Climatic Change
DOI 10.1007/s10584-011-0033-9

The twentieth century contiguous US temperature changes indicated by daily data and higher statistical moments

Samuel S. P. Shen · Ai B. Gurung · Hee-S. Oh · T. Shu · David R. Easterling

Received: 3 June 2009 / Accepted: 14 January 2011

Abstract The daily surface air temperature data are used to assess the climate changes of the contiguous United States during the period of 1901 to 2000. The assessment is made through the first four statistical moments of the daily maximum, minimum, and mean temperature anomalies, the linear trends of the moments, and the changes of the anomalies' probability density functions. The results on the first moment, i.e., the mean, are compared with the existing ones in terms of intra-annual means and their linear trends. Our first moment results agree with known ones and demonstrate a decrease from the 1930s to the 1960s and an increase from the 1970s to 2000. The temperature fluctuation is the smallest in the 1960s among the decades from 1931 to 2000. The trends of the higher (second-, third- and fourth-order) moments of the mean, maximum, and minimum surface air temperatures are calculated for the periods 1901–2000, 1910–1945, 1946–1975, and 1976–2000. The results show a

S. S. P. Shen (✉)
Department of Mathematics and Statistics, San Diego State University,
San Diego, CA 92182, USA
e-mail: shen@math.sdsu.edu

A. B. Gurung · T. Shu
Department of Mathematical and Statistical Sciences, University of Alberta,
Edmonton, AB T6G 2G1, Canada

A. B. Gurung
e-mail: gurungai@yahoo.co.uk

T. Shu
e-mail: tingting-shu@hotmail.com

H.-S. Oh
Department of Statistics, Seoul National University, Seoul, 151-747, Korea
e-mail: heeseok.oh@gmail.com

D. R. Easterling
NOAA/National Climatic Data Center, Asheville, NC 28801-5001, USA
e-mail: David.Easterling@noaa.gov

Published online: 08 March 2011

decreasing trend of the second- and third-order moments of all the temperatures. The fourth-order moments of the mean and maximum surface air temperatures have increasing trends, but that of the minimum surface air temperature has a decreasing trend. The seasonal histograms of the mean, maximum, and minimum surface air temperatures are calculated for the three periods 1910–1945, 1946–1975, and 1976–2000 for the stations which have the largest trend of maximum daily surface air temperature. An obvious change has been identified in the probability density functions. Among the changes of statistical parameters, the ones for the minimum temperature are larger than those for the maximum and mean temperatures.

1 Introduction

Understanding climate change has called on the attention to changes in climate variability and extremes (Karl et al. 1993; Katz and Brown 1992; Meehl et al. 2000; IPCC 2007). Very often, users of weather resources wonder if the climate is getting more or less variable, and news reports may have made many people think that the present climate is more variable, violent, and unstable than that of the earlier part of the last century. However, this perception is subject to careful scrutiny, and it is true for certain climate parameters over a certain region over the Earth, such as the intensity of cyclones over the ocean (Knutson et al. 2010), but it can be false for other parameters over other regions based on our current study and other published literatures. Our study implies that the present surface air temperature (SAT) over the contiguous United States is actually less variable than that of a century ago. The higher statistical moments, such as variance, skewness, and kurtosis, affect the shape of the probability density function (PDF), and the size of the PDF tail is a measure of frequency of climate extremes. The information beyond the mean is thus required in the studies of climate change, particularly the changes of climate extremes. To identify the climate changes beyond the mean and to find whether weather is getting more or less variable are related to the trends in the higher statistical moments and the changes of a climatic variable's PDF.

Vinnikov and Robock (2002) investigated the higher statistical moments for the time series data of several climate indices from 1901 to 2000. No trends of the first four moments were found in the New York sea level, US annual mean Modified Palmer Drought Severity Index, All-India Monsoon Rainfall Index, and Southern Oscillation Index. Parker et al. (1994) compared the variances from the seasonal global SAT anomalies for two 20-year periods (1954–1973 and 1974–1993) for each calendar season separately and found a slight increase in the variance in the later period. Karl et al. (1995) analyzed the temperature and precipitation data of different countries and concluded that "although the notion of a recent increase in inter-annual temperature variability is supported by data from the past few decades, the longer data records indicate that this trend is an aberration." Balling (1997) found that the spatial variance of the global SAT had declined in general. On each hemisphere, the variance was negatively correlated to the mean. Michaels et al. (1998) computed the intra-monthly variance of the global daily maximum and minimum SATs for each January and July. The results showed that the intra-monthly variance of both daily minimum and maximum temperatures had decreasing trends. IPCC (2001, 2007) showed some evidences that the variability of the intra-annual temperature had actually decreased. Gong and Ho (2004) computed the intra-seasonal variance and skewness of the SAT data from 155 stations in China and Korea in the period 1954 to 2001. They found decreasing trends in both the variability and the skewness, but very few trends were statistically significant. Each of the above studies used monthly data. However, to understand the changes of climate extremes, the daily data are highly valuable, and they are used in this paper to assess the temporal changes of the higher moments and the PDFs of the daily mean, maximum, and minimum SAT in the last century for the contiguous US. Our results support a decreasing trend of the SAT variability over the contiguous US.

The rest of the paper is arranged as follows. Section 2 describes the GDCN data. Section 3 describes the method of higher statistical moments. Section 4 contains results and their interpretations, and Section 5 includes conclusions and discussion.

2 Data

We used the US National Climatic Data Center's Global Daily Climatology Network (GDCN) v1.0 (Gleason 2002). This global data set contains the daily maximum and minimum temperatures and daily total precipitation data of 32,857 stations around the world from March 1, 1840 to November 30, 2001. Of course, not all the stations have the complete coverage for the entire period. All the GDCN data (both metadata and data) have been processed through an extensive set of quality-control procedures, or simple datum checks and statistical analysis of observations, to locate and identify potential outliers and errors (Gleason 2002), but these data are not homogenized. The contiguous US contains 8,656 stations of daily maximum SAT and 8,651 stations of daily minimum SAT. The data in the period from January 01, 1901 to December 31, 2000 are used in this study. The daily history of the number of the GDCN stations within the contiguous US is shown in Fig. 1. Although the

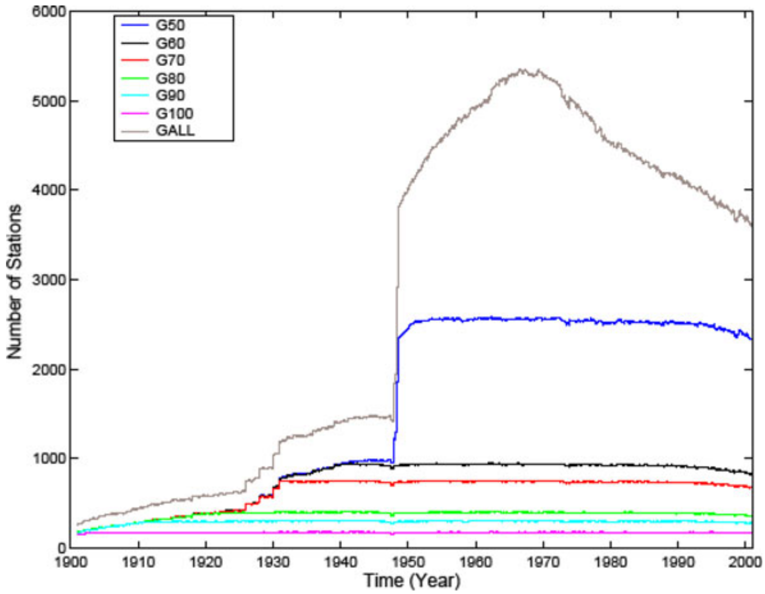

Fig. 1 Number of stations with respect to time for the different record lengths in the GDCN dataset

contiguous US had a total of 8,620 stations in the last century, for a given day the maximum number of stations reporting valid data is less than 5,500, occurred around the 1960s. Most stations started their operations after World War II.

To consider the spatial coverage of the stations with different length of records, we extracted six sub-networks from the GDCN stations based on the minimum record length of 50, 60, 70, 80, 90, and 100 years, in which a maximum of 10% of missing records are allowed. These networks are denoted by G50, G60, G70, G80, G90, and G100, respectively. The seventh subset, denoted by GALL, includes all the US GDCN stations' data irrespective of the minimum record length and the percentage of missing records. The locations of the GALL, G50 and G100 are shown in Fig. 2. Apparently only very few stations have 100 years of continuous records.

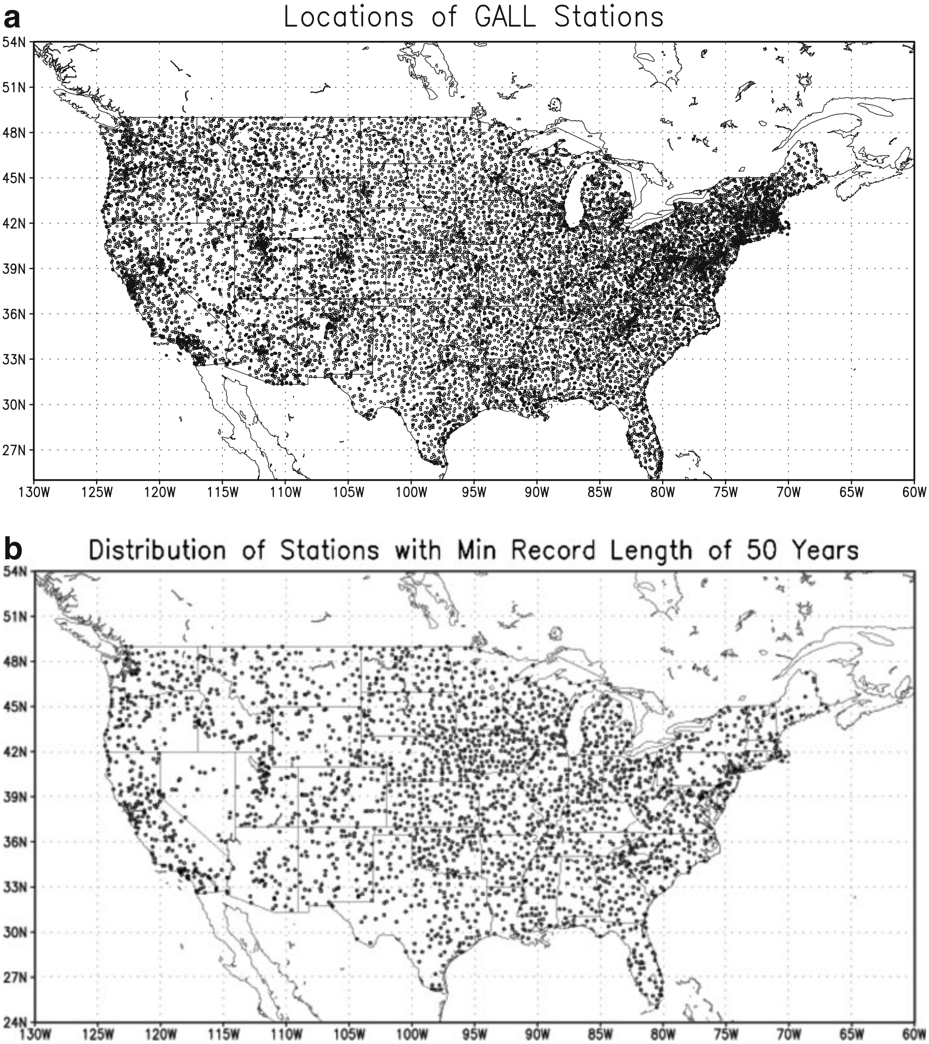

Fig. 2 **a** Locations of GALL stations, **b** locations of G50 stations, and **c** locations of G100 stations

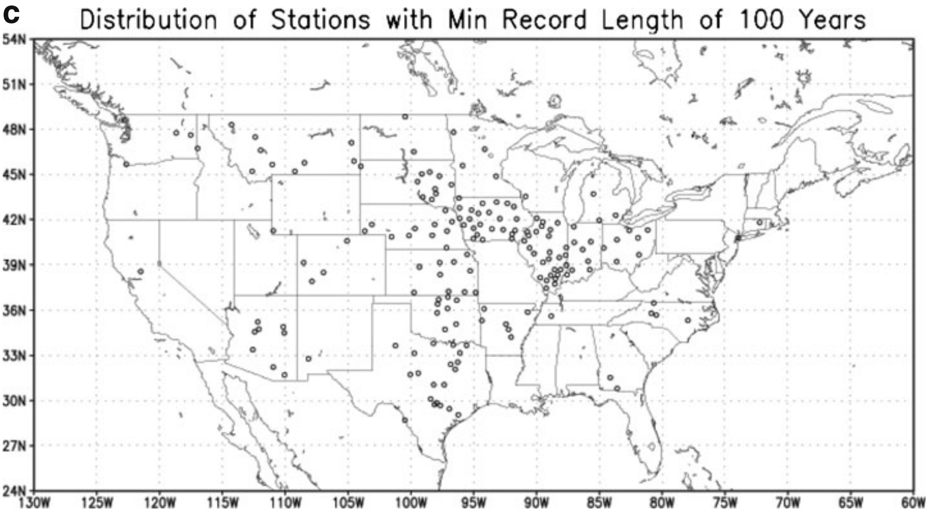

Fig. 2 (continued)

The GDCN source data were mostly collected by the NCDC scientists through cooperative research exchanges. Many stations were collected through previously established data networks such as the Global Climate Observing System (Peterson and Vose 1997). According to Gleason (2002), the most notable problems with temperature data are the presence of both true and false outliers. Whether to remove the outliers depends on the applications. When generating the GDCN 1.0, a principle was taken that uses a minimal set of flags to identify possible outliers. Users can then make the decision whether to include the datum based upon their particular applications. For our application of calculating statistical properties of higher moments and PDFs, an individual outlier will usually not alter our results. Hence we use all the data that, of course, will include some false outliers, and also falsely smoothed data. However, if the application is for catastrophic modeling in the insurance industry, return level of climate extremes, or return period of climate extremes, one should be extremely careful about the false climate outliers.

Other data are also used in this paper to verify our results by comparison. The comparison is mainly made for monthly and annual mean temperatures. The following four monthly and annual mean temperature datasets from 1930 to 2000 reported in Balling and Idso (2002) are used.

1. The raw United States Historical Climatological Network's (USHCN) data (Karl et al. 2000), denoted by RAW. The USHCN contains the raw, time-of-observation-bias (TOB) corrected data from 1,221 stations. The raw data, also known as areal edited data, are the original station data. The TOB indicates that the time-of-observation biases have been removed so that the data are consistent in reflecting the local midnight-to-midnight temperature variation.
2. The FILNET (Fill Missing Original Data in the Network) annual mean temperature time series from 1930 to 2000 (Williams et al. 2007). The FILNET data are the TOB corrected data that have been further adjusted: filling some individual missing data by using the weighted average of data from highly correlated

neighboring stations. The FILNET dataset contains the maximum and minimum SATs of 1,221 USHCN stations.

3. The urban adjusted data, which adjusted the urbanization effects among the FILNET stations by regression (Williams et al. 2007), denoted by URB-ADJ.
4. The IPCC (Intergovernmental Panel on Climate Change) temperature series originally developed by Jones (1994).

3 Methodology

3.1 The mean and higher moments

The climate mean of the SAT in a period of length τ is estimated by the following formula

$$\bar{S}_i = \frac{1}{\tau}\sum_{t=1}^{\tau} S_i\left(t\right), \tag{1}$$

where $S_i(t)$ is the SAT anomaly at station i and time t. Statistically, there is an unbiased estimator of the mean

$$\frac{1}{\tau - 1}\sum_{t=1}^{\tau} S_i\left(t\right).$$

However, physically we want to calculate the average of SAT in a period of time. For monthly or annual data, the length of the period τ may be small for some cases. The statistically unbiased estimate will not be a good estimate of the average. We thus choose to use the biased estimator (Eq. 1) as the first statistical moment. For the daily data, τ is large, and hence there is no noticeable difference between the results from the biased or unbiased estimates.

To assess the change of the second moment, we use standard deviation, which has the same units as the climate mean. The standard deviation is estimated as follows

$$\sigma_i = \left(\frac{1}{\tau}\sum_{t=1}^{\tau}\left[S_i(t) - \bar{S}_i\right]^2\right)^{1/2}. \tag{2}$$

An increase in the variability with or without a change in the mean causes flatter tails on both sides or, at least, on one side of the histogram.

The skewness and kurtosis are estimated based on standardized SAT data by using the following formula

$$M_i^k = \frac{1}{\tau}\sum_{t=1}^{\tau}\left[\frac{S_i\left(t\right) - \bar{S}_i}{\sigma_i}\right]^k, \tag{3}$$

where M_i^3 is defined as the skewness and M_i^4 is the kurtosis (von Storch and Zwiers 1999). The skewness and kurtosis are dimensionless indices. The first two moments M_i^1 and M_i^2 are then assigned to be $M_i^1 = \bar{S}_i$ and $M_i^2 = \sigma_i^2$, respectively.

Positive skewness means that a distribution has its asymmetric tail extending out toward the right. This implies more extreme warm events than cold ones. An increase or decrease of the skewness in time implies an increase or decrease of the warm events. Similarly, if an asymmetric tail extending out toward the left implies more extreme cold events, the skewness is negative. Therefore, a change in the sign of the skewness characterizes a frequency-shift from extreme cold events to the extreme warm events or a shift in opposite direction.

The kurtosis gives the measure of the peakedness of a distribution. If its value is equal to 3, the distribution is said to be neither excessively peaked (leptokurtic) nor flat (platykurtic), but mesokurtic, i.e.,

$$\text{kurtosis} \begin{cases} > 3, \text{leptokurtic, sharply peaked,} \\ < 3, \text{platykurtic, flattened strongly dispersed,} \\ = 3, \text{mesokurtic, equivalent to normal distribution.} \end{cases}$$

A leptokurtic distribution implies a sharp peak compared to a normal distribution, and the corresponding SAT has weaker dispersion. This means that power spectra are more concentrated to the main frequencies. A platykurtic distribution indicates a more flattened and broader distribution again compared to a normal distribution. The SAT then has stronger dispersion and the power spectra are more spread across many frequencies.

For the entire contiguous United States, we consider the spatial average of all the four moments. This can be done by a cosine latitude weight to all the grid values to obtain the area-weighted average moments $\bar{M}^k$:

$$\bar{M}^k = \frac{\sum_{i=1}^{N} M_i^k \cos(\phi_i)}{\sum_{i=1}^{N} \cos(\phi_i)}, \tag{4}$$

where ϕ_i is the latitude of the grid point, and N is the number of grid points considered for the spatial average.

3.2 Data gridding and treatment of missing data

Almost none of the stations in the GDCN dataset or any daily climate dataset around the world has complete records from 1901 to 2000. Missing data have been a common problem encountered in the climate data analysis. Various kinds of methods of filling in the missing values have been proposed in statistics and used in practice. Traditionally, climatologists often choose to retain the data from most reliable stations when assessing climate changes in the monthly scale or longer. In these time scales, the extreme weather conditions have been smoothed out and the climate variance is small relevant to that of daily or weekly data. For example, to calculate regional or global averages of monthly means, Hansen et al. (1999) used only the monthly station data which had a record length of 20 years or longer, claiming that if a data set is too short, then using it does more harm than good. King'uyu et al. (2000) rejected stations if more than 10% of a station's daily records were missing. Our objectives here are to assess the climate changes in terms of higher moments by using the daily data. Both the spatial and temporal variances of climate are very important. Each station's data that have gone through the basic quality assessment and quality control (QA/QC) processes are useful in retaining the spatial and temporal variance. The GDCN dataset has gone through the QA/QC processes. Thus, an effort has been made to utilize all the GDCN station data. Our method is to interpolate the station daily data by using the nearest-station-assignment method, which is equivalent to the Thiessen polygon interpolation and averaging, onto grids of appropriate resolutions

(Shen et al. 2001). Here, three resolutions are used: $0.1° \times 0.1°$, $0.5° \times 0.5°$, and $1° \times 1°$. The gridded daily data are complete in both space and time. The relevant statistics can then be calculated from the gridded data.

The interpolation method is as follows. We create a grid with the designed resolution that covers the contiguous US. The interpolation goal is to assign every grid point the anomaly data of daily maximum temperature, daily minimum temperature, and daily mean temperature on each day from January 1, 1901 to December 31, 2000. For each grid point, a distance table is created in the ascending order of distances between the grid point and the GDCN stations. The anomaly values of the grid point are assigned to be those of the nearest station that has the corresponding anomalies. For a given day, a grid point's daily maximum temperature value comes from only one station, and no spatial averaging is involved. The spatial variance is well preserved this way. For different days, a grid point's anomaly values may come from different stations. The temporal variance of the gridded anomalies may be slightly larger than the reality due to the possible jump discontinuity of the anomaly time series of the two or three stations used for the interpolation of the same grid point. This may happen in the station-sparse regions. However, even when this happens, the amplification of the variance is negligible in the case of contiguous US' area-weighted average. Therefore, we conclude that the method retains the spatial and temporal variances and will not exaggerate the contiguous US' climate variability. The cross validation results in Shen et al. (2001) for the Alberta Province, Canada data support this conclusion.

3.3 Sensitivity testing procedures

To demonstrate the reliability of the results and to test the results' sensitivity to the inclusion of more short-term stations, we use the following methods. First, we compare our results derived from daily data with the existing results on the first moment derived from monthly data. Second, we examine the results calculated from different stations of various lengths of records, including the station records of 50 years, 60 years, ..., and 100 years. The major difference of the networks of stations of the different time lengths is the spatial coverage. The quality of the station data can also be a difference worth probing. Usually, the longer the station record, the higher quality the station data. The details of the comparison and sensitivity studies are described in Section 4.

4 Results

The GDCN daily data are sorted according to the Julian dates. The calendar date January 1, 1901 corresponds to the Julian date 2415386, and December 31, 2000 to 2451910. The daily mean temperature is defined as the average of the daily maximum and daily minimum temperatures. This is the common practice in climate research, but of course, the "mean" here is only an approximation of the true time average of the 24 h temperature of a day, which should be defined by an integration of the temperature function of time divided by the total time of a day. The contiguous US during January 1, 1901–December 31, 2000 had 8,620 stations with records of both maximum and minimum daily temperatures, hence their means can be computed. We will use only these stations in our analysis here to compute climatology, anomalies, and the statistical moments of different orders.

Climatology and anomalies are computed for each station by using the RSM (reference station method) described in the Appendix. Despite the flexibility of the method, the conditions for computing the climatology still exclude about 2,400 stations in the analysis. It leaves only 6,196, 6,236, and 6,234 stations to calculate anomalies of mean, maximum, and minimum temperatures, respectively.

The monthly and annual mean temperatures have a large spatial scale and do not need many stations to get a highly accurate spatial average. Thus, almost all the spatial averages, global or regional, agree with each other well when the anomalies are properly computed. As a matter of fact, the degrees of freedom of the global monthly and annual temperature fields are small, in the order of a few hundreds (Jones 1994; Wang and Shen 1999; Shen et al. 1994). Many stations are redundant for this purpose of finding the global average annual mean SAT. However, in our present analysis of higher moments relevant to variances and extremes, we have retained as many stations as possible. Retaining a large number of stations will certainly introduce some stations of low quality.

The usual approach to calculating the area-weighted average in climate research is to aggregate the station data onto regular grid boxes by simple average and then to use the cosine latitude weighting on the gridded data. However, this is not truly an area-weighted average since the data gridding part is a simple average. The true area-weighted average may be implemented by the Thiessen polygon approach, which is equivalent to a fine-grid-method (FGM) suggested by Shen et al. (2001). The FGM is easier to implement than the direct Thiessen polygon method and interpolates the temperature anomalies of each station to the grids of 0.5° × 0.5° by the nearest-station-assignment method. A grid point mask is made to approximately cover the contiguous US (Fig. 3). The cosine of the grid point is assigned to the weight of the grid box, which is represented by the grid point as the box's approximate centroid.

The SAT's variances at the daily time scale are not only dynamically interesting but also important to many applications, such as defining the first fall frost days, growing degree days, degree days, corn-heat-unit in agriculture, heat index

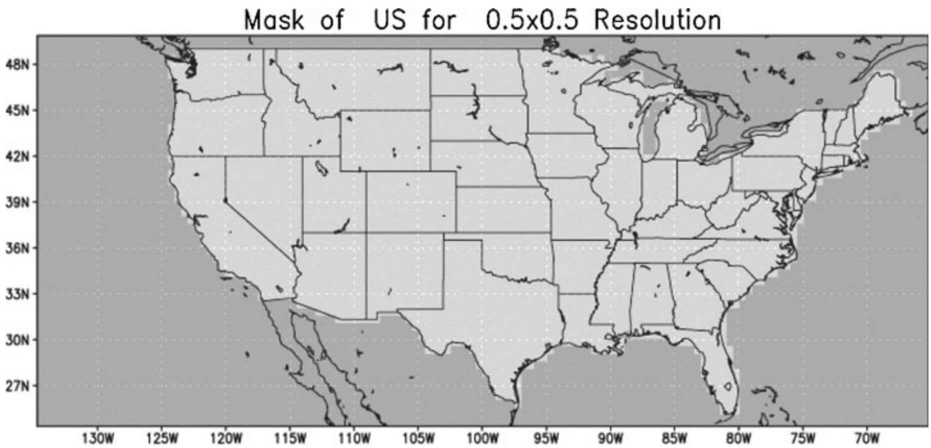

Fig. 3 Mask of the contiguous US with a 0.5° latitude × 0.5° longitude grid

in environmental health, cooling degree days, and heating degree days in utility management. If climate variances have changed in the last century, they should have been reflected more clearly in the daily data than in the monthly or annual data. The daily data analysis is carried out for the SAT anomalies. The 30-year-mean climatology, also called 30-year annual cycle (TAC) (Shen et al. 2005), for daily data is not stable and varies from one 30-year period to another. The daily climate time series may be regarded as piecewise stationary, and hence stationary in a month or in a season. Thus, we choose to define the daily anomalies around the TAC monthly climatology according to Shen et al. (2005), but not to use the nonlinear and non-stationary annual cycle (NAC) for comparison reasons. Thus, the daily anomaly data are generated by subtracting the corresponding monthly climatology defined in the period of 1961–1990 from the daily SAT.

4.1 US average annual mean SAT

We now consider the US average annual mean SAT. The station daily data are gridded onto regular grids by the nearest station assignment method (NSA). Three grids are tested here: 0.1° × 0.1°, 0.5° × 0.5°, and 1.0° × 1.0°. The results from these three grids are almost the same. The one corresponding to the 0.5° × 0.5° grid with the GDCN data are shown in Fig. 4 for the daily-mean SAT together with four other datasets. The annual mean of the daily-mean SATs from G50, G60, ..., and GALL are shown in Fig. 5.

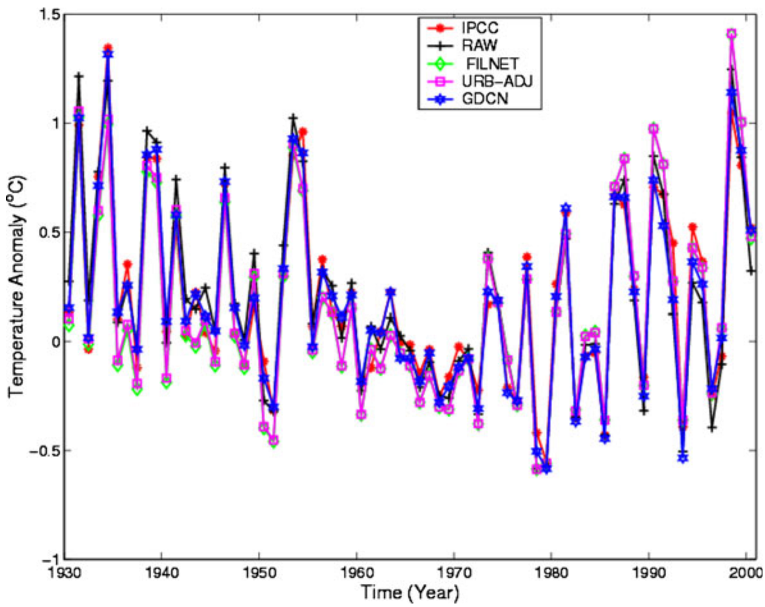

Fig. 4 The US average annual mean of the daily mean SAT computed from different data sets: the United States Historical Climatological Network's (USHCN) raw data, FILNET, and the urban adjusted data set, the IPCC, and the GDCN. The GDCN result is obtained from the area-weighted average with the 0.5° × 0.5° grid and the station's annual anomalies computed from the monthly anomalies. The other results are from Balling and Idso (2002)

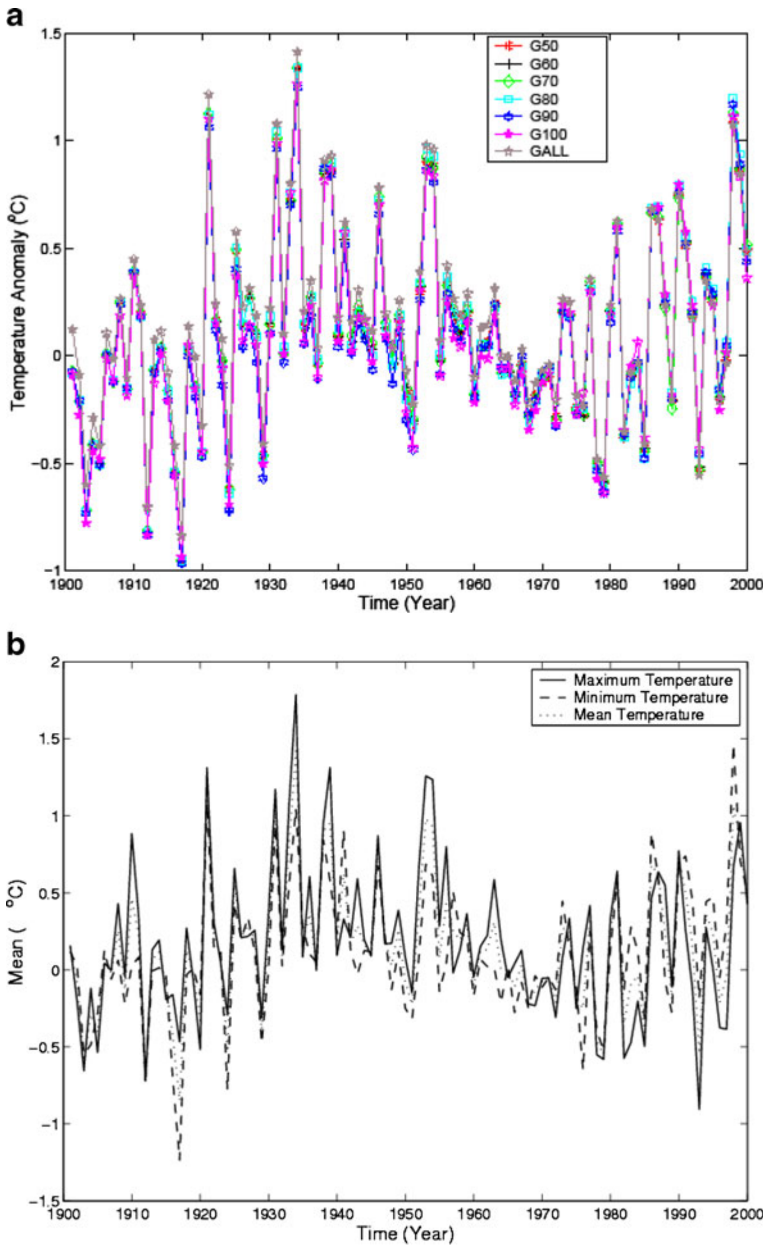

Fig. 5 **a** The US average annual mean of the daily mean SAT calculated from the GDCN's data sets with different minimum record lengths by using the area-weighted average method. The annual data are computed directly from the daily anomalies. **b** The area-weighted US average of annual mean of the daily mean, maximum, and minimum SAT computed from the GDCN daily anomalies

Balling and Idso (2002) computed and compared the average of the mean temperature of four different climatological databases: the United States Historical Climatological Network's (USHCN)'s raw data, the USHCN FILNET data (Fill

Missing Original Data in the Network), the USHCN urban adjusted data, and the IPCC2001's series of the contiguous US for the period 1930–2000. All the US spatial averages reported in Balling and Idso (2002) were computed by the area-weighted method. In April 2005, the Oak Ridge National Lab in collaboration with the NCDC also documented the USHCN data and made some comparison studies (http://cdiac.ornl.gov/ftp/ushcn_monthly/ushcn_monthly_doc.html). The number of stations used in the USHCN's data set (raw, FILNET, and the urban adjusted) is 1,221, while the IPCC's series consists of gridded data sets with a resolution of 5.0° × 5.0°, and all these data sets are at monthly resolution. The results from our seven sub-networks (G50, G60, G70, G80, G90, G100, and GALL) are compared with the results of Balling and Idso (2002), but only the G70 results are shown here (Fig. 4). This G70 network includes 785 stations and has a reasonable spatial coverage of the contiguous US.

We compare our annual mean computed from the daily anomalies with the results reported in Balling and Idso (2002). Figure 4 shows the five curves: four from Balling and Idso (2002) and one from the GDCN G70 sub-network and a 0.5° × 0.5°. The USHCN network and the GDCN G70 sub-network share almost the same stations.

Figure 5a shows an increasing trend of the annual average of the daily-mean SAT (T_{mean}) from 1901 to 1934, a decreasing trend from 1935 to 1975, and an increasing trend from 1976 onward. However, this is only a rough and visual assessment of the change points. Rigorous detection of change points in a time series needs to go through a statistical procedure (Picard 1985), which is not the emphasis of this paper. The characteristics of the trends for annual average of the daily-maximum SAT (T_{max}) and daily-minimum SAT (T_{min}) are similar to those of the daily-mean SAT in the entire period of analysis, but the magnitudes of the changes show some differences (see Fig. 5b). Before the 1970s, the minimum SAT anomalies were mostly below those of maximum SAT anomalies. In the 1980s and 1990s, the order was switched. The switch implies that the daily minimum temperature had changed more than daily maximum temperature. Particularly, the sharp increasing trend of the minimum temperature anomalies from 1917 to 1934 and that from 1976 to 1998 is noticeable. The fast increase of the daily minimum temperature and the relatively slow increase of the daily maximum temperature imply a decrease of the diurnal temperature range (DTR). These conclusions are well known and are well reflected in our results (IPCC 2007, Sub-section 3.2.2.1 and Fig. 3.2).

The linear regression trend of the GDCN data is also compared with those of Balling and Idso (2002). Considering only the data in the period from 1930–2000, a statistically significant trend of 0.05°C/decade at 5% level of significance is obtained in the GDCN's data, which is exactly the same as that observed by Balling and Idso (2002). Therefore, we can conclude that our GDCN results from the area-weighted average of the daily mean SAT for the contiguous US are in very good agreement with those from the USHCN's raw data sets reported by Balling and Idso (2002), although we have used the daily anomalies. This verifies the reliability of using the daily anomalies in climate change assessment. Next, we will use the daily SAT anomalies to assess the climate changes in terms of higher statistical moments.

4.2 The higher statistical moments from the GDCN daily anomaly data

After having justified the reliability of the area-weighted average method and the GDCN daily mean SAT, we have gained the confidence in our daily anomalies.

Equation 2 is used to compute the standard deviation, and its square is the variance of station or grid point SAT data, and Eq. 3 with the value of k equal to 3 and 4 are used for skewness and kurtosis, respectively. Equation 4 is used to compute the area-weighted average.

The GDCN daily anomalies of the mean, maximum, and minimum SAT are used to compute the higher statistical moments. The area-weighted method is used to calculate the contiguous US average. The station anomaly data are interpolated onto the $0.1^\circ \times 0.1^\circ$ grid. The higher order moments are computed at each grid point for each year according to Eqs. 2 and 3 for $\tau = 365$ by excluding February 29 in the leap years. Then the area-weighted average of the moments at the grid points is calculated for the contiguous US according to Eq. 4. Figures 6, 7 and 8 show the results of the variance, skewness, and kurtosis for the mean, maximum, and minimum SATs, respectively.

Figure 6 shows that the general trend of the SAT variance over the contiguous United States is decreasing, although the rate is very slow. We do not intend to conclude a statistically significant decreasing trend of variance, yet we are confident to claim the non-increment of the SAT's variance over the contiguous US. Our claim is not the first conclusion on non-increment of SAT variance over land. Gong and Ho (2004), Balling (1997) and Michaels et al. (1998) also claimed non-increment of climate variances using different climate parameters or over different spatial regions. Our work is more comprehensive, uses significantly more data, and has gone through an extensive comparison.

The two conspicuous large peaks at 1936 and 1989 in Fig. 6 might be caused by strong La Nina events. The El Nino oscillations supposedly imply large variance, but

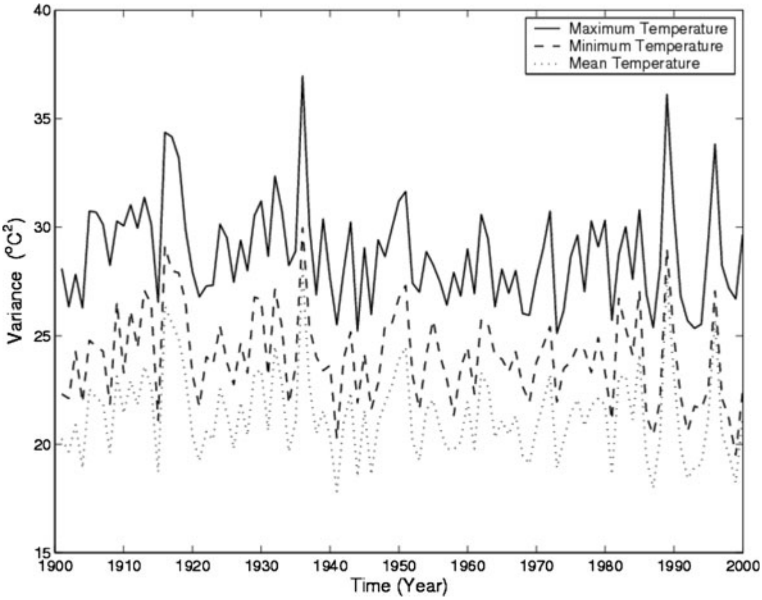

Fig. 6 The area-weighted US average of annual variance of the daily mean, maximum, and minimum SAT computed from the GDCN daily anomalies

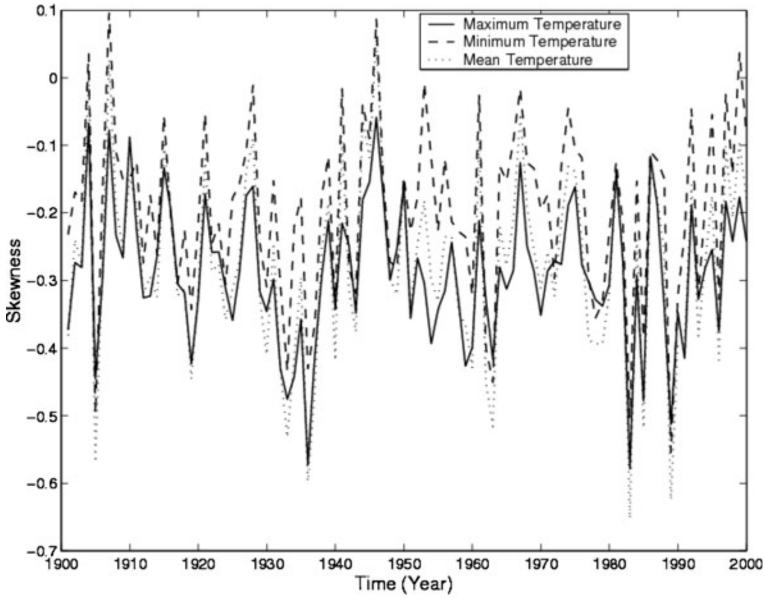

Fig. 7 The area-weighted US average of annual skewness of the daily mean, maximum, and minimum SAT computed from the GDCN daily anomalies

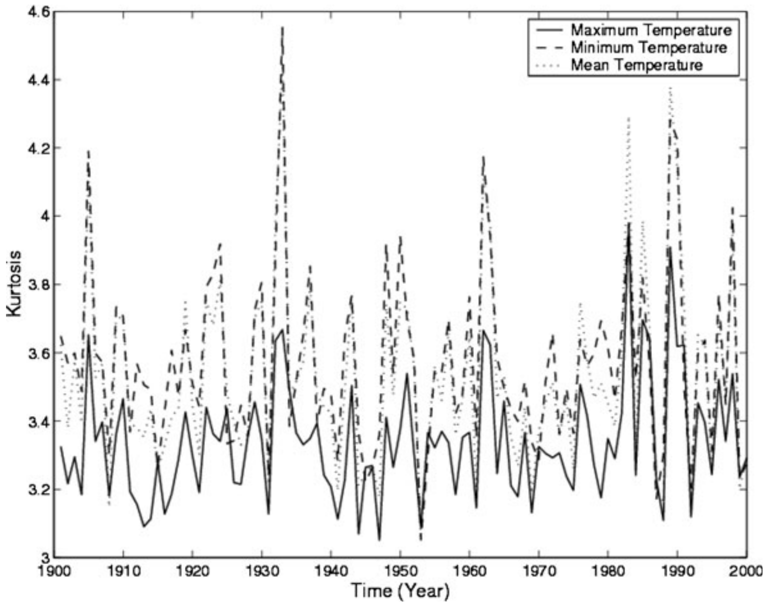

Fig. 8 The area-weighted US average of annual kurtosis of the daily mean, maximum, and minimum SAT computed from the GDCN daily anomalies

they are transient through December and hence the large variance is often divided into two separate years. Between these two peaks, the variance demonstrates little variation. So the annual climate variance changed little from the 1940s to 1980s.

Figure 6 also shows that the variance of the maximum SAT is always the largest, that of the minimum SAT is the second largest, while that of the mean SAT is the smallest. This fact implies that the maximum SAT is most volatile, and the mean SAT is the least volatile. This conclusion is supported by other studies (Michaels et al. 1998; IPCC 2007, Sub-section 3.8.2.1).

Figure 7 shows that the skewness of all the temperatures is negative most times. This implies left-skewed PDF, i.e., skewed toward cold. So the number of cold anomalies is larger than that of warm ones. The general trend of skewness was decreasing (i.e., becoming more asymmetric) from 1901 to 1936, increasing from 1937 to 1946 (i.e., becoming more symmetric), and increasing from 1990–2000. The lowest value of the skewness was at 1983, a year with very strong 1982–1983 El Nino. The El Nino event may enhance the un-symmetric distribution of the SAT, particularly the maximum daily SAT. Amplitude of skewness of the maximum temperature is often the largest, and that of the minimum temperature is often the lowest. It is thus intriguing that the minimum temperature is more symmetrically distributed than the mean temperature. These observations indicate the existence of more cold extremes in all temperatures, but the occurrence of such cold extreme events is more frequent in the maximum temperatures than in the other two.

We also calculated the annual skewness from the monthly data. The skewness is mostly positive. (The figure is not shown here.) This means the cold extreme anomalies reflected in the daily data have been averaged out. Thus, the skewness computed from the monthly data appears not very useful in inferring the climate behavior. The daily data are the best to display the climate distribution that is closely relevant to practical applications of climate resources.

Figure 8 shows the area-weighted US average of annual kurtosis of the daily mean, maximum, and minimum SAT computed from the GDCN daily SAT anomalies. The figure shows two peaks at 1933 and 1989. The kurtosises of the mean, minimum and maximum temperatures are all greater than 3, and hence are excessively peaked (leptokurtic). The kurtosises of the mean and minimum temperatures are comparable in size and interweaving with each other in the 100 years period. They are clearly larger than that of the maximum temperature. These observations imply that the maximum temperature's distribution is more widely spread than those of minimum and mean temperatures, and hence, the maximum temperature has stronger dispersion capability than the other two temperatures. This finding is consistent with the larger variance of the maximum temperature than the other two temperatures.

To examine the dispersion behavior at daily and monthly scales, we have compared the kurtosis calculated from daily and monthly data. For the daily maximum SAT, the 100-year mean kurtosis is around 2.7 for the monthly data and is around 3.3 for the daily data. For the daily minimum SAT, the 100-year mean kurtosis is around 3.0 for the monthly data and is around 3.6 for the daily data. Thus, the distributions of monthly maximum and minimum temperatures are platykurtic (less peaked), and those of the daily temperatures are leptokurtic (more peaked). The month anomalies have stronger dispersion. They are more widely distributed than the daily anomalies, compared with the corresponding fitted normal distributions of monthly and daily temperatures, respectively.

IPCC (2001) reported 1910–1945 and 1976–2000 as the periods of global warming and 1946–1975 as the period of cooling. To learn the behaviors of the higher moments over the United States in these global warming and cooling periods separately as well as through out the whole century, the linear trends (irrespective of the statistical significance) of the different moments $M^k(t)$ are computed for 1901–2000, 1910–1945, 1946–1975, and 1976–2000, respectively. The linear model is expressed in terms of

$$M^k\,(t) = A + Bt + \varepsilon, \tag{5}$$

where A and B are the fitting parameters, t is the time, and ε is the model error. The unit of A would be same as that of the moment under consideration (e.g., °C for the mean and °C^2 for the variance), and that of B would be GT^{-1}, where G is the units of the moment, and T is the units of time, which can be expressed either in days, years, decades, or centuries.

Table 1 lists the slope B values of the different moments for the daily mean, maximum and minimum SAT. The 1901–2000 trend of the annual mean of the daily mean temperature is increasing at a rate of 0.148°C/century, which is close to the figure of 0.16°C/century obtained by Hansen et al. (2001) by using the raw USHCN data. The mean temperature shows increasing trends in the periods 1910–1945 and 1976–2000 and a decreasing trend in 1946–1975. These trends are comparable to those of the global average temperature.

The trend of the variance of the mean temperature is decreasing in all periods. In the period 1901–2000, the decreasing slope is −0.992°C^2 per century; i.e., the mean temperature of the contiguous US become less variable, about a degree Celsius^2, in the past century. The skewness of the mean SAT also shows a small decreasing trend in all periods except 1976–2000, implying the chances for slightly more cold extremes. However, the increased frequency of more extremely cold events was not

Table 1 Trends of different moments for the daily mean temperature

Period	Trends of moments (per century)			
	Mean (°C)	Variance (°C^2)	Skewness	Kurtosis
Mean SAT				
1901–2000	0.148	−0.992	−0.021	0.040
1910–1945	1.772	−4.888	−0.131	0.148
1946–1975	−1.638	−2.733	−0.028	−0.317
1976–2000	2.499	−5.954	0.645	−0.504
Max SAT				
1901–2000	−0.058	−1.684	−0.025	0.130
1910–1945	1.574	−6.089	−0.201	0.220
1946–1975	−2.362	−5.436	−0.032	−0.088
1976–2000	1.428	−2.467	0.288	−0.170
Min SAT				
1901–2000	0.367	−1.755	−0.004	−0.016
1910–1945	1.984	−5.393	−0.048	0.047
1946–1975	−0.927	−2.670	0.027	−0.345
1976–2000	3.627	−11.226	0.756	−0.594

clearly observed due to the fact that the absolute value of the decreasing slope was very small, only −0.021 per century. Further, the increasing trend 0.645 per century of the skewness in 1976–2000 implies the chances for increased warm extremes in this period.

The trend of the kurtosis in 1901–2000 was increasing, but again very small, only 0.040/century. The dispersion was thus becoming slightly weaker, i.e., a slightly reduced variability. However, the kurtosis decreased in 1946–1975 and 1976–2000. The relatively large magnitude of decreasing slope of kurtosis in the 1976–2000 period implies stronger dispersion in the last two decades of the twentieth century.

Table 1 for max SAT reveals that the annual mean of the daily maximum temperature had a decreasing trend in 1901–2000 and 1946–1975 and an increasing trend in 1910–1945 and 1976–2000. The 30 years decrease from 1946 to 1975 prevailed the changes in other periods and was about 0.7°C. The variance has a decreasing trend in all the periods, with an overall trend of $-1.68^{\circ}C^2$ per century. The prevailed change was from 1910 to 1945 and the net change was about $2.2^{\circ}C^2$. The skewness also showed a decreasing trend in all periods except in 1976–2000. Again, the magnitude of the changes is small. The kurtosis showed an increasing trend in 1901–2000 and 1910–1945 and a decreasing trend in 1946–1975 and 1976–2000 with small changing amplitude as well.

Table 1 for min SAT shows a significant increase of the daily minimum temperature in the last century and the net increase was 0.37°C. The outstanding result is the large positive slope 3.63°C/century in the period of 1976–2000 and the net increase in this short period is about 0.9°C. It was this fast change that alarmed the scientific community and general public on the climate change issues at the last decade of the twentieth century. The decrease of the minimum temperature from 1946 to 1975 was a reflection of the overall post-World War cooling trend. The variance showed a decreasing trend in all periods, with an overall trend of $-1.76^{\circ}C^2$ per century. The large decreasing slope $-11.23^{\circ}C^2$ per century from 1976 to 2000 is a reflection of the sharp decrease of the diurnal cycle's amplitude (or called the diurnal temperature range, or DTR) (IPCC 2007). The skewness demonstrated a decreasing trend in 1901–2000 and 1910–1945 and an increasing one in 1946–1975 and 1976–2000. The large increasing slope 0.76/century indicates that the daily minimum temperature tended to be more symmetrically distributed from 1976 to 2000. The changes in other periods had very small magnitudes. The kurtosis was decreasing in all periods except in 1910–1945. The fastest change was in 1976–2000 (−0.59 per century). The decrease implies that the minimum temperature's dispersion became stronger from 1976 to 2000.

4.3 PDF changes

Although the changes of the first four statistical moments can reflect much about the SAT's variations, a more comprehensive demonstration of the changes of the SAT can be reflected in the changes of the SAT's PDF. The full assessment of the PDF changes, however, requires many more data, which are not available. A compromise is to assess the PDF changes in selected periods by assuming the property of piecewise stationarity of the SAT. Further, the PDF changes are assessed by only using the data from the stations which experienced large climate changes. This will result in dropping many stations, and the remaining stations will not completely cover the entire contiguous US.

The SAT histograms in the periods 1910–1945, 1946–1975, and 1976–2000 are calculated and compared. The trend of the different temperatures of the individual stations in the G70 network is analyzed. The daily data sets of the mean, maximum, and minimum temperatures in the summer (June, July, and August) and winter (December, January, and February) seasons for 50 stations where the largest increasing trend in the period of 1910–2000 is detected are used to construct the histograms. Locations of the 50 most influential stations in terms of largest climate change of the maximum and minimum SAT are shown in Fig. 9a and b, respectively. These stations are mostly located in the western US.

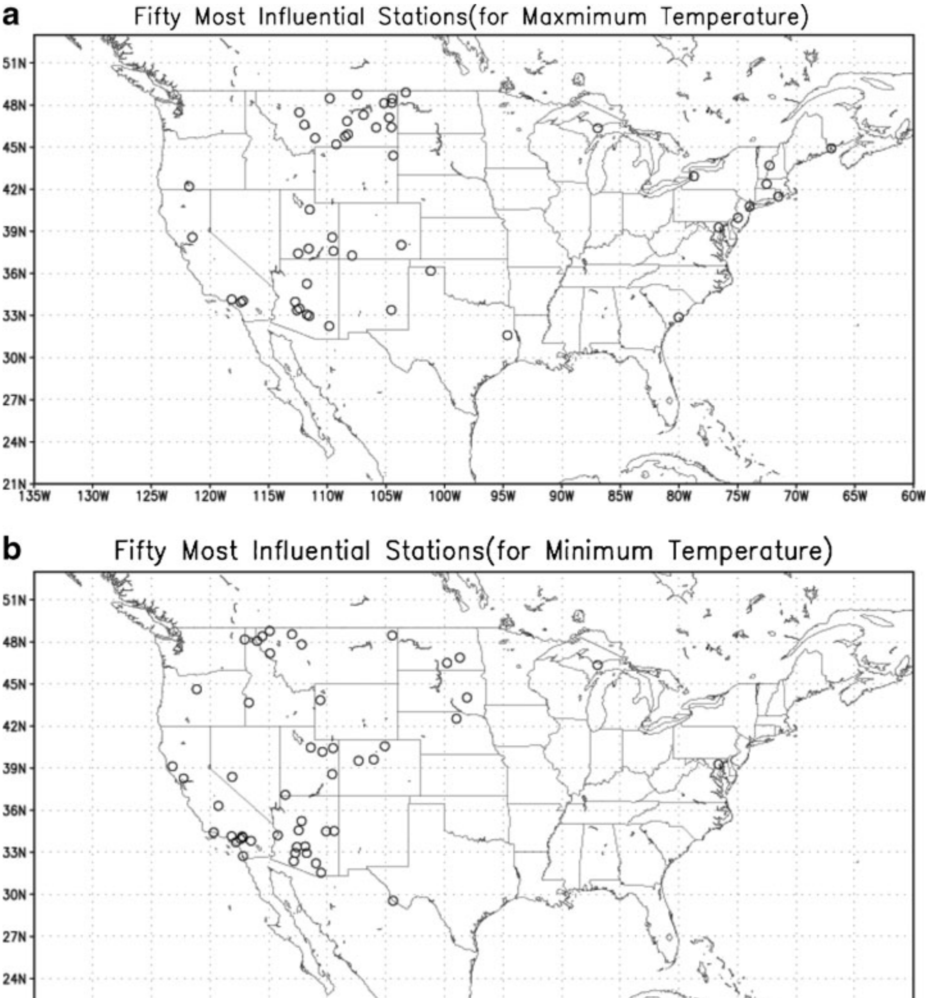

Fig. 9 **a** Locations of the 50 most influential stations for the maximum surface air temperature. **b** Locations of the 50 most influential stations for the minimum surface air temperature

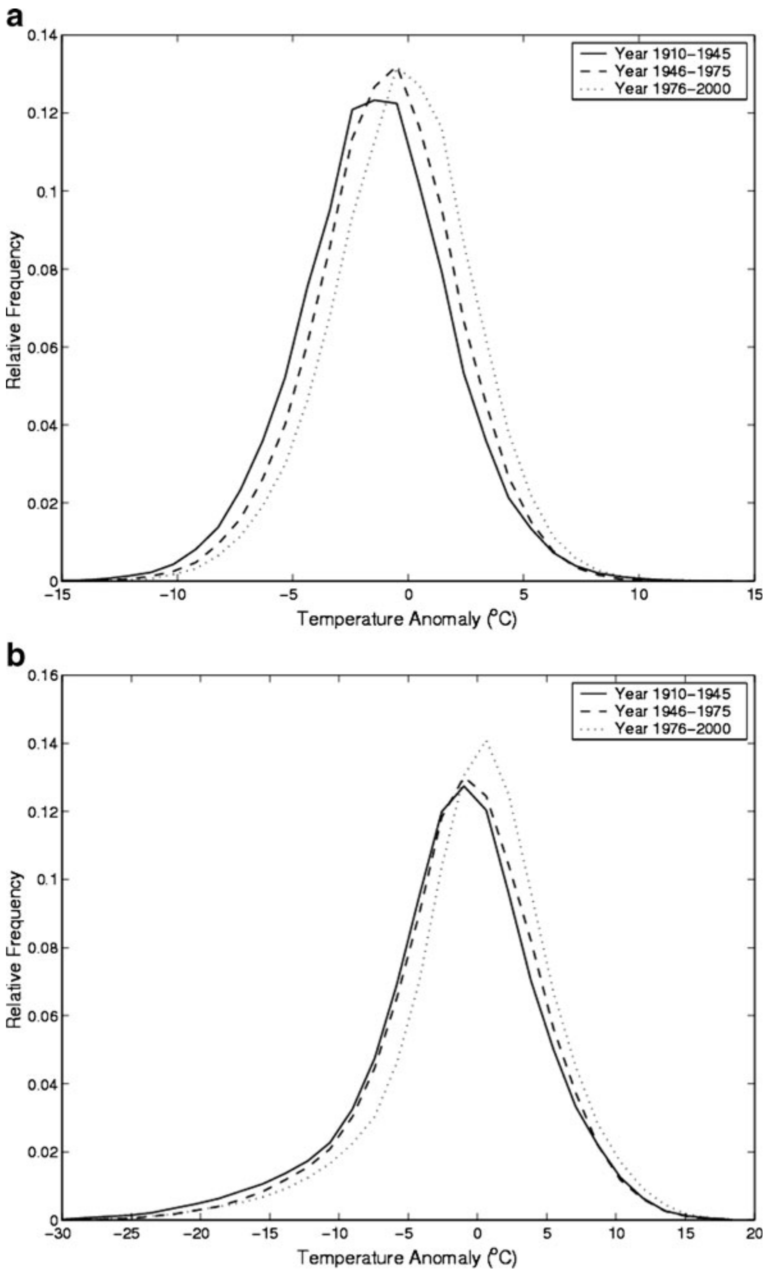

Fig. 10 Three histograms of the daily mean SAT of 50 stations where the highest increasing trend occurred in the summer season (**a**), and in the winter (**b**). Three histograms of the daily maximum SAT of 50 stations where the highest increasing trend occurred in the summer season (**c**), and in the winter (**d**). Three histograms of the daily minimum SAT of 50 stations where the highest increasing trend occurred in the summer season (**e**), and in the winter (**f**)

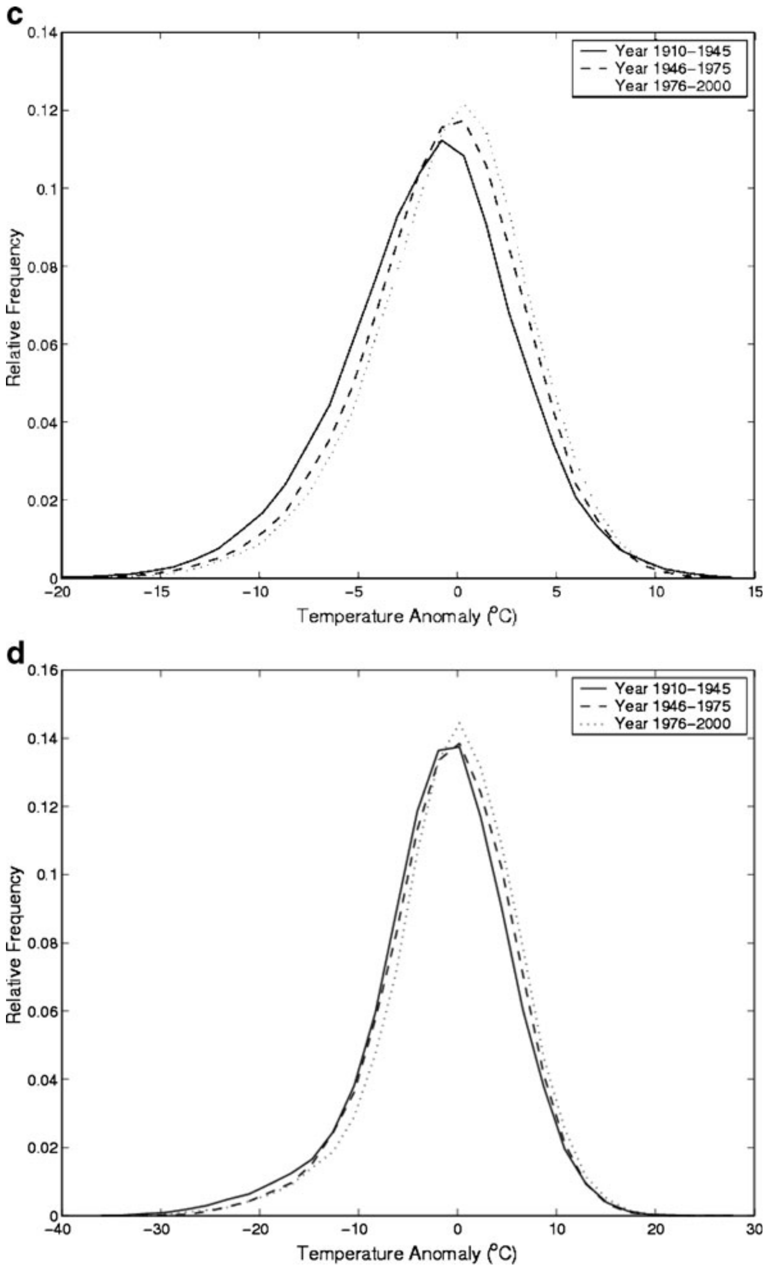

Fig. 10 (continued)

Since the selected most influential stations do not cover the entire contiguous US, the results do not fully reflect the spatial average statistics of the US. However, our results raise some interesting climate-change questions. For example, how are the moments' changes reflected in the PDF changes? Do the PDF changes reflect

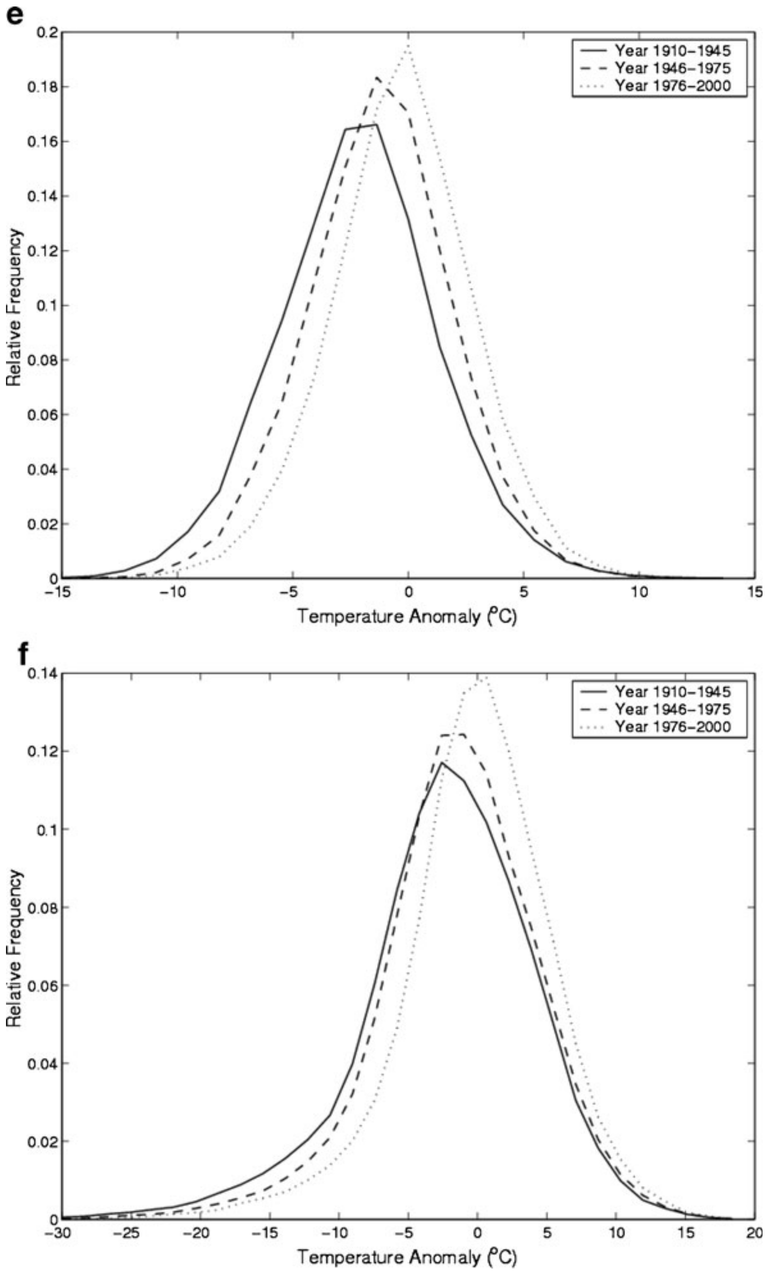

Fig. 10 (continued)

changes of climate dynamics or atmospheric circulation patterns? Questions like these may not be apparent from the histogram constructed by using the data of all the stations for the entire year. The use of daily data in different periods and different seasons for the most influential stations may help answer such questions.

Our computed histogram results and statistical moments are shown in Fig. 10 and Table 2 for six cases with three parameters (maximum SAT, minimum SAT, and mean SAT) and two seasons (winter and summer). The results imply the following:

1. Right shift of the histograms: A clear shift to the right is observed for all the histograms for both winter and summer and for all the three temperatures as the time period moved from 1910–1945 to 1946–1975 to 1976–2000 (see Fig. 10a–f). However, the magnitudes of the right shift are not the same for the six cases.
2. Increase of the mean: The net increases of the winter (summer) season's mean from 1910–1945 to 1976–2000 are 1.52°C (1.17°C) for the daily mean SAT, 1.20°C (1.07°C) for the daily maximum SAT, and 2.22°C (1.89°C) for the daily minimum SAT. The largest increment is in the daily minimum SAT as concluded in many studies.
3. Decrease of the variance: The net decrease of the winter (summer) season's variance from 1910–1945 to 1976–2000 are $-5.41°C^2$ ($-1.10°C^2$) for the daily mean SAT, $-5.06°C^2$ ($-2.89°C^2$) for the daily maximum SAT, and $-10.77°C^2$ ($-2.56°C^2$) for the daily minimum SAT. The large decrease of the winter minimum SAT's variance over the locations of the 50 stations is mainly due to the decreasing number of extremely cold days. These changes are reflected in an increase of the height, and hence a decrease of the width, of the histograms.
4. No significant changes of skewness and kurtosis: The histograms do not show a clear change of shapes or symmetry, except Fig. 10e and f. Thus, the changes of skewness and kurtosis are only noticeable in the minimum SAT in winter. The

Table 2 Moments of the average temperature in the winter and summer seasons for different periods of time in the 50 stations where the highest increasing trends occurred

Season	Moments				
	Period	Mean (°C)	Variance ($°C^2$)	Skewness	Kurtosis
Mean SAT					
Winter	1910–1945	−0.916	38.886	−0.692	4.409
	1946–1975	−0.359	33.295	−0.560	4.052
	1976–2000	0.608	33.475	−0.732	4.672
Summer	1910–1945	−0.965	10.737	−0.043	3.448
	1946–1975	−0.462	9.521	−0.151	3.353
	1976–2000	0.209	9.642	−0.173	3.430
Max SAT					
Winter	1910–1945	−0.627	48.818	−0.610	4.128
	1946–1975	−0.020	43.035	−0.467	3.606
	1976–2000	0.569	43.758	−0.590	4.087
Summer	1910–1945	−0.907	19.696	−0.280	3.538
	1946–1975	−0.263	17.148	−0.401	3.577
	1976–2000	0.166	16.802	−0.392	3.700
Min SAT					
Winter	1910–1945	−1.499	40.998	−0.656	4.424
	1946–1975	−0.659	34.375	−0.549	4.517
	1976–2000	0.720	30.228	−0.602	4.654
Summer	1910–1945	−1.500	11.926	0.054	3.411
	1946–1975	−0.568	9.872	−0.026	3.300
	1976–2000	0.393	9.368	−0.068	3.403

net changes of skewness and kurtosis are 0.05 and 0.23, respectively, implying that the histogram had become more peaky and more symmetric.

The above findings suggest that the changes of SAT's third and fourth moments are more or less independent of the first moment. However, further study is needed to support this conclusion. The statistics corresponding to the histograms of the winter season experienced larger changes than those of the summer season. Among all of the three temperatures, the statistics of the minimum temperature of winter had the largest changes.

5 Conclusions and discussion

We have investigated the climate changes using the daily minimum, maximum, and mean SAT data from January 1, 1901 to December 31, 2000 over the contiguous US by considering the mean, variance, skewness, and kurtosis in different seasons, different sub-periods, and different sub-networks of stations. The statistical moments at station locations have been spatially averaged by the area-weighted average method. The US averages of the mean SATs have been compared with the existing results obtained from other station networks and analysis methods. There are no existing results to make comparisons with for variance, skewness, and kurtosis. An increase of the annual mean of the SATs and a decrease of the annual variance are obvious, but no clear changes have been observed in the skewness and kurtosis in the last century.

The SAT's PDF changes have been assessed through the 50 most influential stations in three periods: 1901–1945, 1946–1975, and 1976–2000. Obvious changes have been observed in the increase of the PDF's mean and the decrease of the PDF's variance, but no clear changes have been identified for the skewness and kurtosis. However, the winter's minimum daily SAT might have become more peaky and more symmetric. These conclusions suggest that the mean and variance are closely related, while the skewness and kurtosis are independent to the first two moments.

Our observation of the widespread reduction in variability agrees with the results of Karl et al. (1995), Balling (1997), Michaels et al. (1998) and IPCC (2001). Our conclusions of no large changes of skewness and kurtosis do not have existing results for comparison. To our knowledge, only Vinnikov and Robock (2002) attempted to compute the third- or fourth-order moments of the climatic time series. However, their approach and their data sets are differed from ours. Gong and Ho (2004) analyzed only up to the third order moment.

The conclusion of the reductions in the temperature variability seems counterintuitive since many reports state that climate has become warmer and more violent particularly with stronger cyclones over the ocean (Knutson et al. 2010; Price 2009; Del Genio et al. 2007). However, intensity of extreme weather does not automatically imply more temperature variance over the land. There may be a difference of the temperature variance between the land and the ocean, and between the winter and the summer. IPCC (2007) reported a reduction in the temperature variability in winter (Sub-section 3.8.2.1. of IPCC 2007). This is perhaps due to the fact that in winter, the increase of the minimum temperature anomalies plays the dominant role in the climate change, and the variance of the minimum SAT is smaller than that of maximum SAT. Because understanding the behavior of change of the variance, skewness and kurtosis is very important as they contain key information about the extremes of the climate system besides helping detect changes of mean, more studies on the higher moments are needed to verify or support the findings of this study and other similar ones.

Several points are worth discussing. The first is the problem of data homogenization and update. Most of the long-term climatological series have been affected by a number of non-climatic factors that make these series' data sets unrepresentative of the actual climate variation occurring over time (Peterson et al. 1998). Therefore, such data sets must be homogenized before they can be used for climatological analysis (Vincent et al. 1999). For homogenization, a base or reference station and the candidate stations (whose data are to be adjusted based on those of the reference station) are required. However, adjustments of the temperature at a daily scale would be very difficult due to the high variability of the daily observations (Vincent et al. 1999). For global or large-area data selecting the base and candidate stations and doing such adjustments would be difficult. Very often, quantitative information required to adjust the station data is not available. Fortunately, the random component of such error tends to average out in large area averages and in calculations of temperature change over long periods; therefore, stations' data do not always need to be homogenized (Hansen et al. 1999). The updated dataset of the GDCN is the GHCN-D (Global Historical Climatology Network-Daily) (Durre et al. 2008; Menne et al. 2009). More data have been included and more quality control procedures have been implemented. Over the contiguous US, the GDCN data coverage is almost as complete as GHCN-D. The differences between GHCN-D and GDCN are reflected in some individual stations and have little effect on the regional average results like those presented in this paper.

The second is about the methods of spatial averages and spatial coverage of stations. We have concluded that the area-weighted average is less sensitive to the station coverage than the simple arithmetic average. Neither of the averaging methods has considered climate dynamics and atmospheric circulation. Empirical orthogonal functions (EOF) in a moving time window can summarize the principal properties of climate dynamics and atmospheric circulation. If the EOF type of optimal spatial averaging method, called the reduced spectral optimal averaging (RSOA) by Folland et al. (2001) or spectral optimal averaging (SOA) by Shen et al. (1994, 1998), is applied, the results will be even less sensitive to the spatial coverage. The main reason is that the SOA maximizes the variance representativeness by using only a few eigenfunctions and eigenvalues, and hence the spatial length scales are taken into account, and consequently the limited degrees of freedom are well represented by the corresponding number of stations and many other stations have become redundant. Although the SOA has been applied to the global SAT and regional SST, it has not been applied to the regional SAT yet.

The third is about the time resolution. It is well known that the longer the time scale, the smaller the variance for land stations. The daily time scale is the minimum we can work on now due to the availability of data. Thus, it is very important for the climate research community to explore the observed data at the shortest possible time scales: day, hour, or minutes. The wind direction and speed need data in minutes. The flood watching needs rain data in hours. The temperature stress on human lives needs data in days. The study of the climate change at the shortest possible scale may make it possible to assess the changes through PDFs, which are very important in optimizing an operational decision-making system.

The fourth is about the linearity of the trend. Vinnikov and Robock's conclusion was based on linear trends. However, the trends may be nonlinear and instantaneous. It is worthwhile to consider the nonlinear trends of the higher moments of the detrended series (Cunderlik and Burn 2003). The wavelet analysis and the empirical mode decomposition methods may be helpful in climate data analysis (Huang and Shen 2005).

The fifth is about the change of the sea surface temperature (SST) variance, which may not follow the same trend as the surface air temperature (SAT). Xue et al. (2003) analyzed the SST data from 1871 to 2000 and discovered both positive and negative trends over Atlantic and Pacific from two different reconstructed SST data. The inconsistency of the results led them to conclude that the seasonal standard deviation of the SST over either the Atlantic or Pacific might not be well resolved by the reconstructed data. Improved reconstruction of the SST data would be needed for this study. It is still interesting to find out whether the change of SST variance in the twentieth century was synchronized with the land SAT variance. Recently, Knutson et al. (2010) demonstrated an increase of tropical cyclone intensity as the global climate becomes warmer in the twenty-first century. Therefore, the variances of the land SAT and the SST may not be synchronized and demonstrate different behaviors in time, which is associated with different climate dynamics. It is interesting to explore the multi-decadal variations of the variances.

Acknowledgements The data used to compare our results on the first moment with those of the published papers were provided to us by Robert C. Baling Jr. of Arizona State University. The mask data for the grids of 0.1° latitude × 0.1° longitude, 0.5° latitude × 0.5° longitude, and 1.0° latitude × 1.0° longitude resolutions were generated by Xiaodong Zeng of the University of Arizona during his residence at Information Research Laboratory of the University of Alberta. The plots of the masks for the land and ocean and the contiguous US were generated by the Grid Analysis and Display System (GrADS) scripts provided by the Center for Ocean-Land-Atmosphere Studies (COLA). Thomas R. Karl of the US National Climatic Data Center and Gerald R. North of Texas A&M University were consulted for using the higher statistical moments to detect climate changes.

This work is financially supported by the US NOAA Office of Global Programs, the Canadian MITACS, and the University of Alberta, Canada.

Appendix: RSM for calculating climatology and anomalies

This appendix describes details of the RSM (reference station method) for calculating climatology and anomalies, which was developed and applied in Shen et al. (2004), but the paper did not provide detailed documentation of the procedures. This appendix fulfills the gap. The RSM is a hybrid of CAM (climate anomaly method) (Jones et al. 1982) and FDM (first difference method) (Peterson et al. 1998). It extends the flexibility of both CAM and FDM and can include more stations. To make the RSM description self-contained in this paper, we also briefly describe the procedures of CAM and FDM.

It is well known that the SAT data from the stations at different altitudes and different seasons cannot be directly averaged to make a physical sense for assessing the climate changes. The spatially and annually averaged quantities should be the departures from the normal of each station. The normal is called the climatology, and the departure is called the anomaly. One usually uses the climate mean of the period from 1961–1990, called the normal period, to approximate the climatology. The anomaly is defined as the deviation from the mean. The anomaly $S_i(t)$ at time t and at station i is given by

$$S_i\,(t) = T_i\,(t) - c_i\,(m)\,, \tag{6}$$

where $T_i(t)$ is the observed temperature data at time t (day or month) and $c_i(m)$ is the 1961–1990 monthly mean (or gross climatology) for month m at station i.

The CAM of calculating the climatology and anomalies works in the following way (Jones et al. 1982). Since the monthly climatology is used, the daily temperature data need to be used to compute the monthly mean temperature, which is in turn used to compute the 1961–1990 mean, i.e., the climatology. However, in practice many stations do not have the full 28, 29, 30 or 31 days of data in a month. The mean over a subset of the days is thus used to approximate the full month mean. A condition is imposed that the subset has to include at least 21 days of data in a month in order for this month's mean to be calculated. In this way, the monthly mean SAT from January to December and from 1961 to 1990 may be computed for the stations and months that satisfy the 21-day condition. Of course, many stations and months do not satisfy the condition and hence their means are not calculated. The 30-year mean from 1961–1990 is computed for stations also from a subset of the months. The condition is again at least 21 months in the 30 months from 1961 to 1990. When the climatology is calculated, the anomalies can be calculated according to Eq. 1. If not enough records are available for a station in the normal period, then the climatology for that station cannot be computed and anomalies cannot be computed either. Thus, the CAM will result in many stations whose anomalies cannot be calculated, since many stations did not cover much of the period of 1961–1990.

Shen et al. (2004) introduced the reference station method (RSM) of calculating anomalies. This method is a mixture of the CAM and the FDM (first difference method) and requires less strict conditions. A reference station and the first differences are used in this method. The RSM is implemented in this study, which works as follows. Suppose the climatology of a station (Station A) has to be computed, but this station does not have enough records for the period 1961–1990 to satisfy the conditions required to compute the climatology.

We still compute the monthly data from the 21-day rule. First differences for the monthly data at every station and every month are calculated, except for the first year. For a station, each month has a sequence of first differences. If a station has a temporal gap, then the first difference cannot continue, and a new reference year starts from the first year of the next section of observations. Because of the change of the reference year, the integration of the first difference may not recover the anomaly values of the station after the gap. In order to keep as many stations as possible while maintaining acceptable accuracy, the long term mean values (not the anomalies) replace the missing values if a gap is less than or equal to two consecutive years.

The procedure of anomaly computations for the candidate station A is demonstrated as follows. Let Y_b be the beginning year of the data of the station A, and Y_e the end year of the station A. If the first difference series of the station A includes

1961 and overlaps more than 21 years with the 30-year climatology base period 1961–1990, the anomaly of the station is directly integrated from the first differences. This 21-year criterion is the same as that required by Jones et al. (1997). If the overlap is less than 21 years, then a reference station will be used to extend the first difference series of the candidate station A. In this case, a reference station, called B, needs to be found such that

1. it overlaps with station A for at least 1 year,
2. its first difference series overlaps with 1961–1990 for at least 21 years (with 1961 inclusive), and
3. the distance between A and B is less than 200 km.

These constraints still exclude many observational stations from our analysis. For instance, if station A cannot find a reference station, then A is dropped from our analysis. Of course, one can increase the distance from 200 km to a larger value in order to avoid dropping off too many stations. Numerical experiments were performed for 150, 300, 500, and 1,000 km. It appeared that 200 km was the best value, and 300 km and 500 km yielded slightly weaker results in the sense of the accuracy of historical field reconstruction (Shen et al. 2004). Considering the e-fold correlation length, one may think that 1,000 km or larger can be used as the distance to find the reference stations, as Jones et al. (1997) concluded that the e-fold correlation length is in the range of 750–4,500 km. Although this large distance did retain almost all the stations, our numerical experiments showed that the anomaly data derived did not yield good reconstructions (Shen et al. 2004). This result may be due to the inhomogeneities within the data. For instance, some short term stations may have introduced a bias, which makes the data incompatible with the patterns of global circulation. Jones' data, on the other hand, are more homogeneous. However, in terms of global average, the results computed from the two datasets should be similar, because the global average is a very robust signal and can be easily observed with several dozen or a few hundred stations (Jones 1994; Shen et al. 1994). However, in terms of the spatial field reconstruction, there are many more degrees of freedom, and the accuracy of the data becomes much more important. When choosing reference stations, one has adopted the assumption that the first difference time series for two stations less than 200 km apart does not have a significant discontinuity when joined together, regardless of their background climatology, such as one station in a valley and another on top of a mountain.

Since Station B goes through the climatology reference period for at least 21 years, one has no problem in calculating its anomalies. Described below is our procedure for calculating anomalies of the candidate station A, which uses station B as its reference station through the climatology period. Suppose that station A goes from 1891–1944, and station B goes from 1938–1998. Then one can calculate the following first difference sequences:

$$\mathrm{A} : \delta_A (1892) , \delta_A (1893) , \cdots , \delta_A (1944) , \tag{7}$$

$$\mathrm{B} : \delta_B (1939) , \delta_B (1940) , \cdots , \delta_B (1998) . \tag{8}$$

The time series of the first difference of Station A needs to be extended from 1944 to 1990 in order to compute anomalies for the Station A. The extended time series is defined by

$$\hat{\delta}_A(1892) = \delta_A(1892), \hat{\delta}_A(1893) = \delta_A(1893), \cdots, \hat{\delta}_A(1938) = \delta_A(1938),$$
$$\hat{\delta}_A(1939) = \frac{\delta_A(1939) + \delta_B(1939)}{2}, \cdots, \hat{\delta}_A(1944) = \frac{\delta_A(1944) + \delta_B(1944)}{2},$$
$$\hat{\delta}_A(1945) = \delta_B(1945), \cdots, \hat{\delta}_A(1990) = \delta_B(1990).$$

Namely, during the overlapping period between the two stations, the extended time series is equal to the average of the time series of the two stations. Before the overlapping time, the extended time series is equal to the time series of the Station A. After the overlapping time, the extended time series is equal to the time series of the Station B.

Therefore, the time series of the real values of the temperature data of the Station A during the climatology base period 1961–1990 can be derived from its value in the first year and the first differences of the following years:

$\tilde{T}_A(1891)$ the beginning year temperature data,

$$\tilde{T}_A(1892) = \tilde{T}_A(1891) + \hat{\delta}_A(1892),$$
$$\tilde{T}_A(1893) = \tilde{T}_A(1891) + \hat{\delta}_A(1892) + \hat{\delta}_A(1893),$$
$$\cdots\cdots\cdots\cdots$$
$$\tilde{T}_A(1961) = \tilde{T}_A(1891) + \hat{\delta}_A(1892) + \cdots + \hat{\delta}_A(1961),$$
$$\tilde{T}_A(1962) = \tilde{T}_A(1891) + \hat{\delta}_A(1892) + \cdots + \hat{\delta}_A(1961) + \hat{\delta}_A(1962),$$
$$\cdots\cdots\cdots\cdots$$
$$\tilde{T}_A(1990) = \tilde{T}_A(1891) + \hat{\delta}_A(1892) + \cdots + \hat{\delta}_A(1961) + \hat{\delta}_A(1962) + \cdots + \hat{\delta}_A(1990).$$

The climatology is defined as the average of above data over the climatology reference period 1961–1990:

$$c = \frac{\tilde{T}_A(1961) + \tilde{T}(1962) + \cdots + \tilde{T}(1990)}{30}.$$

From this definition and the formulas above, the climatology can be expressed in terms of the beginning year temperature and the first differences in the following years:

$$c = \tilde{T}_A(1891) + \left[\hat{\delta}_A(1892) + \cdots + \hat{\delta}_A(1960)\right]$$
$$+ \frac{30\,\hat{\delta}_A(1961) + 29\,\hat{\delta}_A(1962) + \cdots + \hat{\delta}_A(1990)}{30}.$$

Then the anomaly time series of the Station A, still denoted by $\tilde{T}$, is equal to the temperature minus the climatology. The following formulas for calculating the anomalies can be derived:

$$\begin{aligned}
\tilde{T}_A(1891) &= -\left[\hat{\delta}_A(1892) + \hat{\delta}_A(1893) + \cdots + \hat{\delta}_A(1960)\right] \\
&\quad - \frac{30\,\hat{\delta}_A(1961) + 29\,\hat{\delta}_A(1962) + \cdots + \hat{\delta}_A(1990)}{30}, \\
\tilde{T}_A(1892) &= \tilde{T}_A(1891) + \hat{\delta}_A(1892), \\
&\quad \ldots\ldots\ldots\ldots, \\
\tilde{T}_A(1944) &= \tilde{T}_A(1943) + \hat{\delta}_A(1944).
\end{aligned}$$

Of course, if the reference station B goes through only part of the 1961–1990 period, say 1961–1982, then $\hat{\delta}_A$ needs to be extended to only 1982, and the anomaly time series should be computed by

$$\begin{aligned}
\tilde{T}_A(1891) &= -\left[\hat{\delta}_A(1892) + \hat{\delta}_A(1893) + \cdots + \hat{\delta}_A(1960)\right] \\
&\quad - \frac{22\,\hat{\delta}_A(1961) + 21\,\hat{\delta}_A(1962) + \cdots + \hat{\delta}_A(1982)}{22}, \\
\tilde{T}_A(1892) &= \tilde{T}_A(1891) + \hat{\delta}_A(1892), \\
&\quad \ldots\ldots\ldots\ldots, \\
\tilde{T}_A(1944) &= \tilde{T}_A(1943) + \hat{\delta}_A(1944).
\end{aligned}$$

If some records are missing, then the first difference series cannot be computed. The missing values are replaced by their long-term mean if the missing records are only one or two consecutive years. For example if data are missing for January of 1932, and the observation covers for the period January 1920 to December 1995, then the January mean from 1920 to 1995 except in 1932 is used to approximate for January of 1932. However, this value is used to estimate the FDS only. If any B station has its first observation made on or after January 01, 1971, then we cannot have sufficient (21 years) records in the normal period and cannot compute the climatology of this station. Therefore, we discard all such stations.

If Station A satisfies the CAM conditions, then the CAM, FDM, and RSM yield the same climatology and anomalies.

References

Balling RC (1997) Analysis of daily and monthly spatial variance components in historical temperature records. Phys Geogr 18:544–552

Balling RC, Idso CD (2002) Analysis of adjustments to the United States historical climatological network (USHCN) temperature database. Geophys Res Lett 29:25-1–25-3

Cunderlik JM, Burn DH (2003) Non-stationary pooled flood frequency analysis. J Hydrol 276:210–223

Del Genio AD, Yao M-S, Jonas J (2007) Will moist convection be stronger in a warmer climate? Geophys Res Lett 34:L16703. doi:10.1029/2007GL030525

Durre I, Menne MJ, Vose RS (2008) Strategies for evaluating quality-control procedures. J Clim Appl Meteorol 47:1785–1791
Folland CK, Rayner NA, Brown SJ, Smith TM, Shen SSP, Parker DE, Macadam I, Jones PD, Jones RN, Nicholls N, Sexton DMH (2001) Global temperature change and its uncertainties since 1861. Geophys Res Lett 28:2621–2624
Gleason BE (2002) Data documentation for data set 9101, Global Daily Climatology Network, V1.0. Technical report, National Climatic Data Center, 151 Patton Ave, Asheville, NC 28801-5001, USA
Gong DY, Ho CH (2004) Intraseasonal variability of wintertime temperature over East Asia. Int J Climatol 24:131–144
Hansen J, Ruedy R, Glascoe J, Sato M (1999) GISS analysis of surface temperature change. J Geophys Res 104:30997–31022
Hansen J, Ruedy R, Sato M, Imhoff M, Lawrence W, Easterling D, Peterson T, Karl T (2001) A closer look at United States and global surface temperature change. J Geophys Res 106:23947–23963
Huang NE, Shen SSP (2005) Hlibert-Huang transform and its applications. World Scientific Press, Singapore
IPCC (2001) Climate change 2001: the scientific basis. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, New York
IPCC (2007) Climate change 2007: the physical science basis. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
Jones PD (1994) Hemispheric surface air temperature variations: a reanalysis and an update to 1993. J Clim 7:1794–1802
Jones PD, Wigley TML, Kelly PM (1982) Variations in surface air temperature Part I, northern hemisphere 1881–1980. Mon Weath Rev 110:59–70
Jones PD, Osborn TJ, Briffa KR (1997) Estimating sampling errors in large-scale temperature averages. J Clim 10:2548–2568
Karl TR, Jones PD, Knight RW, Kukla G, Plummer N, Razuvayev V, Gallo KP, Lindseay J, Charlson RJ, Peterson TC (1993) A new perspective on recent global warming: asymmetric trends of daily maximum and minimum temperature change. Bull Am Meteor Soc 74:1007–1023
Karl TR, Knight RW, Plummer N (1995) Trends in high-frequency climate variability in the twentieth century. Nature 377:217–220
Karl TR, Easterling DR, Hughes PY, Mason EH, Lawrimore J (2000) United States Historical Climatology Network (HCN) serial temperature and precipitation data. Technical Report, National Oceanic and Atmospheric Administration National Climatic Data Center, 151 Patton Avenue, Room 120, Asheville, North Carolina 28801-5001
Katz RW, Brown BG (1992) Extreme events in a changing climate: variability is more important than averages. Clim Change 21:289–302
King'uyu SM, Gallo LA, Anima EK (2000) Recent trends of minimum and maximum surface temperature over eastern Africa. J Clim 13:2876–2886
Knutson TR, McBride JL, Chan J, Emanuel K, Holland G, Landsea C, Held I, Kossin JP, Srivastava AK, Sugi M (2010) Tropical cyclones and climate change. Nature-Geosci 3:157–163
Meehl GA 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
Menne MJ, Williams CN, Vose RS (2009) The United States Historical Climatology Network monthly temperature data—version 2. Bull Am Meteorol Soc 90:993–1007
Michaels PJ, Balling RC, Vose RS, Knappernberger PC (1998) Analysis of trends in the variability of daily and monthly historical temperature measurements. Clim Res 10:27–33
Parker DE, Jones PD, Folland CK, Bevan A (1994) Interdecadal change of surface temperature since the late nineteenth century. J Geophys Res 99:14373–14399
Peterson TC, Vose RS (1997) An overview of the Global Historical Climatology Network temperature data base. Bull Am Meteorol Soc 78:2837–2849
Peterson TC et al (1998) Homogeneity adjustments of in situ atmosphere climate data: a review. Int J Climatol 18:1493–1517
Picard D (1985) Testing and estimating change-points in time series. Adv Appl Prob 17:841–867

Price C (2009) Thunderstorms, lightning, and climate change. In: Dieter Betz H, Schumann U, Laroche P (eds) Lightning: principles, instruments and applications. Springer, New York. doi:10.1007/978-1-4020-9079-0
Shen SSP, North GR, Kim KY (1994) Spectral approach to optimal estimation of the global average temperature. J Clim 7:1999–2007
Shen SSP, Smith TM, Ropelewski CF, Livezey RE (1998) An optimal regional averaging method with error estimates and a test using tropical Pacific SST data. J Clim 11:2340–2350
Shen SSP, Dzikowski P, Li G, Griffth D (2001) Interpolation of 1961–97 daily temperature and precipitation data onto Alberta polygons of ecodistrict and soil landscapes of Canada. J Appl Meterol 40:2162–2177
Shen SSP, Basist AN, Li G, Williams C, Karl TR (2004) Prediction of sea surface temperature from the Global Historical Climatology Network data. Environmetrics 15:233–249
Shen SSP, Shu T, Huang NE, Wu Z, Karl TR, Easterling DR (2005) HHT analysis of the nonlinear and non-stationary annual cycle of daily surface air temperature data. In: Huang NE, Shen SSP (eds) Hilbert-Huang transform and its applications. World Scientific, Singapore, pp 187–210
Vincent LA, Zhang X, Bonsal BR, Hogg WD (1999) Homogenization of daily temperatures over Canada. J Clim 15:1322–1334
Vinnikov KY, Robock A (2002) Trends in moments of climatic indices. Geophys Res Lett 29:14-1–14-4
von Storch H, Zwiers FW (1999) Statistical analysis in climate research. Cambridge University Press, Cambridge
Wang X, Shen SSP (1999) Estimation of spatial degrees of freedom of a climate field. J Clim 12:1280–1291
Williams CN, Menne MJ, Vose RS, Easterling DR (2007) United States historical climatology network monthly temperature and precipitation data. ORNL/CDIAC-187, NDP-019. Carbon Dioxide Information Analysis Center. Oak Ridge National Laboratory, Oak Ridge, Tennessee
Xue Y, Smith TM, Reynolds RW (2003) Interdecadal changes of 30-yr SST normals during 1871–2000. J Clim 16:1601–1612