

## Reversible Record Breaking and Variability: Temperature Distributions across the Globe

AMALIA ANDERSON AND ALEXANDER KOSTINSKI

*Michigan Technological University, Houghton, Michigan*

(Manuscript received 5 October 2009, in final form 10 March 2010)

### ABSTRACT

Based on counts of record highs and lows, and employing reversibility in time, an approach to examining natural variability is proposed. The focus is on intrinsic variability; that is, variance separated from the trend in the mean. A variability index  $\alpha$  is suggested and studied for an ensemble of monthly temperature time series around the globe. Deviation of  $\langle \alpha \rangle$  (mean  $\alpha$ ) from zero, for an ensemble of time series, signifies a variance trend in a distribution-independent manner. For 15 635 monthly temperature time series from different geographical locations (Global Historical Climatology Network), each time series about a century-long,  $\langle \alpha \rangle = -1.0$ , indicating decreasing variability. This value is an order of magnitude greater than the  $3\sigma$  value of stationary simulations. Using the conventional best-fit Gaussian temperature distribution, the trend is associated with a change of about  $-0.2^\circ\text{C} (106 \text{ yr})^{-1}$  in the standard deviation of interannual monthly mean temperature distributions (about 10%).

### 1. Introduction

While there is a vast literature on detecting globally averaged interannual mean temperature trends (global warming), the possible trend in the interannual variance has received less attention. Yet, a change in variance may be just as important (e.g., Katz and Brown 1992). According to the IPCC report, “Evidence for changes in observed inter-annual variability (such as standard deviations of seasonal averages) is still sparse” (Trenberth et al. 2007, p. 300) and previous studies have led to varied conclusions (Scherrer et al. 2005).

Many recent studies devoted to interannual variability have addressed regional and seasonal variability, motivated largely by European heat waves of the early twenty-first century. For example, Della-Marta et al. (2007) find a variance trend of  $+6\% \pm 2\%$  for daily summer maximum temperatures in western Europe. While Della-Marta et al. explore distributions of extreme temperatures (tails of distributions), variance trends from entire distributions are also important. Although such attempts (Schär et al. 2004; Scherrer et al. 2005) inspired much study and

received support from climate model simulations (e.g., Vidale et al. 2007), because of data sparsity and the weak signals, results remain statistically inconclusive.

Since detection of a statistically significant global trend in interannual temperature variance remains an important, yet open problem, we propose a different approach. Motivated by greater sensitivity of outstanding values to subtle trends and distribution-independence of the results, we reformulate the problem in the language of record-breaking statistics and explore global trends in variance of average temperatures.

Record-breaking statistics have been used to elucidate a variety of weather and climate related trends (e.g., Glick 1978; Vogel et al. 2001), but their use in the analysis of temperature trends is more recent and has focused on observing trends in the mean (as opposed to variance) of temperature time series (see Basset 1992; Basset and Lin 2003; Benestad 2003, 2004; Redner and Petersen 2006; Meehl et al. 2009). To the best of our knowledge, record-breaking statistics have not been used to extract variance trends, although the notion appears in the mathematical literature, for example, seminal works Foster and Stuart (1954), Cox and Stuart (1955), and Ury (1966), and more recently, Gulati and Padgett (2003), Hofmann and Balakrishnan (2006), and Krug (2007). We are, however, aware via Meehl et al. (2009), that R. W. Portmann et al. (2009, personal

---

*Corresponding author address:* Amalia Anderson, Dept. of Physics, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931-1295.  
E-mail: amlander@mtu.edu

communication), may address the same issue. We now proceed to show that record-breaking statistics appear promising for exploring trends in interannual temperature variance.

## 2. Background for record-breaking statistics

The  $i$ th entry in a time series,  $x_i$ , is a record-breaking event (record) if it exceeds all previous values in the sequence. In other words,  $x_i$  is a record high if

$$x_i > \max(x_1, x_2, \dots, x_{i-1}), \quad (1)$$

and is a record low if

$$x_i < \min(x_1, x_2, \dots, x_{i-1}). \quad (2)$$

The first entry is always a record high and a record low; see Fig. 1.

The expected number of records (successes) in a time series is the sum over trial (term in a series) probabilities of being a record (success). For a set of trials that are independent, identically distributed (i.i.d.), and continuous, the  $n$ th trial has an equal chance of having the greatest value as all preceding trials:  $1/n$ . Thus the expected number of successes (records),  $E(R)$ , for a time series with  $n$  events is

$$E(R) = 1 + 1/2 + 1/3 + \dots + 1/n, \quad (3)$$

and, by Euler's formula for harmonic series,

$$E(R) \approx \ln(n) + \gamma, \quad (4)$$

where  $\gamma = 0.577\dots$  is the Euler constant. These results are occasionally attributed to Rényi (1962) [e.g., Eliazar and Klafter (2009) refers to the "celebrated theorem by Rényi"]. However, the results originate with Foster and Stuart (1954). We stress the distribution independence of these results; that is, they hold for any continuous probability densities. If the i.i.d. assumption is violated by a trend or correlations, the number of records will deviate from the logarithmic dependence in Eq. (4) and trends can, perhaps, be detected in a distribution-independent manner. However, one difficulty remains.

Measured quantities are never continuous as instrumental precision is finite, e.g., GHCN monthly temperatures are reported with a resolution of a tenth of a degree (C). This allows for a possibility of ties, thus ruining the beautifully simple result [Eq. (4)]. [See, e.g., Vervaat (1973), Gouet et al. (2001), and Key (2005) for details regarding records in the discrete case.]

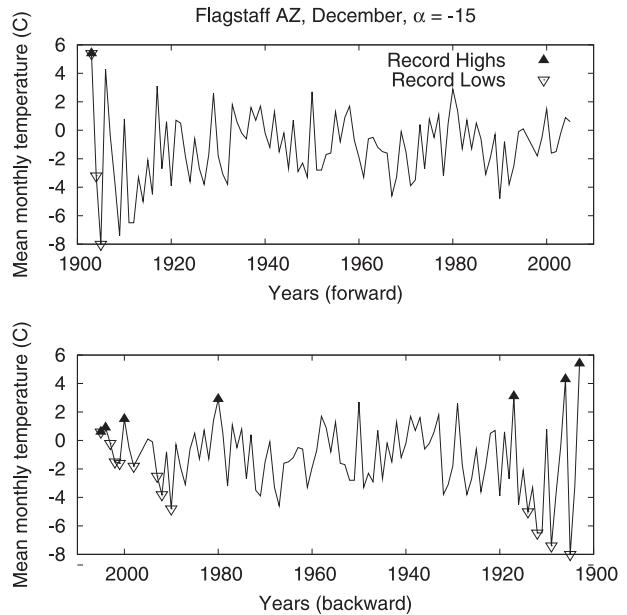


FIG. 1. Example of record counting: monthly means for Flagstaff, AZ (December). The filled (open) triangles show locations of record-breaking highs (lows). The variability index  $\alpha = (RH_{\text{fwd}} - RH_{\text{bwd}}) + (RL_{\text{fwd}} - RL_{\text{bwd}})$ , where RH (RL) is the number of record-breaking highs (lows), and the subscript indicates direction in time. A negative (positive) value of  $\alpha$  corresponds to a decreasing (increasing) variance for the forward time series. For this example,  $\alpha = -15$ , both before and after detrending. This is the lowest value of  $\alpha$  for the GHCN dataset, see section 4. Years (top) forward and (bottom) backward.

## 3. Problem formulation and the variability index

To circumvent the above difficulty, we pursue an approach that is blind to ties, yet sensitive to variance. To that end, note that stationary time series have no trends in mean, variance, etc., and, therefore are invariant with respect to time-reversal (e.g., Foster and Stuart 1954). This invariance implies that the expected number of record highs or lows does not change upon time-reversal, regardless of continuity (possibility of ties) or serial correlation (clusters of extremes may occur, but without preference for direction). Conversely, trends break the time-reversal invariance and deviation from such invariance can possibly be used to detect time-dependence of various parameters, as is illustrated in Fig. 1.

In addition to time-reversal symmetry (or lack thereof), there is a possible symmetry between highs and lows. For example, the global warming signal (trend in the mean annual global temperature) causes excess number of record highs when compared with record lows (see appendix A; Meehl et al. 2009). To disentangle the two symmetries, we subtract the mean trend from our time series (section 4b) and focus on detecting variance trends.

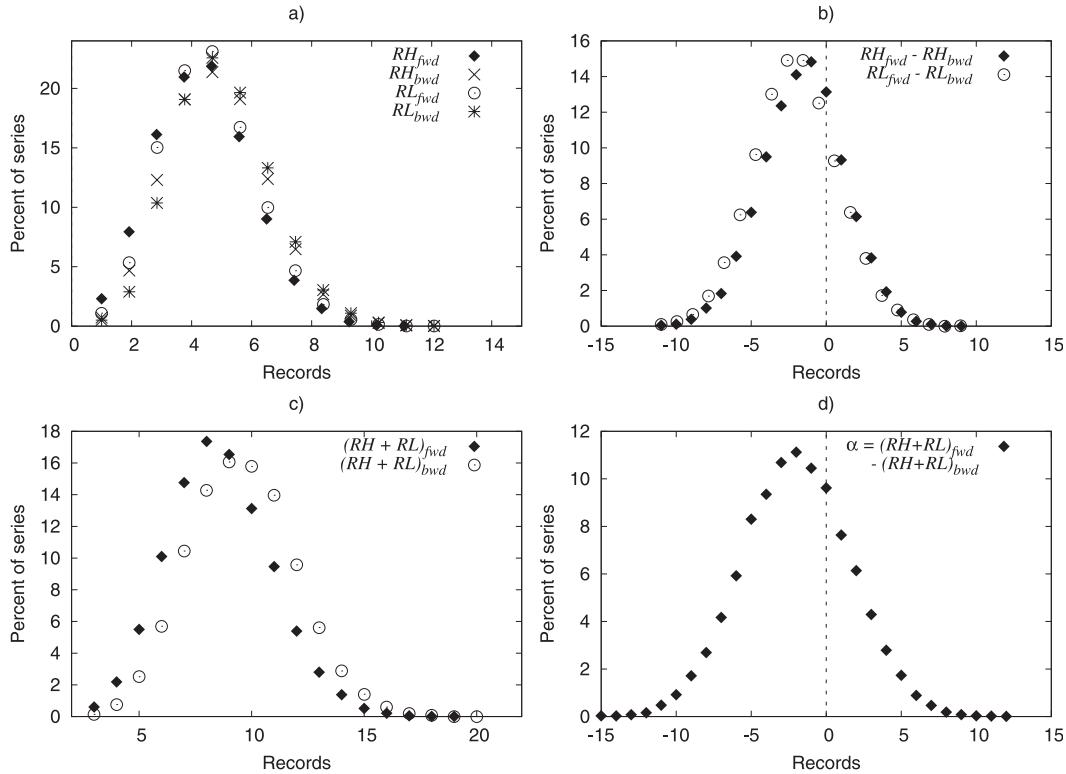


FIG. 2. Histograms of  $\alpha$  components for the 15 635 GHCN time series, each with its mean trend removed. Record highs (lows) are signified by RH (RL) with subscripts denoting direction in time. (a) Each of four components are plotted; (b) the “sum of differences” perspective (section 3) is emphasized,  $RH_{fwd} - RH_{bwd}$  and  $RL_{fwd} - RL_{bwd}$  are plotted. Both histograms are shifted left, indicating a decreasing variance. The sum of these components is  $\alpha$ . (c) The “difference of sums” perspective is emphasized,  $(RH + RL)_{fwd}$  and  $(RH + RL)_{bwd}$  are plotted. The difference of the components of these two histograms is  $\alpha$ . The shift between the histograms hints at a non-0  $\alpha$ . (d) The histogram of  $\alpha$  values for the GHCN data. Our main result:  $\langle \alpha \rangle = -1.0$  is evident here in the shift of the peak from  $\langle \alpha \rangle = 0$ , indicating a decreasing variance.

Given these symmetries, we construct a variance detector  $\alpha$  as follows:

$$\begin{aligned} \alpha &= (RH_{fwd} - RH_{bwd}) + (RL_{fwd} - RL_{bwd}) \\ &= (RH + RL)_{fwd} - (RH + RL)_{bwd}, \end{aligned} \tag{5}$$

where RH and RL are the numbers of record-breaking highs and record-breaking lows, respectively, and the subscripts denote direction in time (forward and backward). The two forms in the above equation differ only by rearrangement but contrast “sum of differences” versus “difference of sums” perspectives, see Fig. 2 in section 5. Note that for stationary time series  $\alpha$  is entirely independent of distribution, continuity, and serial correlation;  $E(\alpha) = \langle \alpha \rangle = 0$ . Here we introduce  $\langle \cdot \rangle$  to denote an average over an ensemble of time series, used in the remainder of this paper. A positive (negative)  $\langle \alpha \rangle$  indicates an increasing (decreasing) variance as demonstrated in Fig. 1. Thus, a qualitative trend can be

extracted without making any assumptions about the data (section 5).

For subsequent development, we note that some symmetries also apply for the nonstationary case where only a mean trend (e.g., monotonic, “hockey stick,” or possessing odd–even symmetry) is present. For example,  $\langle RH_{fwd} \rangle = \langle RL_{bwd} \rangle$ ,  $\langle RH_{bwd} \rangle = \langle RL_{fwd} \rangle$ , and the number of ties is not affected by the time reversal. Hence,  $\langle \alpha \rangle = 0$ , regardless of distribution. Conversely, if the mean trend were removed, these symmetries would remain as the time series would return to being stationary. (These statements are elaborated below, see Fig. 4 in section 5 and Fig. B1 in appendix B).

Correlations between station pairs can reduce the statistical significance of  $\langle \alpha \rangle$  as the effective number of time series is reduced in their presence. In the GHCN data (section 4a), we account for correlations between stations by reducing the number of simulated time series used to evaluate statistical significance of the results. Serial correlations can also affect  $\langle \alpha \rangle$  by altering its

magnitude but for the GHCN data, this effect has been found negligible, as shown in appendix C. Recall that because of the time-reversal symmetry built into the definition of  $\alpha$ , correlated clusters of records tend to cancel. (See the appendix C for details concerning both temporal and spatial correlations.) To quantify a variance trend via  $\alpha$ , a particular distribution must be assumed and our simulations and data analysis to that end are reported in section 4c.

To summarize, the novelty of the proposed approach lies in combining record-breaking statistics with the separation of trends in order to ask about intrinsic variability. Indeed, while it is known that records can indicate trends and, furthermore, time reversibility has already been applied to finding mean trends, Foster and Stuart (1954) and Benestad (2004), we find no studies that use the method to find trends in variability. Perhaps no one asked the question because it is difficult to draw conclusions about a variance trend when entwined with mean trends. Below, we discuss detrending and trend simulation methods and then proceed to extract a weak but significant trend using our metric  $\alpha$ .

#### 4. Data and simulations

##### a. Description of the dataset

The Global Historical Climatology Network dataset (version 2) is maintained by the National Climatic Data Center (NCDC) and has been used widely in studies of global temperature trends. Notably, it has been used by Benestad (2004) in his studies of temperature and record-breaking statistics, by Hansen et al. (2001) of the Goddard Institute of Space Studies (GISS) in climate analysis, and by Trenberth et al. (2007) of the International Panel on Climate Change (IPCC) as a reference dataset. IPCC reports that this data has a linear global mean temperature trend of  $0.64^{\circ}\text{C} \pm 0.16^{\circ}\text{C} (100 \text{ yr})^{-1}$  for 1901–2005 (Trenberth et al. 2007). Additional information about the data can be found in Peterson and Vose (1997) and Peterson et al. (1998).

The GHCN dataset consists of land surface monthly mean temperatures for  $\approx 7000$  stations distributed globally. We use the dataset adjusted for nonclimatic irregularities, for example, processing errors, mislocated stations, changes in location of data collection, and instrument changes, Peterson and Vose (1997). Our analysis was performed on time series between 90 and 130 years in duration. Also, in order to exclude correlated data, we omit time series designated as duplicates. The filtering rejected  $\approx 80\%$  of time series, mostly as too short. Thus,  $\approx 1500$  stations qualified for our analysis, most containing data for all twelve months. We constructed time series for

each month individually so, for example, monthly means for January at a given station comprise one time series, February values present a new time series and 15 635 time series were thus obtained.

##### b. Extracting intrinsic variance

To extract a possible trend in variance, independent of the “global warming” signal, one must remove the mean trend. From the outset, we note that the results reported below are robust; that is, relatively insensitive to the method of detrending. We chose a local linear regression, LOWESS (see Cleveland 1979) as one of the methods. LOWESS detrending uses a smoothing parameter ( $f$ ), the percent of values (nearest neighbors) used to calculate each local regression. Since the GHCN time series length ( $l$ ) is variable, we refer to  $f' \equiv fl$ , the number of nearest neighbors used in calculating each local regression. We chose  $f' = 50$ : given that  $l \approx 100$  yr, this corresponds to  $f \approx 0.5$ , the starting point recommended by Cleveland (1979). For each of the 15 635 GHCN time series we subtract off resulting mean trends (regressions) individually. As detailed in appendix B, other detrending methods, such as linear regression, LOWESS regressions with other  $f'$  values, and various least squares polynomial fits, yield similar variance trends.

The resulting time series, stationary in the mean, may still have variance trends as indicated by the metric  $\alpha$  [Eq. (5)]. Does the detrending method or magnitude of the mean trend affect the variability index  $\langle \alpha \rangle$ ? It is shown in appendix B that  $\langle \alpha \rangle$  is far more sensitive to variance trends than mean trends and it is quite robust with respect to the methods of detrending. Also, a test of mean trend removal in the language of records (deviation of  $\langle \beta \rangle \equiv \langle \text{RH} - \text{RL} \rangle$  from 0) is discussed in appendix A. For the GHCN data, even with no detrending at all, a significant variance trend is detected.

##### c. Simulation of variance trends

The variability metric  $\langle \alpha \rangle$  is calculated using solely the GHCN dataset, without any assumptions about distributions. However, in order to associate a quantitative trend with a numerical value of  $\langle \alpha \rangle$  we employ simulations, based on the commonly used Gaussian fits to temperature distributions (see discussion below for more details). The important notion here is to keep constant the mean while letting the variance of the Gaussian distribution,  $1/\sqrt{2\pi\sigma^2} \exp(-T^2/2\sigma^2)$ , depend on time, for example,  $\sigma = \sigma(t)$ . The simulation is then performed by drawing random numbers from a discrete Gaussian distribution (separation of  $0.1^{\circ}\text{C}$  between

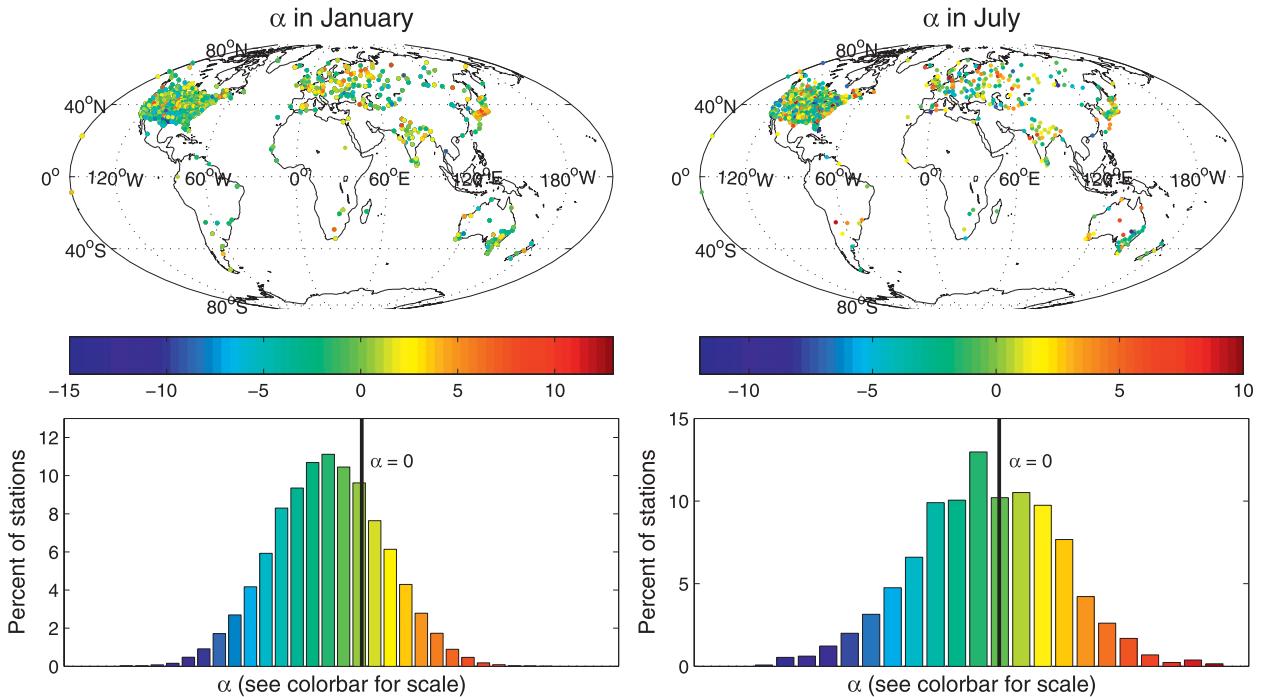


FIG. 3. (top) Spatial distribution of stations in the GHCN dataset and the corresponding  $\alpha$  values (color bar) for (left) January and (right) July (each month provides a set of over a thousand time series). Here the integer-valued  $\alpha$  values are from the LOWESS detrended data. The most popular values of  $\alpha$  are  $-1$  and  $-2$ . (bottom) Histograms for all 12 months are shifted from  $\alpha = 0$  ( $\langle \alpha \rangle = 0$  for no trend) again showing a decreasing trend in variance.

values, the resolution of the GHCN data) with time dependent standard deviation,  $\sigma(t)$ :

$$\sigma(t) = \sigma_0 + \delta t \tag{6}$$

with  $\sigma_0 \equiv$  the initial standard deviation,  $\delta \equiv$  the trend per year, and  $t \equiv$  the time in years. We mimicked the data by using an initial standard deviation of  $1.8^\circ\text{C}$ , the mean standard deviation for the data (per time series), and a length of 106 years, the mean time series length in the data.

The conventional Gaussian choice for the temperature distribution was confirmed via the moment test (e.g., Thode 2002), with 92% of the data within the 99th percentile for the third moment of a Gaussian and 97% is within the 99th percentile for the fourth moment of a Gaussian, each with coinciding means. We also considered truncated Gaussians, as recommended by Redner and Petersen (2006), with the range between boiling points of nitrogen and water, but found that for  $\sigma \approx 2^\circ\text{C}$ , simulated monthly mean values were not affected.

### 5. Results and discussion

In addition to proposing a distribution-independent method, based on reversible record breaking, our

GHCN-based results are also noteworthy. The main result is the ensemble-averaged (15 635 time series)  $\langle \alpha \rangle = -1.0$ , calculated via Eq. (5) for the detrended GHCN dataset. Histograms of  $\alpha$  and its components are shown in Fig. 2, while the spatial distribution of stations and monthly components of  $\langle \alpha \rangle$  (January and July) are illustrated in Fig. 3. The histograms of  $\alpha$  are peaked to the left of  $\langle \alpha \rangle = 0$ , supporting the  $\langle \alpha \rangle = -1.0$  result, and indicating a decreasing variance. This is so for each of the 12 months.

Returning to the result  $\langle \alpha \rangle = -1.0$ , we attempt to quantify its statistical significance and uncertainty, via a Monte Carlo estimate of confidence intervals (e.g., Press et al. 2007). To that end, we turn to i.i.d. simulations. Each single stationary (i.i.d.) simulation is a time series consisting of independent draws, mimicking GHCN conditions. One thousand datasets (each consisting of 10 000 time series) were generated. Rather than using the 15 635 time series as in GHCN, we used a conservative 10 000 stationary simulations to account for possibly correlated time series pairs in the GHCN data (see appendix C). The overwhelming 99.9% of our stationary (i.i.d.) simulations, resulted in  $|\langle \alpha \rangle| < 0.12$ . None approached the GHCN value of  $-1.0$ . Furthermore, insofar as  $\langle \alpha \rangle$  is obtained by summation of (integer valued)  $\alpha$ s, it

is Gaussian-distributed by the central limit theorem. The  $3\sigma$  value or the 99.7% confidence interval for these stationary time series is  $\pm 0.12$ . Thus, the GHCN data value of  $\langle \alpha \rangle = -1.0$  is an order of magnitude larger than the  $3\sigma$  uncertainty.

While  $\langle \alpha \rangle = -1.0$  is statistically significant and free of assumptions, attributing a variance trend to  $\alpha$  depends on a distribution. Simulations based on the conventional, best-fit Gaussian distributions with time-variable variance described in section 4c are shown in Fig. 4. These associate a variance trend, as defined by time-dependent standard deviation  $\sigma = \sigma(t)$ , with a particular value of the variability index  $\alpha$ . For the GHCN value of  $\alpha = -1.0$ , this results in a standard deviation trend of year-to-year monthly mean temperature distributions of about  $-0.2^\circ\text{C} (106 \text{ yr})^{-1}$  [change in variance of  $-(0.04^\circ\text{C})^2 (106 \text{ yr})^{-1}$ ], a 10% relative change. This is for data with most time series between 1900 and 2005, see Fig. C2 in appendix C for durations of time series.

## 6. Concluding remarks

Natural variability is an often used but vaguely defined notion. Linking natural variability with trends in variance is one way to render it precise. For example, Whitcher et al. (2002) review various tests for stationarity of variance and point out that the Gaussian assumption is often employed. Indeed, in climatology assuming a distribution is often unavoidable and yet unrealistic (e.g., Wilcox 2003; Ghil et al. 2002; Gluhovsky and Agee 2007). In contrast, our approach to intrinsic natural variability, based on the index  $\alpha$ , is independent of underlying probability distributions. Also, the index is simple to use and, perhaps, most importantly, our results are insensitive to mean trends (see appendix B for details). Additionally, it is notable that for all reasonable trends, the sign of  $\langle \alpha \rangle$  is independent of mean trends and the built-in reversibility ensures that  $\langle \alpha \rangle = 0$ .

We found  $\langle \alpha \rangle = -1.0$  for the 15 635 detrended GHCN time series, indicating a decreasing variance. This result is statistically significant:  $\langle \alpha \rangle = -1.0$  is an order of magnitude larger than the  $3\sigma$  uncertainty. To attribute a measure to this trend, we turned (alas) to conventional, best-fit Gaussian distributions with time-variable variance. Interpolation results in a standard deviation trend of year-to-year monthly mean temperature distributions of about  $-0.2^\circ\text{C} (106 \text{ yr})^{-1}$ , a 10% relative change. This is for data with most time series between 1900 and 2005, see Fig. C2 in appendix C for durations of time series.

Decreasing variance of interannual (year-to-year) monthly temperatures may seem surprising in light of research done by Della-Marta et al. (2007), Schär et al. (2004), and Scherrer et al. (2005), who found increasing

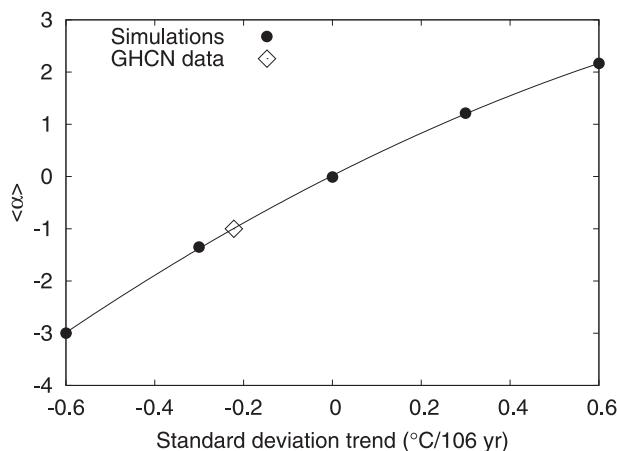


FIG. 4. Plot of  $\langle \alpha \rangle$  vs trend in variance. The ensemble mean value  $\langle \alpha \rangle = -1.0$  indicates decreasing variance. This value is statistically significant as it exceeds, by an order of magnitude, the  $3\sigma$  value for stationary simulations. To attribute a variance trend to  $\langle \alpha \rangle = -1.0$ , results of simulations mimicking the GHCN data are shown. Simple interpolation yields a trend of about  $-0.2^\circ\text{C} (106 \text{ yr})^{-1}$ , a 10% relative change (open diamond). Each filled circle represents the mean of 10 000 time series, see appendix C for details. Temperature time series were simulated by drawing numbers from a discrete Gaussian distribution (separation of  $0.1^\circ\text{C}$  between values, the resolution of the GHCN dataset) with a time-dependent standard deviation (variance trend). The abscissa (trends) is in degrees Celsius per 106 years with an initial standard deviation of  $1.8^\circ\text{C}$ , chosen to equal the mean time series length and mean standard deviation for the data. Error bars are within the markers.

variability for European summers. However, it appears less surprising within the context of intra-annual (values within a single year) variance trends, which are more widely distributed spatially and seasonally. In the past two decades a decrease in intra-annual temperature variance has been observed via diurnal and seasonal changes and it is thoroughly documented (e.g., Karl et al. 1993; Easterling et al. 1997; Michaels et al. 1998; Vinnikov et al. 2002; Klein Tank and Können 2003; Vose et al. 2005; Alexander et al. 2006; Trenberth et al. 2007; Sen Roy and Yuan 2009). These studies also suggest asymmetry between low and high temperature rates, for example, various lows warming faster than highs, and possibly relate to an interannual variance decrease, reported here. Hence, “global calming” may be related to global dimming.

Despite the statistical significance of the results, there is much to explore. The spatial distribution of the GHCN stations is not uniform as can be seen in Fig. 3. Specifically, the longer time series used here are found primarily in the United States, Europe, India, Japan, and Southern Australia. Furthermore, this is a land-only dataset, resulting in the further underrepresentation of the Southern Hemisphere. Hence, the reported decrease in variance may be more localized than the global

extent of the data suggests. However, the question of intrinsic variance and the method based on the distribution-independent variability index  $\alpha$ , along with a first feasibility study are important and should be tried on datasets of better precision and more uniform spatial distribution.

*Acknowledgments.* This work was supported by NSF Grant ATMO5-5467 and the Michigan Space Grant Consortium. We thank Katie Schalk and Jeffrey Pierce for useful discussions and three anonymous reviewers for helpful suggestions.

## APPENDIX A

### Record Highs and Lows versus the Mean Trend

#### a. Mean trend detection via high–low asymmetry

Mean trends break the high–low record symmetry, for example, one expects the global warming trend to cause excess number of record highs. This is, indeed, the case as has recently been reported in Meehl et al. (2009). Meehl et al. examined the mean trend via a ratio of records using NCDC U.S. Cooperative Observer Program network station data:  $\approx 2000$  stations spanning the United States and 56 years of data (1950–2006). Rather than examine globally distributed records in time series of monthly mean temperatures, Meehl et al. consider U.S. time series of both daily maxima and minima [extremes are also good indicators of change, see Della-Marta et al. (2007) who also examine daily temperatures]. They report the ratio of record daily maxima to record daily minima that occur in 2006 to be  $\approx 2$ . Similarly, we can report the ratio of total record highs to lows using monthly means in the GHCN data:  $\sum \text{RH} / \sum \text{RL} = 1.14$ , where the sums are over all 15 635 time series and RH (RL) are the number of record highs (lows) per time series. However, note that the number of expected record events increases with the length of time series. When a mean ratio is used instead of the sums, record highs are compared to record lows for time series of the same length. Our time series vary in length (90 to 130 years) and the mean ratio is better suited for evaluating trends:  $\langle \text{RH} / \text{RL} \rangle = 1.46$ .

#### b. Detrending tests via high–low asymmetry

As above, we employ high–low asymmetry to check our detrending. For this purpose, we define  $\beta \equiv \text{RH} - \text{RL}$ . For strictly stationary time series, there is a complete symmetry between highs and lows so that  $\langle \beta \rangle = 0$  (either forward or backward). Indeed,  $\langle \beta \rangle = 0$  indicates absence of (nonperiodic) mean trend, although the

condition is sufficient but not necessary. For example, a low-clipped distribution would favor record highs, but could still be stationary. We calculate  $\langle \beta \rangle$  to demonstrate the diminishing high–low asymmetry for the detrended GHCN data. For the LOWESS detrended data used in this paper ( $f' = 50$ ),  $\langle \beta \rangle = -0.2$  (forward and backward), where the angular brackets denote an ensemble average (over the 15 635 time series). This is a decrease in  $|\langle \beta \rangle|$  of 0.45 forward and 1.21 backward, indicating the removal of mean trend; for the GHCN data before detrending  $\langle \beta \rangle_{\text{fwd}} = 0.65$  and  $\langle \beta \rangle_{\text{bwd}} = -1.41$ . We also give results for  $\langle \beta \rangle$  for multiple detrending methods—see Table A1 and appendix B. In all cases  $|\langle \beta \rangle|$  decreases after detrending, again confirming the removal of mean trend.

| Method     | $f'/\text{order}$ | $\langle \alpha \rangle$ | $\langle \beta \rangle_{\text{fwd}}$ | $\langle \beta \rangle_{\text{bwd}}$ |
|------------|-------------------|--------------------------|--------------------------------------|--------------------------------------|
| None       | —                 | −0.84                    | 0.65                                 | −1.41                                |
| Lowess     | 10                | −0.92                    | −0.07                                | 0.00                                 |
| Lowess     | 20                | −0.96                    | −0.10                                | −0.02                                |
| Lowess     | 30                | −0.99                    | −0.14                                | −0.08                                |
| Lowess     | 40                | −1.00                    | −0.18                                | −0.12                                |
| Lowess     | 50                | −1.00                    | −0.22                                | −0.17                                |
| Lowess     | 60                | −1.00                    | −0.23                                | −0.27                                |
| Lowess     | 70                | −0.98                    | −0.22                                | −0.38                                |
| Lowess     | 80                | −0.97                    | −0.17                                | −0.48                                |
| Polynomial | 1                 | −0.93                    | −0.12                                | −0.58                                |
| Polynomial | 2                 | −0.95                    | −0.07                                | −0.51                                |
| Polynomial | 3                 | −1.00                    | −0.46                                | −0.09                                |
| Polynomial | 4                 | −0.94                    | −0.16                                | −0.18                                |
| Anomaly    | —                 | −0.91                    | −0.04                                | 0.40                                 |

condition is sufficient but not necessary. For example, a low-clipped distribution would favor record highs, but could still be stationary. We calculate  $\langle \beta \rangle$  to demonstrate the diminishing high–low asymmetry for the detrended GHCN data. For the LOWESS detrended data used in this paper ( $f' = 50$ ),  $\langle \beta \rangle = -0.2$  (forward and backward), where the angular brackets denote an ensemble average (over the 15 635 time series). This is a decrease in  $|\langle \beta \rangle|$  of 0.45 forward and 1.21 backward, indicating the removal of mean trend; for the GHCN data before detrending  $\langle \beta \rangle_{\text{fwd}} = 0.65$  and  $\langle \beta \rangle_{\text{bwd}} = -1.41$ . We also give results for  $\langle \beta \rangle$  for multiple detrending methods—see Table A1 and appendix B. In all cases  $|\langle \beta \rangle|$  decreases after detrending, again confirming the removal of mean trend.

## APPENDIX B

### Sensitivity of $\alpha$ to Detrending

The purpose of this appendix is to demonstrate that 1) the variability index  $\alpha$  is insensitive to method of detrending (see Table A1) and 2)  $\alpha$  is far more sensitive to changes in variance than changes in the mean. Both of these conclusions tend to render the specifics of detrending moot: we elaborate next.

### a. Sensitivity to methods of detrending

We subtracted off the mean trend fit for each time series and chose two different techniques to calculate the mean trends: 1) local linear regression, LOWESS (recommended by one of the anonymous reviewers); 2) least squares polynomial fit to the time series. Between the two techniques we tested 12 distinct detrending versions: eight LOWESS tests for different smoothing parameters (section 4b) and four least squares regressions for different order polynomials. All 12 mean trend removal calculations yielded values an order of magnitude larger than the  $3\sigma$  uncertainty,  $\pm 0.12$ :  $\langle\alpha\rangle$  ranges between  $-0.92$  and  $-1.00$  (see Table A1). In fact, even without detrending, the GHCN dataset still gives statistically significant results:  $\langle\alpha\rangle = -0.84$ .

In addition, we tried detrending by subtracting off the global temperature anomaly for that year. The temperature anomaly is commonly defined as a deviation of a globally averaged temperature for a given year from the globally averaged temperature of the entire time span (1880–2006): see the NCDC Internet site (online at <http://www.ncdc.noaa.gov/oa/climate/ghcn-monthly/index.php>) for the relevant deviations. Thus processed time series from this time period resulted in  $\langle\alpha\rangle = -0.91$ .

### b. Sensitivity to magnitude of the trends

In Fig. B1, we compare sensitivities of  $\alpha$  to linear trends in mean and variance (see section 4c for details on trend simulation). When the mean and variance trends are comparable, the mean trend has a negligible effect on  $\langle\alpha\rangle$ , compared to the effect of the variance trend. This is because there is no direct measure of mean trend in  $\alpha$ ; that is, the difference or ratio of record highs and lows in the same temporal direction. Thus, while detrending of the mean trend will often be imperfect,  $\langle\alpha\rangle$  will still give a reliable measure of variance trend as shown in Table A1.

## APPENDIX C

### Spatial and Temporal Correlations

#### a. Spatial correlations

In addition to GHCN filtering, we tested for “geographical independence” of the GHCN time series by calculating correlation coefficients ( $\rho$ ) for all station pairs, calculating each month separately. For the GHCN dataset with no detrending, 45% of series pairs are uncorrelated ( $\rho$  between  $\pm 0.10$ ,  $1\sigma$  for uncorrelated time series) and 83% are uncorrelated or weakly correlated

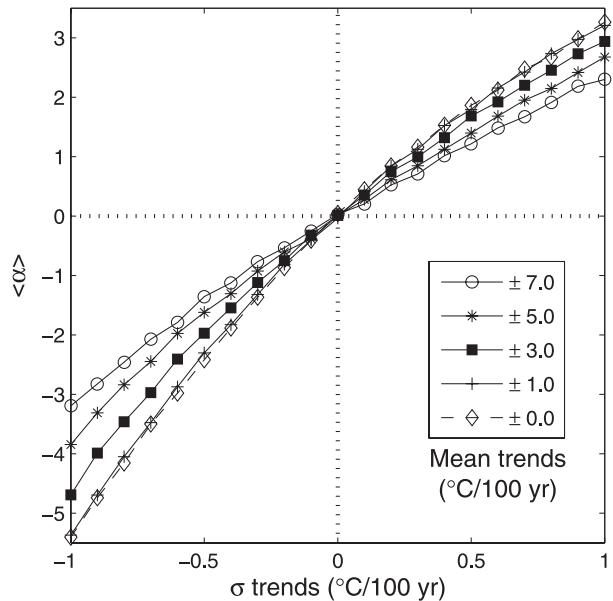


FIG. B1. Variability index  $\alpha$  is relatively insensitive to mean trends. Plots of  $\langle\alpha\rangle$  vs the linear trend in standard deviation ( $\sigma$ ) for several values of a parameter (linear mean trend). Realistic mean trends,  $<|\pm 1^\circ\text{C} (100 \text{ yr})^{-1}|$  (comparable to GHCN data), do not affect  $\alpha$  but larger mean trends do. In particular, the figure shows that only variance changes cause  $\langle\alpha\rangle \neq 0$ ; At a fixed variance trend,  $|\langle\alpha\rangle|$  decreases with increasing mean trend, rendering the  $\alpha$ -estimated variance trend conservative; and mean trends can change the magnitude of  $\langle\alpha\rangle$ , but not its sign. Each point is the mean for 10 000 simulations (see section 4c for details regarding the simulations).

( $\rho < 0.30$ ), see Fig. C1. Additionally, 11% of correlated time series ( $\rho > 0.30$ ) are not in the same country (widely separated), so only  $\approx 15\%$  of time series pairs are possibly significantly correlated. Correlation coefficients for the LOWESS detrended data ( $f' = 50$ ) are similar: 46% of series are uncorrelated and 86% are uncorrelated or weakly correlated (see Fig. C1). To account for these possible correlations, we conservatively reduced the number of simulated time series to 10 000,  $\approx 64\%$  of the GHCN number of time series.

Note that while intercorrelated time series could lead to clustering of records, the clustering effects tend to cancel in the calculation of  $\alpha$  because of the time-reversal symmetry. Furthermore, we show that clustering of records is minimal in the GHCN data, by examining the placement of records in time. Figure C2 shows the fraction of time series that have a record each year; record highs and lows, forward and backward. The vertical axis on the right shows the number of time series available each year. Note that deviations from apparent noise in the annual fraction of records often coincides with changes in the number of time series, most notably

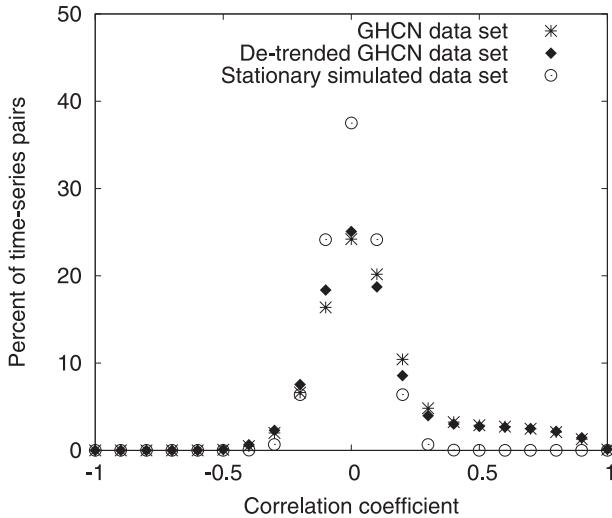


FIG. C1. Statistical independence of geographically separated time series is shown here by comparing histograms of correlation coefficients for pairs from the 15 635 GHCN time series, the LOWESS detrended GHCN data ( $f' = 50$ ), and a simulated dataset of stationary and independent time series. The overlap of the histograms indicates low frequency of correlations in the GHCN data; 85% of GHCN time series pairs and 86% of the LOWESS detrended GHCN time series pairs are uncorrelated or weakly correlated ( $\rho < 0.30$ ). However, with 15% possibly correlated we conservatively reduced the number of simulated time series to 10 000,  $\approx 64\%$  of the GHCN number of time series.

around 1990 where there is a large spike in records counted backward, but also a 20% increase in the number of time series.

*b. Temporal correlations*

GHCN time series (no detrending) are slightly skewed toward positive serial correlations (see Fig. C3a; 52% of time series have no significant correlation, between  $\pm 0.10$  for 1-yr lag,  $1\sigma$  for “noise” in autocorrelation coefficients beyond 2-yr lags, and 93% have no serial correlation or weak serial correlation, between  $\pm 0.30$  for 1-yr lag. Meanwhile, serial correlations for the LOWESS detrended GHCN time series ( $f' = 50$ ) have the same distribution as those for stationary simulations: 64% of time series have none, between  $\pm 0.10$  for 1-yr lag,  $1\sigma$  for “noise” in autocorrelation coefficients beyond 2-yr lags, and 99% have no serial correlation or weak serial correlation, between  $\pm 0.30$  for 1-yr lag. This demonstrates that the excess of positively correlated series is associated with the mean trend as the detrending leaves no appreciable serial correlations.

We verify that serial correlations are negligible by calculating  $\langle \alpha \rangle$  for two simulated datasets with identical variance trends: one with serial correlations (average autocorrelation at 1-year lag  $\approx 0.9$ ), the other with

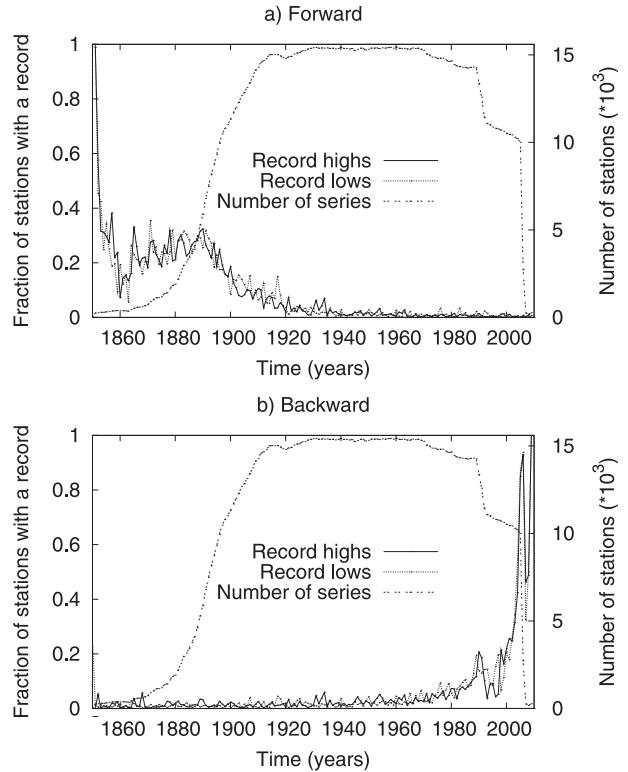


FIG. C2. For the GHCN data, we show here the annual fraction of series with a record high or low, (a) forward and (b) backward. Record highs and lows in the forward (backward) direction decrease (increase) toward 2000. The number of series that are available each year are shown also (right y axis). Notably, there do not seem to be any outstanding years with record clustering; the largest fluctuations in record numbers correspond to increases or decreases in the number of time series, recall the first value is always a record.

no serial correlations. The results were statistically indistinguishable, see Fig. C3b. Because serial correlations in the GHCN data are weaker and less frequent than those in the simulated time series with serial correlations, the weak serial correlations in the data do not affect the value of  $\alpha$  nor the attribution of the trend.

To conclude, in the absence of detrending we find significant serial correlation, but note that distinguishing between strictly stationary correlated time series and uncorrelated time series with a trend, is dubious. Indeed, the very notion of serial correlation is valid only for strictly stationary processes. In fact, a trend in variance itself can induce spurious serial correlations. We are not alone in taking this view; for example, Meehl et al. (2009) do not even consider study of autocorrelation worthwhile in their record high/low study: “Though this simple summation does not take into account station record length or any autocorrelation effects, it is not unexpected that there would be more record high

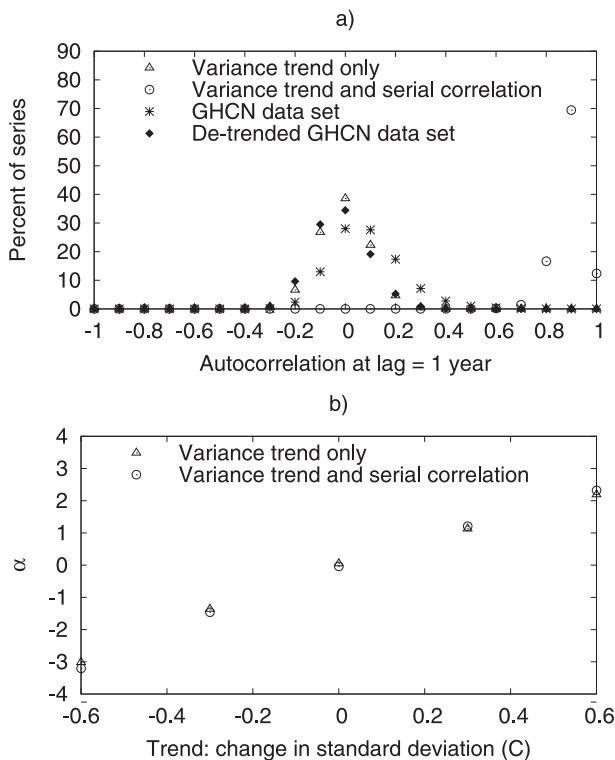


FIG. C3. Simulations of time series with variance trends and serial correlations that are stronger and more frequent than those in the GHCN time series are shown here to be indistinguishable in  $\alpha$  from the same simulations without serial correlations. We conclude, that based on time-reversibility, serial correlations in the GHCN time series do not significantly affect the value of  $\alpha$  or the attribution of the trend. (a) A histogram of autocorrelation coefficients at lag of one year for the GHCN time series before and after LOWESS detrending ( $f' = 50$ ) and the two simulated datasets with variance trends: one with serial correlations (average autocorrelation at 1-year lag  $\approx 0.9$ ) and one without. Serial correlation is notably reduced for the detrended data; (b)  $\langle \alpha \rangle$  for the two simulated datasets with variance trends, with and without serial correlations. The results are statistically indistinguishable (error bars are within the markers).

maximum temperatures being set than record low minima simply because the annual U.S. average surface temperatures have been increasing since the 1970s.”

#### REFERENCES

Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.*, **111**, D05109, doi:10.1029/2005JD006290.

Basset, G. W., 1992: Breaking recent global temperature records. *Climatic Change*, **21**, 303–315.

—, and Z. Lin, 2003: Breaking global temperature records after Mt. Pinatubo. *Climatic Change*, **25**, 179–184.

Benestad, R. E., 2003: How often can we expect a record event? *Climate Res.*, **25**, 1–13.

—, 2004: Record-values, non-stationarity tests and extreme value distributions. *Global Planet. Change*, **44**, 11–26.

Cleveland, W. S., 1979: Robust locally weighted regression and smoothing scatterplots. *J. Amer. Stat. Assoc.*, **74**, 829–836.

Cox, D. R., and A. Stuart, 1955: Some quick sign tests for trend in location and dispersion. *Biometrika*, **42** (1/2), 80–95.

Della-Marta, P. M., M. R. Haylock, J. Luterbacher, and H. Wanner, 2007: Doubled length of Western European summer heat waves since 1880. *J. Geophys. Res.*, **112**, D15103, doi:10.1029/2007JD008510.

Easterling, D. R., and Coauthors, 1997: Maximum and minimum temperature trends for the globe. *Science*, **277**, 364–367.

Eliazar, I., and J. Klafter, 2009: Record events in growing populations: Universality, correlation, and aging. *Phys. Rev. E*, **80**, 061117, doi:10.1103/PhysRevE.80.061117.

Foster, F. G., and A. Stuart, 1954: Distribution-free tests in time-series based on the breaking of records. *J. Roy. Stat. Soc., Ser. B*, **16** (1), 1–22.

Ghil, M., and Coauthors, 2002: Advanced spectral methods for climatic time series. *Rev. Geophys.*, **40**, 1003, doi:10.1029/2000RG000092.

Glick, N., 1978: Breaking records and breaking boards. *Amer. Math. Mon.*, **85** (1), 2–26.

Gluhovskiy, A., and E. Agee, 2007: On the analysis of atmospheric and climatic time series. *J. Appl. Meteor. Climatol.*, **46**, 1125–1129.

Gouet, R., F. J. Lopez, and M. San Miguel, 2001: A martingale approach to strong convergence of the number of records. *Adv. Appl. Probab.*, **33**, 864–873.

Gulati, S., and W. J. Padgett, 2003: *Parametric and Nonparametric Inference from Record-Breaking Data*. Springer, 113 pp.

Hansen, J. E., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and T. Karl, 2001: A closer look at United States and global surface temperature change. *J. Geophys. Res.*, **106D**, 23 947–23 963.

Hofmann, G., and N. Balakrishnan, 2006: A nonparametric test for trend based on initial ranks. *J. Stat. Comput. Sim.*, **76**, 829–837.

Karl, T. R., and Coauthors, 1993: A new perspective on recent global warming: Asymmetric trends of daily maximum and minimum temperature. *Bull. Amer. Meteor. Soc.*, **74**, 1007–1023.

Katz, R. W., and B. G. Brown, 1992: Extreme events in a changing climate: Variability is more important than averages. *Climatic Change*, **21**, 289–302.

Key, E., 2005: On the number of records in an iid discrete sequence. *J. Theor. Probab.*, **18**, doi:10.1007/s10959-004-2578-2.

Klein Tank, A. M. G., and G. P. Können, 2003: Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *J. Climate*, **16**, 3665–3680.

Krug, J., 2007: Records in a changing world. *J. Stat. Mech. Theor. E*, **7**, P07001, doi:10.1088/1742-5468/2007/07/P07001.

Meehl, G. A., C. Tebaldi, G. Walton, D. Easterling, and L. McDaniel, 2009: Relative increase of record high maximum temperatures compared to record low minimum temperatures in the U.S. *Geophys. Res. Lett.*, **36**, L23701, doi:10.1029/2009GL040736.

Michaels, P. J., R. C. Balling Jr., R. S. Vose, and P. C. Knappenberger, 1998: Analysis of trends in the variability of daily and monthly historical temperature measurements. *Climate Res.*, **10**, 27–33.

Peterson, T. C., and R. S. Vose, 1997: An overview of the global historical climatology network temperature database. *Bull. Amer. Meteor. Soc.*, **78**, 2837–2849.

—, —, R. Schmoyer, and V. Razuvaev, 1998: Global historical climatology network (GHCN) quality control of monthly temperature data. *Int. J. Climatol.*, **18**, 1169–1179.

- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 2007: *Numerical Recipes: The Art of Scientific Computing*. 3rd ed. Cambridge University Press, 1235 pp.
- Redner, S., and M. R. Petersen, 2006: Role of global warming on the statistics of record-breaking temperatures. *Phys. Rev. E*, **74**, 061114, doi:10.1103/PhysRevE.74.061114.
- Rényi, A., 1962: On the extreme elements of observations. *Selected Papers of Alfréd Rényi*, P. Turán, Ed., Akadémia Kiadó, Vol. 3, 50–66.
- Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. *Nature*, **427**, 332–336.
- Scherrer, S. C., C. Appenzeller, M. A. Liniger, and C. Schär, 2005: European temperature distribution changes in observations and climate change scenarios. *Geophys. Res. Lett.*, **32**, L19705, doi:10.1029/2005GL024108.
- Sen Roy, S., and F. Yuan, 2009: Trend in extreme temperature in relation to urbanization in the Twin Cities metropolitan area, Minnesota. *J. Appl. Meteor. Climatol.*, **48**, 669–679.
- Thode, H. C., Jr., 2002: *Testing for Normality*. CRC Press, 479 pp.
- Trenberth, K. E., and Coauthors, 2007: Observations: Surface and atmospheric climate change. *Climate Change 2007: The Physical Science Basis*. S. Solomon et al., Eds., Cambridge University Press, 241–301.
- Ury, H. K., 1966: Large sample sign tests for trend in dispersion. *Biometrika*, **53** (1/2), 289–291.
- Vervaat, W., 1973: Limit theorems for records from discrete distributions. *Stoch. Proc. Appl.*, **1** (4), 317–334.
- Vidale, P. L., D. Lüthi, R. Wegmann, and C. Schär, 2007: European summer climate variability in a heterogeneous multi-model ensemble. *Climatic Change*, **81**, 209–232.
- Vinnikov, K. Y., A. Robock, and A. Basist, 2002: Diurnal and seasonal cycles of trends of surface air temperature. *J. Geophys. Res.*, **107**, 4641, doi:10.1029/2001JD002007.
- Vogel, R. M., A. Zafirakou-Koulouris, and N. C. Matalas, 2001: Frequency of record-breaking floods in the United States. *Water Resour. Res.*, **37**, 1723–1731.
- Vose, R. S., D. R. Easterling, and B. Gleason, 2005: Maximum and minimum temperature trends for the globe: An update through 2004. *Geophys. Res. Lett.*, **32**, L23822, doi:10.1029/2005GL024379.
- Whitcher, B., S. D. Byers, P. Guttorp, and D. B. Percival, 2002: Testing for homogeneity of variance in time series: Long memory, wavelets, and the Nile River. *Water Resour. Res.*, **38**, doi:10.1029/2001WR000509.
- Wilcox, R. R., 2003: *Applying Contemporary Statistical Techniques*. Academic Press, 608 pp.