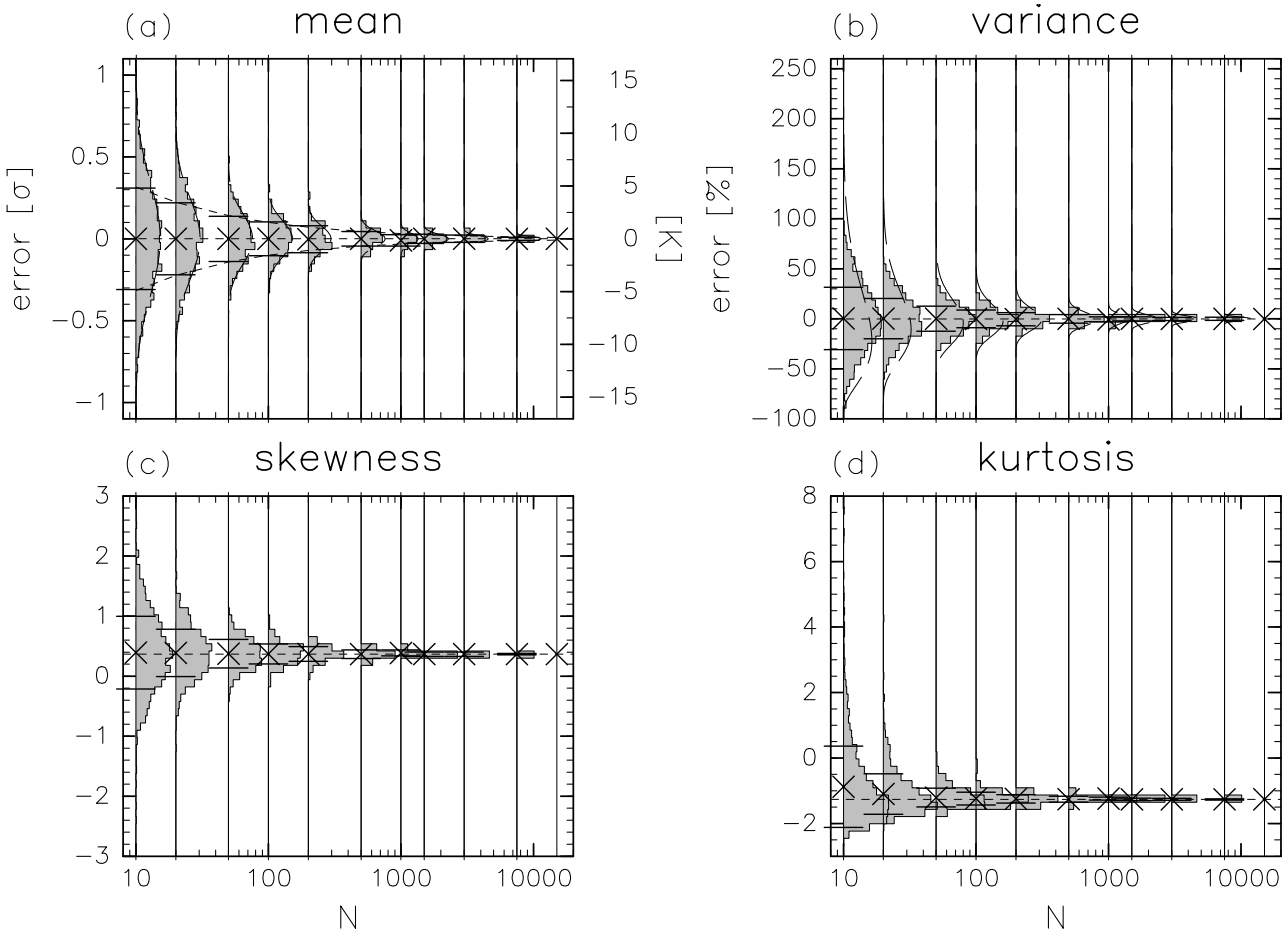

Figure 5. The histograms of (a) error of sample mean normalized by the standard deviation derived from the 15,200-year data, (b) rate of sample variance to the variance derived from the 15,200-year data, (c) sample skewness, and (d) sample kurtosis of the monthly mean polar temperature at 2.6 h Pa in January for several data length, N. The crosses and bars denote the mean and standard deviation of each histogram, respectively. The thick dashed line in Figure 5a represents the theoretical line of the standard deviation, and the thin dashed lines in Figures 5a and 5b represent the pdf of the normal distribution and $\chi^2(N-1)$, respectively.

sphere, on the other hand, the monthly mean temperature has an almost normal distribution throughout of the year (not shown).

[26] Though the shape of the histogram is similar to that derived from a 1000-year data set [*Taguchi and Yoden*, 2002b], 1000 years may be insufficient to estimate the moments of the distribution because of irregularities of the histograms. Here we examine dependence of the estimation error of sample moments on the data length. We calculate the sample moments from $\lfloor 1{,}520/N \rfloor \times 10$ nonoverlapping continuous N-year data for some Ns, where $\lfloor x \rfloor$ is the floor function that gives the largest integer less than or equal to x.

[27] Figure 5a shows histogram of the normalized difference of the sample mean from the mean derived from the 15,200-year data. The values are normalized by the standard deviation derived from the 15,200-year data. The histogram is similar to the pdf of a normal distribution. The standard deviation of the histogram becomes smaller as N increases. It almost corresponds to the theoretical result, $\sigma N^{-1/2}$ [e.g., *Shao*, 2003], where σ is the population standard deviation. With this result, we can roughly estimate the magnitude of the error of the sample mean. For example, the standard deviation of internal variability is about 15 K in the winter stratosphere, so that the error of sample mean is about 3.4 K ($\approx 15\ \mathrm{K} \times 20^{-1/2}$) with 20 years of data and 2.1 K ($\approx 15\ \mathrm{K} \times 50^{-1/2}$) with 50 years of data.

[28] The error of the sample variance also becomes smaller as N increases (Figure 5b). In theory, error of sample variance has a χ^2 distribution with degree of freedom $N-1$, $\chi^2(N-1)$, when the variability has a normal distribution. In general cases, the standard deviation of the sample variance is $\sqrt{\beta_2 N^{-1} + 2(N-1)^{-1}}$ times the population variance, where β_2 is the population kurtosis [e.g., *Kenney and Keeping*, 1951]. For example, when $N \sim 100(\beta_2 + 2)$, the error of sample variance is about 10%, and when $N \sim 10(\beta_2 + 2)$ the error is about 30%. The histogram has some deviation from the χ^2 distribution, and the standard deviation of the histogram is smaller than that of the χ^2 distribution because of the negative kurtosis. The histogram in the troposphere has good agreement with the χ^2 distribution (not shown).

[29] The standard deviation of the histogram of the sample skewness also becomes smaller as N increases

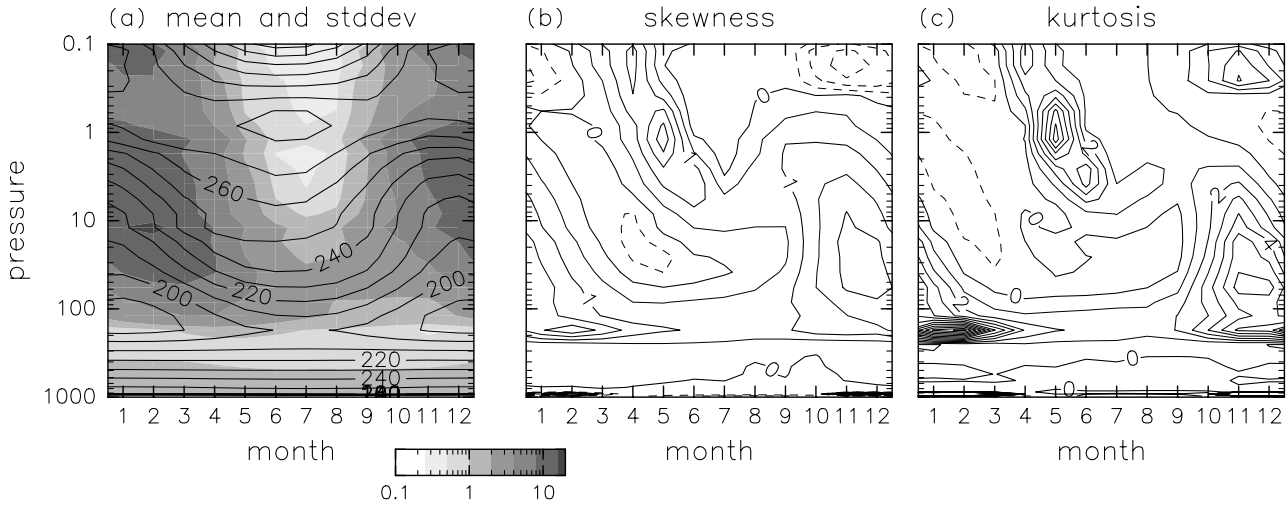

Figure 6. Meridional sections of (a) mean (contours) and standard deviation (shading), (b) skewness, and (c) kurtosis of the monthly mean polar temperature. The contour interval is 10 K, 0.5, and 1 for the mean, the skewness, and the kurtosis, respectively.

(Figure 5c). The mean of the histogram for small N may differ from the skewness derived from the 15,200-year data, because sample skewness is not an unbiased estimator of population skewness. In this case, the standard deviation of the sample skewness derived by least squares fitting is about $2N^{-0.55}$, and the error of sample skewness is about 0.1 for $N \sim 250$ and about 0.4 for $N \sim 20$.

[30] The sample kurtosis also has a similar dependence on N (Figure 5d). Sample kurtosis is also not an unbiased estimator of population kurtosis, so that the mean of the histogram for small N is different from the kurtosis derived from the 15,200-year data. The histogram has large skewness in both the stratosphere and the troposphere (not shown). In this case, the standard deviation of the sample kurtosis derived by least squares fitting is about $4N^{-0.6}$, and the error of sample kurtosis is about 0.1 for $N \sim 500$ and about 0.7 for $N \sim 20$.

3.3. Spatial and Seasonal Distribution of Moments

[31] The moments of the atmospheric internal variability depend on longitude, latitude, height, and season. We study spatial and seasonal distribution of the moments of the internal variability with the 15,200-year data.

[32] The year-to-year variation is the largest in the polar region at almost every month and level. Figure 6 shows month-pressure sections of some moments of monthly averaged zonal mean temperature in the Northern polar region. The standard deviation has a maximum of about 16 K around 12 hPa in February, and it is small in the summer upper stratosphere. The large standard deviation is associated with the occurrence of stratospheric sudden warmings. In fall it is largest around 1 hPa, and the pressure level at which it is the largest shifts down to about 100 hPa in summer. In the troposphere, it is almost constant through the year. In midlatitudes, the pattern of the standard deviation is almost the same as that at high latitudes but it is largest in December and the maximum value is about 8 K around 12 hPa (not shown).

[33] The skewness and kurtosis are large in late fall and winter in the lower stratosphere and around 1 hPa in May. The maximum value of the skewness is about 2, while the kurtosis is above 6 in the winter lower stratosphere and around 1 hPa in May. In these regions and seasons, the variability has a nonnormal distribution. In midlatitudes, they are nearly 0, and the distribution of internal variability is nearly normal (not shown).

4. A Numerical Experiment on Spurious Trend

[34] Since the internal variability depends on space and season, so does the spurious trend. We performed an ensemble experiment on the spurious trend with the same model, in which the radiative heating was assumed to have a linear cooling trend in the stratosphere.

4.1. Experimental Design

[35] The radiative equilibrium temperature for the Newtonian heating/cooling was linearly changed for 50 years from that in the control experiment to a final profile shown in Figure 3b, which was made with a smaller static stability $N^2 = 4.75 \times 10^{-4}$ s^{-2} in the stratosphere. Figure 3c shows the difference of the radiative equilibrium temperature profile between the initial and the final years. The externally given trend is about -0.25 K/year around 50 km.

[36] After 1-year spin-up integration with the radiative temperature of the control experiment, a 96-member ensemble of 50-year integrations was performed with the trend in radiative temperature.

4.2. Dependence of Spurious Trend on Space and Season

[37] Figure 7a shows an example of time sequences of the polar temperature in January for all the 96 runs: it is a monthly mean at 12 hPa. The externally given trend is about -0.18 K/year at this level. The range of the year-to-year variation is about 60 K, and the highest temperature reached in the last year is much higher than the lowest temperature in the first year. Thus we can find out that some runs show warming trend for 50 years.

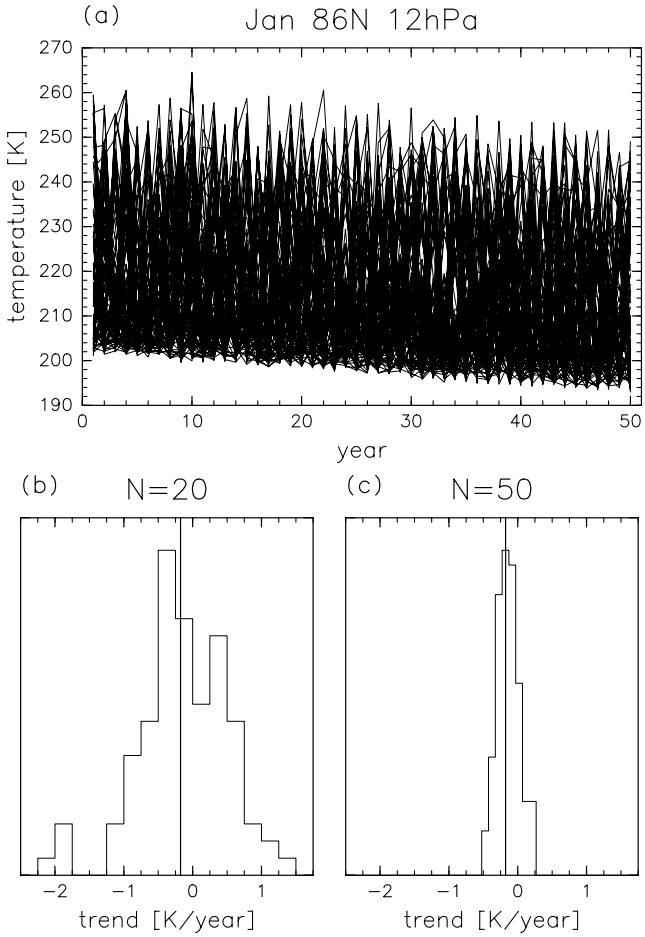

Figure 7. (a) Time sequences of the monthly mean polar temperature at 12 hPa in January in all the 96 runs, (b) histogram of the estimated trend from the first 20-year data, and (c) that from the 50-year data. The vertical lines in Figures 7b and 7c denote the value of the externally imposed trends.

[38] Histograms of the estimated linear trend by the method of least squares with the first 20- and 50-year data are shown in Figures 7b and 7c, respectively. Though the ensemble average of the estimated trend is −0.13 K/year for the first 20 years and −0.14 K/year for the 50 years, the estimated trend for each ensemble member has both signs: 39 of the 96 runs showed a warming trend for the 20-year estimation, and 18 runs for the 50-year estimation. The ensemble average approaches the given value and the variability of the spurious trend becomes smaller, as the data length increases.

[39] Figure 8 shows time sequences of the polar temperature and histograms of the estimated linear trend in July. The year-to-year variation is much smaller than that in winter. The variation range is about 8 K, and it is comparable to the externally given cooling trend for 50 years. The histograms show the estimated trend converges to the externally given trend for the 50-year data.

[40] The distribution of the estimated spurious trends depends on space and season. Figure 9 shows month-pressure sections at the polar region of the ensemble mean of the estimated trend and the standard deviation of the spurious trends for the first 20- and 50-year data. The ensemble mean in the winter stratosphere is different from the externally given trend (see Figure 3c), and the difference is smaller for the 50-year data. The standard deviation of the spurious trends is also large in the winter stratosphere, indicating that the ensemble mean of the estimated trend has a larger error than in the summer. The pattern of the standard deviation is similar to that of the standard deviation of the internal interannual variability (Figure 6a), and the magnitude of the standard deviation for the 50-year data is about $(50/20)^{-3/2} \sim 0.25$ of that for the first 20-year data. These results are consistent with the theoretical result (Equation 4).

[41] In middle and low latitudes the ensemble mean of the estimated trends is not very different even for 20-year data in all the season, and the standard deviation of the spurious trend is also small (not shown).

4.3. Statistical Significance of Estimated Trend

[42] Hypothesis tests for statistical significance of the estimated trends are performed for each cooling trend run. We use the t test and the bootstrap test, and compare the results with the more accurate test using the distribution function derived by the Edgeworth expansion up to $O(N^{-1})$ (hereafter we call this test Edgeworth test).

[43] Figure 10 shows the statistical significance of the linear trend estimated from the first 20-year data derived by

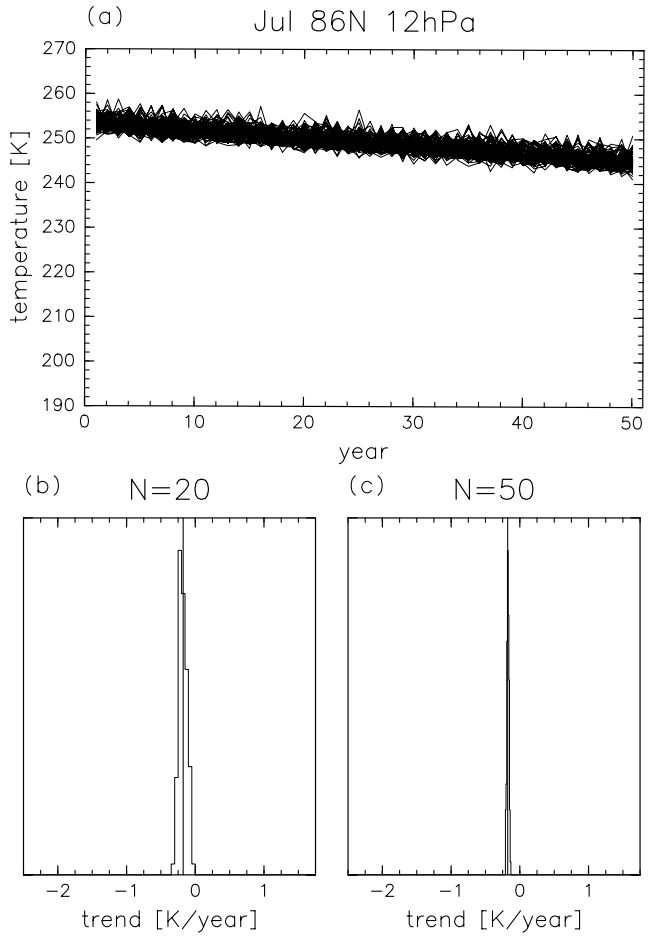

Figure 8. Same as Figure 7 but in July.

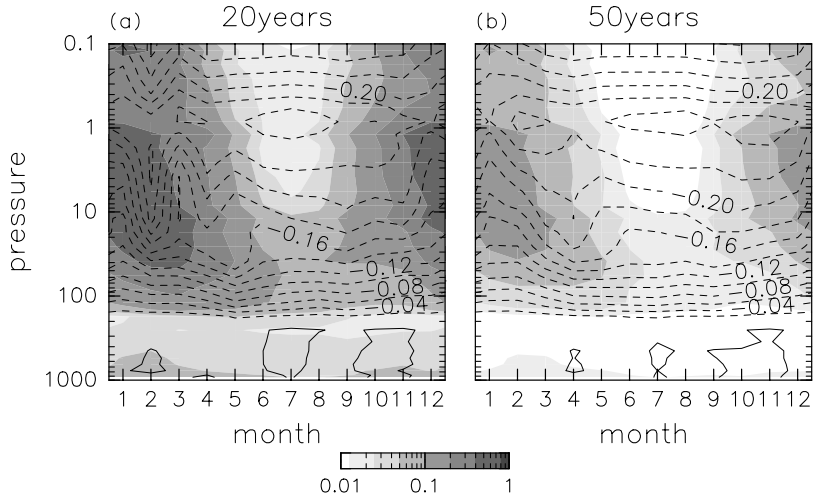

Figure 9. Meridional sections of (contours) the ensemble mean of the estimated trend and the standard deviation (shading) of the spurious trend of the monthly mean polar temperature from (a) the first 20 and (b) 50 years of data.

the three tests in the run for which the difference between the results by the *t* test and the Edgeworth test is largest. Here the difference is defined as the number of grid points at which the significance is higher than 99% by one test and lower than 99% by the other test. In the summer stratosphere, the estimated linear trend normalized by the standard deviation of the spurious trend, which is also estimated by (4) with the 20-year data for the *t* test and bootstrap test and with the 15,200-year data for the Edgeworth test, is large and it is significant (>99%) in all three tests. The statistical significance is not very different between these tests, though the *t* test and bootstrap test overestimate the statistical significance in April and May in the upper stratosphere.

[44] In winter, on the other hand, the normalized estimated trend is small, because the internal variability is large. It is impossible to obtain a statistically significant wintertime trend estimation with 20-year data set.

[45] If the kurtosis of the internal variability is large and the time series is not long enough, the *t* test cannot be used because the assumption of a normal distribution of spurious trend is not satisfied. Locations (height and month) where this occurs are indicated by the symbol × in Figure 10a. The *t* test shows the significance higher than 99% in May in the upper stratosphere, but careful assessment is necessary because of the large kurtosis. Indeed, the Edgeworth test shows a smaller significance. Large values of the kurtosis

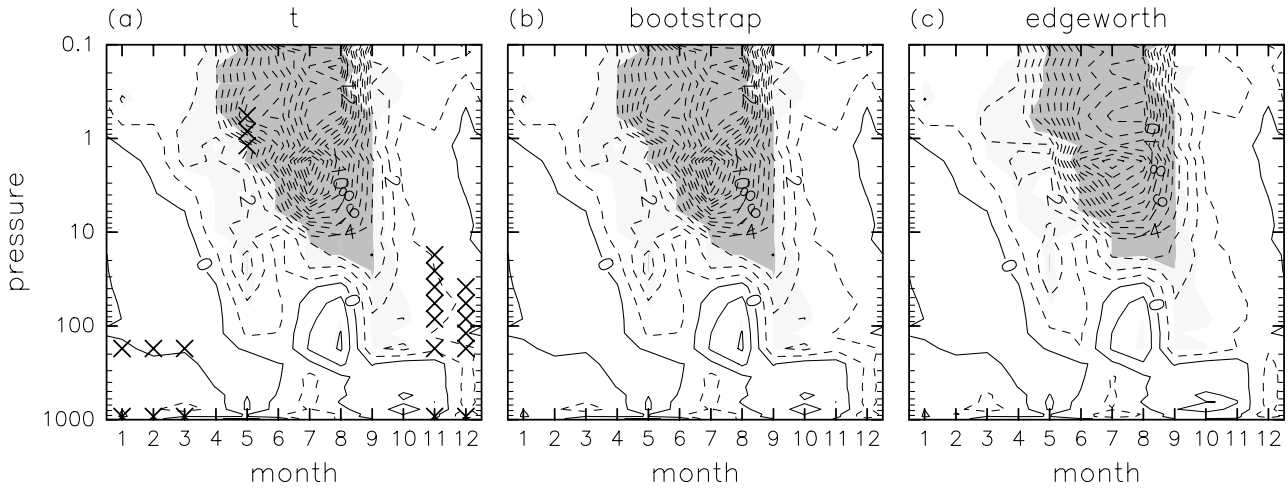

Figure 10. Meridional sections of the estimated trend (contours) and the statistical significance (shading) of the estimated trend of polar temperature in one run derived by (a) the *t* test, (b) the bootstrap test, and (c) the Edgeworth test. The estimated trend for the first 20 years is normalized by the standard deviation of the spurious trend estimated by (4) with the sample standard deviation of the internal variability in Figures 10a and 10b and by the standard deviation obtained by the control experiment in Figure 10c. The light and dark shading represent regions with significance which is larger than 90% and 99%, respectively. The crosses in Figure 10a denote grid points at which the normal distribution approximation is not valid because of large kurtosis.

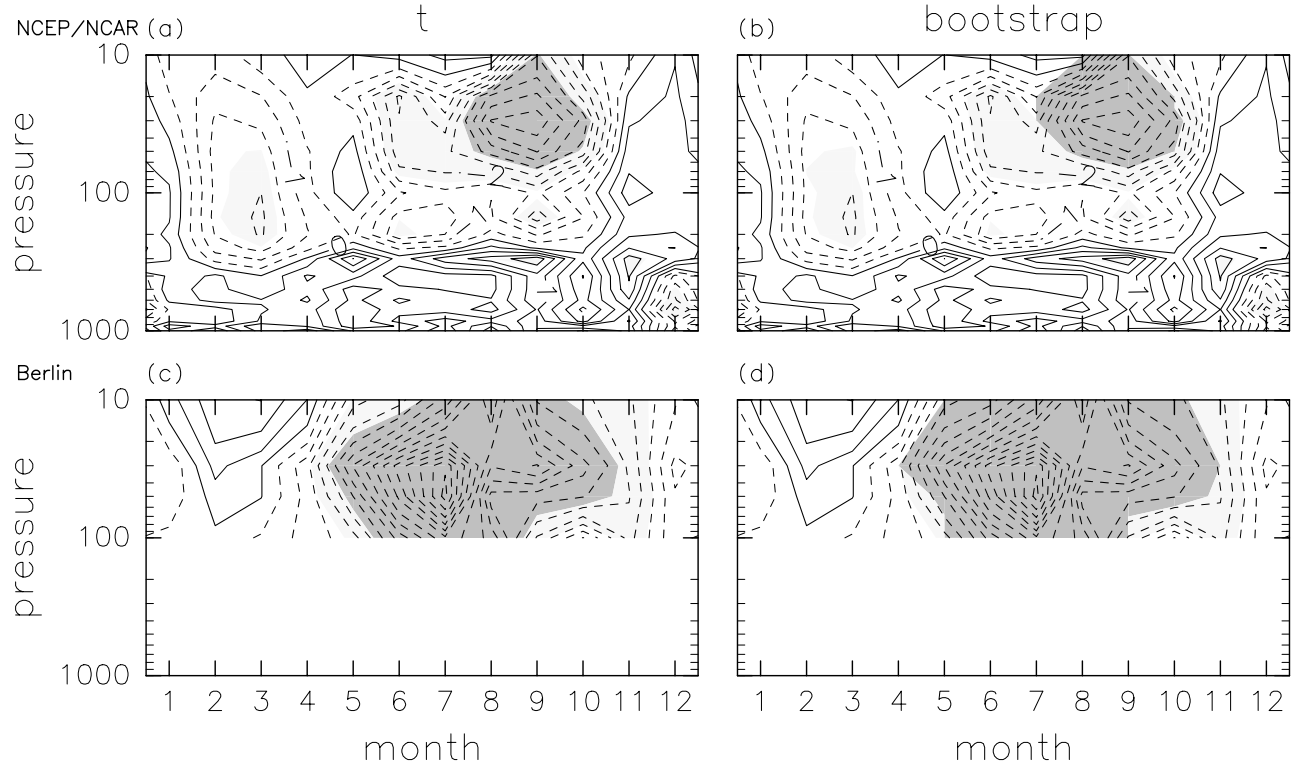

Figure 11. Meridional sections of the Studentized estimated trend (contours) and its statistical significance (shading) derived by the *t* test (Figures 11a and 11c) and bootstrap test (Figures 11b and 11d) of the monthly mean polar temperature in (a and b) the NCEP/NCAR and (c and d) Berlin data.

are also obtained in the winter lower stratosphere, though the estimated trend is not significant due to large standard deviation of the internal variability.

5. Discussion

[46] In the estimation of a linear trend from the real atmospheric data, global and annual averaging has been often applied to obtain a statistically significant result by reducing the standard deviation of the internal variability. However, internal interannual variations have large seasonal dependence in the polar stratosphere as shown in Figures 4 and 6. Estimation of a trend for limited space and season might be important for studies of some localized phenomena such as the ozone hole in spring over Antarctica. Figure 11 shows height and seasonal dependence of the Studentized estimated trend and its statistical significance obtained from the NCEP/NCAR reanalysis data [*Kalnay et al.*, 1996] for 20 years from 1981 to 2000 and from the Berlin Free University stratospheric data [*Labitzke et al.*, 2002] for 39 years from 1963 to 2001. As in the numerical result shown in Figure 10, the normalized trend is large in the summer stratosphere and the estimated cooling trend is statistically significant, in both of the *t* test and the bootstrap test. The significance is higher in the Berlin data than in the NCEP/NCAR data, because of the longer data length, particularly from May to July. *Randel and Wu* [1999] obtained similar height and seasonal dependence of significant cooling of the polar stratosphere by computing the mean temperature differences for the time segments between 1970s and 1990s.

[47] *Labitzke and van Loon* [1994] and *Labitzke et al.* [2005] examined temperature trend with the Berlin data. They noted that the trend can be appreciably different depending on the segment of the data due to large fluctuations on a decadal scale. In the winter stratospheric polar region, as we have shown, the internal variability is so large that the estimated trend has appreciable spurious component and is not statistically significant. It is difficult to distinguish the signal of variations on a decadal scale from the spurious trends due to the high frequency internal variability. Therefore careful argument is necessary for the variation of trends on a decadal scale.

[48] It is difficult to estimate moments of internal variability of the real atmosphere, such as the standard deviation and kurtosis, because the time series of observed data is not long enough; the error of the sample standard deviation is about 10 to 20%, and that of the sample kurtosis is about 0.4 to 0.7 for the data length of 20 to 40 years (see section 3.2). Significance derived by the *t* test may be untrustworthy when the kurtosis of the internal variability is large and erroneous, because we cannot assume that the spurious trend has a normal distribution. On the other hand, the Edgeworth test is not always more accurate than the other tests, because the sample moments of the internal variability may have considerable errors from the population moments. This is the reason why the Edgeworth test was not applied for the real atmospheric data sets in this study. Under these circumstances, long integrations of state-of-the-art general circulation models (GCMs) with purely periodic annual forcing might be a useful way to estimate the moments of the atmospheric internal variability. With the moments

obtained by such a control experiment, approximate distribution function of the spurious trend can be derived by the Edgeworth expansion, and statistical significance of the linear trend estimated with the real atmospheric data can be tested by the Edgeworth test with the moments obtained by GCM experiments.

[49] The internal variability of the monthly averaged zonal mean temperature does not have large kurtosis in summer when the estimated trend is statistically significant. In the winter stratosphere, on the other hand, the variation has a nonnormal distribution (Figure 4). *Hio and Yoden* [2005] pointed out the importance of such nonnormal distribution of stratospheric variability when they assessed the rarity of the stratospheric sudden warming event in the Southern Hemisphere firstly observed in 2002. In the atmospheric variability, there are some other variations having nonnormal distribution. For example, precipitation approximately has Gamma distribution [e.g., *Wilks and Eggleston*, 1992], wind speed approximately has Weibull distribution [e.g., *Conradsen et al.*, 1984; *Pavia and O'Brien*, 1986], and extreme weather events such as heavy rain or gusty wind approximately have extreme value distributions [e.g., *Gumbel*, 1958]. Spurious trends of these variations may have a nonnormal distribution, and the t test may give an untrustworthy measure of significance. Statistical significance of the estimated trend should be tested by the Edgeworth test.

[50] *Hare et al.* [2004] examined stratospheric temperature trends by comparing two 5-member ensembles of 20-year Unified Model transient runs with or without a linear ozone trend. They adopted a Student's t test to discuss the statistical significance of the difference of the ensemble mean trend between the two ensemble runs, under the prerequisite that the estimated trends in each ensemble run have a normal distribution. This assumption can be justified by the present theory stated in section 2: The distribution of the estimated trend is normal when the internal variability has a normal distribution or length of data is large enough. They argued the distribution of the internal variability was approximately normal. On the basis of the statistics, they also discussed the ensemble size to obtain significant estimate in the presence of large internal variability. Similarly, our 15,200-year data set to study the internal variability can be used to answer the question how may years are needed to obtain significant linear trend in the presence of large variability.

[51] *Meehl et al.* [2000] argued changes in extreme events with the changes in the mean and standard deviation of variability in the case that the variation has a normal distribution. However, in order to discuss the changes in frequency or intensity of these extreme events, one must consider nonnormal distributions for each variability. Statistical considerations on the spurious trend in section 2, particularly careful treatments on the tails of the distributions, might be useful for such arguments.

[52] Some atmospheric variations have a bimodal distribution, such as the westerly or easterly phase of the equatorial quasi-biennial oscillation (QBO). If we try to apply the present method to such data sets, we have to consider the mechanism that produces the bimodal distribution. If there are multiple equilibrium states or regimes which have the same response to a small change in external forcing, the distribution of the internal variability is only shifted without changing its shape. Then the present method could be applied for the trend argument. More generally, each regime may have independent response to the change in external forcing. In such a case, the present method cannot be applied. The trend argument should be much more complicated.

6. Conclusions

[53] Distribution functions of a spurious trend due to the finite length of data with random internal variability were investigated theoretically and numerically. Here the spurious trend was defined as the difference between an estimated trend and the true trend caused by changes in external conditions and parameters. Some moments and distribution functions of the spurious trend were derived theoretically in a simple linear trend model with random internal variability. The moments of odd order are 0, and the standard deviation is proportional to the standard deviation of the random internal variability and decreases approximately as $N^{-3/2}$ for the increase of data length, N. The kurtosis is proportional to the kurtosis of the random internal variability and decreases approximately as N^{-1}. The distribution function of the spurious trend depends on the distribution function of the internal variability. The spurious trend has a normal distribution when the internal variability has a normal distribution. For general cases of nonnormal distributions, we derived the Edgeworth expansion of the distribution function of the spurious trend. The distribution function converges to a normal distribution as N^{-1}. By the Edgeworth expansion we can obtain the approximate distribution function with a few low-order moments of the internal variability. For example, the standard deviation and kurtosis of the internal variability are needed for fourth-order accuracy.

[54] Some atmospheric variations have a nonnormal distribution; for example, the polar temperature in the winter stratosphere shows highly skewed or bimodal distributions. In order to examine such random internal variability, we performed a long integration with a simple global circulation model and 15,200-year monthly averaged zonal mean temperature data were analyzed. The error of sample moments estimated with finite length data sets was examined and the dependence of the estimation error on the data length was obtained (Figure 5).

[55] We also investigated spatial and seasonal distributions of the moments of the internal variability. The standard deviation in polar regions is large in the winter stratosphere, while it is small in summer in the upper stratosphere (Figure 6). The skewness and kurtosis are large in the winter lower stratosphere and around 1 hPa in May, so that the distribution of the internal variability shows a nonnormal distribution.

[56] An ensemble experiment on spurious trend was performed with the same model, in which the radiative heating was assumed to have a linear cooling trend in the stratosphere. The 96 runs for 50 years show that the standard deviation of the spurious trend has similar pattern to that of the standard deviation of the internal interannual variability with the magnitude depending on the data length (Figure 9), which is consistent with the theoretical result.

The standard deviation of the spurious trend is large in the winter polar stratosphere.

[57] Hypothesis tests for statistical significance of the estimated trend were also performed by the t test and the bootstrap test, and the results were compared with the more accurate test with the distribution function derived by the Edgeworth expansion (Figure 10). The statistical significance is high in the summer upper stratosphere where the internal variability is small, while the significance is low in the winter stratosphere. The t test cannot be used in the regions and seasons in which the kurtosis of the internal variability is large, because the assumption that the spurious trend has a normal distribution is not satisfied. On the other hand, the test using the Edgeworth expansion does not need any prerequisite for the distribution.

[58] Possible application of the present statistical considerations on the spurious trends to the changes in frequency or intensity of extreme weather events was discussed. Such variations are described by nonnormal distributions, and careful treatment of the tails of distribution is necessary.

Appendix A: Spurious Trend by the Method of Least Squares

[59] In the simple trend model represented by (1), let estimators of a and b obtained by the method of least squares be $\hat{a}$ and $\hat{b}$, respectively. Then the sum of squares of residual,

$$L = \sum_{n=1}^{N} \left\{ X(n) - \left(\hat{a}n + \hat{b} \right) \right\}^2, \tag{A1}$$

satisfies

$$\frac{\partial L}{\partial \hat{a}} = -2 \sum_{n=1}^{N} n \left\{ X(n) - \left(\hat{a}n + \hat{b} \right) \right\} = 0, \tag{A2}$$

$$\frac{\partial L}{\partial \hat{b}} = -2 \sum_{n=1}^{N} \left\{ X(n) - \left(\hat{a}n + \hat{b} \right) \right\} = 0. \tag{A3}$$

After some algebra, we have

$$\hat{a} = S_{nn}^{-1} \sum_{n=1}^{N} \left(n - \frac{N+1}{2} \right) X(n), \tag{A4}$$

$$\hat{b} = \frac{1}{N} \sum_{n=1}^{N} X(n) - \frac{1}{2}(N+1)\hat{a}. \tag{A5}$$

Because of (1), we obtain

$$\hat{a} = a + S_{nn}^{-1} \sum_{n=1}^{N} \left(n - \frac{N+1}{2} \right) \epsilon(n). \tag{A6}$$

Then the spurious trend a', which is the difference between $\hat{a}$ and a, is

$$a' = \hat{a} - a = S_{nn}^{-1} \sum_{n=1}^{N} \left(n - \frac{N+1}{2} \right) \epsilon(n). \tag{A7}$$

Appendix B: Edgeworth Expansion of Spurious Trend

[96] We can derive the Edgeworth expansion of a'_s analogously to that of sample mean [e.g., *Shao*, 2003]. The ch.f. of a'_s is obtained as

$$\psi_{a'_s}(\omega) = E\left(\exp\left(i\omega a'_s\right)\right) = \prod_{n=1}^{N} \psi_\epsilon \left(\frac{1}{S_{nn}^{-1/2}\sigma_\epsilon} S_{nn}^{-1} \left(n - \frac{N+1}{2} \right) \omega \right). \tag{B1}$$

Because of the definition of cumulant,

$$\log \psi_\epsilon(\omega) = \sum_{j=1}^{\infty} \frac{1}{j!} \kappa_j (i\omega)^j, \tag{B2}$$

$\psi_{a's}(\omega)$ becomes

$$\prod_{n=1}^{N} \exp\left\{ \sum_{j=1}^{\infty} \frac{1}{j!} S_{nn}^{-j/2} \left(n - \frac{N+1}{2} \right)^j \frac{\kappa_j}{\kappa_2^{j/2}} (i\omega)^j \right\} \tag{B3}$$

$$= \exp\left(-\frac{\omega^2}{2} \right) \left[1 + \frac{3}{40} \frac{\kappa_4}{\kappa_2^2} (i\omega)^4 N^{-1} + \left\{ \frac{3}{560} \frac{\kappa_6}{\kappa_2^3} (i\omega)^6 + \frac{9}{3200} \frac{\kappa_4^2}{\kappa_2^4} (i\omega)^8 \right\} N^{-2} + \cdots \right]. \tag{B4}$$

We consider a polynomial, $P_l(\omega)$, which satisfies

$$\psi_{a'_s}(\omega) = \sqrt{2\pi}\phi(\omega) \sum_{l=0}^{\infty} P_l(i\omega) N^{-l}. \tag{B5}$$

Then $P_l(\omega)$ of the first three ls are

$$P_0(\omega) = 1, \tag{B6}$$

$$P_1(\omega) = \frac{3}{40} \frac{\kappa_4}{\kappa_2^2} \omega^4, \tag{B7}$$

$$P_2(\omega) = \frac{3}{560} \frac{\kappa_6}{\kappa_2^3} \omega^6 + \frac{9}{3200} \frac{\kappa_4^2}{\kappa_2^4} \omega^8. \tag{B8}$$

[97] Because of

$$\int_{-\infty}^{\infty} (ix)^k \phi(x) \exp(i\omega x) dx = \sqrt{2\pi} \frac{d^k}{d\omega^k} \phi(\omega), \tag{B9}$$

the cdf of a'_s is

$$F_{a'_s}(x) = \sum_{l=0}^{\infty} P_l\left(\frac{d}{dx}\right)\Phi(x)N^{-l}. \tag{B10}$$

We consider a polynomial, $Q_l(x)$, which satisfies

$$F_{a'_s}(x) = \Phi(x) + \sum_{l=1}^{\infty} Q_l(x)\phi(x)N^{-\frac{l}{2}}. \tag{B11}$$

Then we obtain

$$Q_{2m-1}(x) = 0, \tag{B12}$$

$$Q_2(x) = -\frac{3}{40}\frac{\kappa_4}{\kappa_2^2}H_3(x), \tag{B13}$$

$$Q_4(x) = -\frac{3}{560}\frac{\kappa_6}{\kappa_2^3}H_5(x) - \frac{9}{3200}\frac{\kappa_4^2}{\kappa_2^4}H_7(x), \tag{B14}$$

where m is positive integer.

Notation

a	true trend.
$\hat{a}$	estimated trend.
a'	spurious trend.
a'_s	standardized spurious trend.
b	constant term.
$\hat{b}$	estimation of b.
$E(x)$	expectation of x.
$F_{a'_s}(x)$	cumulative distribution function of a'_s.
$f_{a'}(x)$	probability density function of a'.
$H_k(x)$	kth Hermite polynomial.
L	sum of squares of residual.
N	length of data set in year.
$N(\mu, \sigma^2)$	normal distribution with mean μ and standard deviation σ.
n	time in year.
$s_{a'}$	estimation of $\sigma_{a'}$.
S_{nn}	sum of squares of n subtracted its mean.
s_ϵ	estimation of σ_ϵ.
t	Studentized spurious trend.
$t(\nu)$	t distribution with degree of freedom of ν.
$We(a, b)$	Weibull distribution with scale parameter of a and shape parameter of b.
w_α	100α% percentile of $F_{a'_s}(x)$.
$X(n)$	observed data at time n.
z_α	100α% percentile of $\Phi(x)$.
$\beta_{2a'}$	kurtosis of a'.
$\beta_{2\epsilon}$	kurtosis of $\epsilon(n)$.
$\epsilon(n)$	internal variability at time n.
κ_k	kth cumulant of $\epsilon(n)$.
σ_ϵ	standard deviation of $\epsilon(n)$.
$\sigma_{a'}$	standard deviation of a'.
$\Phi(x)$	cumulative distribution function of standard normal distribution.
$\phi(x)$	probability density function of standard normal distribution.
$\psi_{a'}(\omega)$	characteristic function of a'.
$\psi_{a'_s}(\omega)$	characteristic function of a'_s.
$\psi_\epsilon(\omega)$	characteristic function of ϵ.
$\chi^2(\nu)$	χ^2 distribution with degree of freedom of ν.

[98] **Acknowledgments.** We are most grateful to the editor for the helpful comments. Present graphic tools are based on the codes in the GFD-DENNOU Library (SGKS Group, DCL-5.3 (in Japanese), available at http://www.gfd-dennou.org/library/dcl/). Calculations were performed on VPP800 system of Academic Center for Computing and Media Studies, Kyoto University, and VPP5000 system of Information Technology Center, Nagoya University. The first author (S.N.) is supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists. This research was supported in part by the Kyoto University Active Geosphere Investigations for the 21st Century COE (KAGI21), which was approved by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan.

References

Conradsen, K., L. B. Nielsen, and L. P. Prahm (1984), Review of Weibull statistics for estimation of wind speed distributions, *J. Clim. Appl. Meteorol.*, *23*, 1173–1183.

Efron, B. (1979), Bootstrap methods: Another look at the jackknife, *Ann. Stat.*, *7*, 1–26.

Frei, C., and C. Schär (2001), Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region, *J. Clim.*, *14*, 1568–1584.

Graham, N. E., and H. F. Diaz (2001), Evidence for intensification of North Pacific winter cyclones since 1948, *Bull. Am. Meteorol. Soc.*, *82*, 1869–1893.

Gumbel, E. J. (1958), *Statistics of Extremes*, 375 pp., Columbia Univ. Press, New York.

Hall, P. (1988), Theoretical comparison of bootstrap confidence intervals, *Ann. Stat.*, *16*, 927–953.

Hare, S. H. E., L. J. Gray, W. A. Lahoz, A. O'Neill, and L. Steenman-Clark (2004), Can stratospheric temperature trends be attributed to ozone depletion?, *J. Geophys. Res.*, *109*, D05111, doi:10.1029/2003JD003897.

Hegerl, G. C., F. W. Zwiers, P. A. Stott, and V. V. Kharin (2004), Detectability of anthropogenic changes in annual temperature and precipitation extremes, *J. Clim.*, *17*, 3683–3700.

Hio, Y., and S. Yoden (2005), Interannual variations of the seasonal march in the Southern Hemisphere stratosphere for 1979–2002 and characterization of the unprecedented year 2002, *J. Atmos. Sci.*, *62*, 567–580.

Intergovernmental Panel on Climate Change (2001), *Climate Change: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*, edited by J. T. Houghton et al., 944 pp., Cambridge Univ. Press, New York.

Iwashima, T., and R. Yamamoto (1993), Long-term trend of heavy daily precipitation, *J. Meteorol. Soc. Jpn.*, *71*, 37–640.

Kalnay, E. M., et al. (1996), The NCEP/NCAR reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471.

Kenney, J. F., and E. S. Keeping (1951), *Mathematics of Statistics*, part 2, 2nd ed., 429 pp., Van Nostrand Reinhold, Hoboken, N. J.

Labitzke, K., and H. van Loon (1994), Trends of temperature and geopotential height between 100 and 10 hPa on the Northern Hemisphere, *J. Meteorol. Soc. Jpn.*, *72*, 643–652.

Labitzke, K., et al. (2002), The Berlin stratospheric data series [CD-ROM], report, Meteorol. Inst., Free Univ. Berlin, Berlin.

Labitzke, K., B. Naujokat, and M. Kunze (2005), The lower Arctic stratosphere in winter since 1952: An update, *SPARC Newsl. 24*, pp. 27–28, Stratospheric Processes and Their Role in Clim., Geneva, Switzerland.

Landsea, C. W., N. Nicholls, W. M. Gray, and L. A. Avila (1996), Downward trends in the frequency of intense Atlantic hurricanes during the past five decades, *Geophys. Res. Lett.*, *23*, 1697–1700.

Meehl, G. A., et al. (2000), An introduction to trends in extreme weather and climate events: Observations, socioeconomic impacts, terrestrial ecological impacts, and model projections, *Bull. Am. Meteorol. Soc.*, *81*, 413–416.

Osborn, T. J., and M. Hulme (2002), Evidence for trends in heavy rainfall events over the UK, *Philos. Trans. R. Soc. London*, *360*, 1313–1325.

Palmer, T. N., and J. Rälsänen (2002), Quantifying the risk of extreme seasonal precipitation events in a changing climate, *Nature*, *415*, 512–514.

Pavia, E. G., and J. J. O'Brien (1986), Weibull statistics of wind speed over the ocean, *J. Clim. Appl. Meteorol.*, *25*, 1324–1332.

Ramaswamy, V., et al. (2001), Stratospheric temperature trends: Observations and model simulations, *Rev. Geophys.*, *39*, 71–122.

Randel, W. J., and F. Wu (1999), Cooling of the Arctic and Antarctic polar stratospheres due to ozone depletion, *J. Clim.*, *12*, 1467–1479.

Schär, C., and G. Jendritzky (2004), The European heatwave of 2003: Was it merely a rare meteorological event or a first glimpse of climate change to come? Probably both, is the answer, and the anthropogenic contribution can be quantified, *Nature*, *432*, 559–560, doi:10.1038/432559a.

Shao, J. (2003), *Mathematical Statistics*, 2nd ed., 591 pp., Springer, New York.

Taguchi, M., and S. Yoden (2002a), Internal interannual variability of the troposphere-stratosphere coupled system in a simple global circulation model. Part I: Parameter sweep experiment, *J. Atmos. Sci.*, *59*, 3021–3036.

Taguchi, M., and S. Yoden (2002b), Internal interannual variability of the troposphere-stratosphere coupled system in a simple global circulation model. Part II: Millennium integrations, *J. Atmos. Sci.*, *59*, 3037–3050.

Tiao, G. C., G. C. Reinsel, D. Xu, J. H. Pedrick, X. Zhu, A. J. Miller, J. J. DeLuisi, C. L. Mateer, and D. J. Wuebbles (1990), Effects of autocorrelation and temporal sampling schemes on estimates of trend and spatial correlation, *J. Geophys. Res.*, *95*, 20,507–20,517.

Weatherhead, E. C., et al. (1998), Factors affecting the detection of trends: Statistical considerations and applications to environmental data, *J. Geophys. Res.*, *103*, 17,149–17,161.

Wilks, D. S., and K. L. Eggleston (1992), Estimating monthly and seasonal precipitation distributions using the 30- and 90-day outlooks, *J. Clim.*, *5*, 252–259.

Yoden, S., M. Taguchi, and Y. Naito (2002), Numerical studies on time variations of the troposphere-stratosphere coupled system, *J. Meteorol. Soc. Jpn.*, *80*, 811–830.

S. Nishizawa and S. Yoden, Department of Geophysics, Kyoto University, Kyoto, 606-8502, Japan. (seiya@kugi.kyoto-u.ac.jp)