NOAA Satellite and Information Service
National Environmental Satellite, Data, and Information Service (NESDIS)

National Climatic Data Center
U.S. Department of Commerce

An Assessment of the Contiguous US Climate Change by Higher Statistical Moments and GDCN Data from 1901 to 2000

By

S. S. P. Shen[1], A. B. Gurung[1], H. -S. Oh[1], T. Shu[1] and D. R. Easterling[2]

[1]Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada

[2]NOAA/National Climatic Data Center, Asheville, NC 28801-5001, USA

Abstract

The Global Daily Climatology Network's (GDCN) 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 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 distribution 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 results agree with known ones and demonstrate a decrease from the 1930s to the 1960s and an increase from 1970s to present. 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 average, maximum, and minimum surface air temperatures are calculated for the periods 1901-2000, 1910-1945, 1946-1975, and 1976-2000. The results show a decreasing trend of the second- and third-order moments of all the temperatures. The fourth-order moment of the average and maximum surface air temperatures has 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 observed in the probability distribution functions. Among the changes of statistical parameters, the ones for the minimum temperature are larger than those for the maximum and mean temperatures.

Keywords: surface air temperature, daily maximum temperature, daily minimum temperature, daily mean temperature, Global Daily Climatology Network, contiguous US, anomaly, climatology, area weighted average, trend, higher statistical moments, histogram, variance, skewness, kurtosis, climate change

Table of Contents

1 Introduction

The surface air temperature (SAT) is one of the most robust indicators of short term (one-or-two-centuries-long) climate change. The global temperature is widely believed to be increasing, and the most authentic study on climate change (Houghton et al., 2001) documented a warming trend in the global temperature. However, this trend is highly non-uniform in both space and time. In space wise, while most places experienced a warming trend for a specified period, other a few areas experienced cooling. In time wise, a SAT time series is highly non-stationary with variable mean, variance, skewness, kurtosis, and probability distribution functions (PDF) in general. The linear trends of the statistical moments vary from a time period to another.

The temporal non-stationarity is attributed to both mean and extreme conditions of weather and climate. Extreme weather and climate conditions are well known to cause loss of lives and to damage economies. Understanding climate change demands the attention to changes in climate variability and extremes (Karl et al., 1993; Katz and Brown, 1992). A comprehensive understanding of what has happened and what may happen in terms of changes in weather and climate extremes can help the assessment of what these changes could mean in a variety of different human and natural contexts (Meehl et al., 2000). Very often, users of weather resources wonder if the climate is getting more or less variable, and 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 requires careful scrutiny and is, very likely, wrong, for recent studies by us and others suggest that the present climate is actually less variable than that of a century ago.

In terms of a PDF, the higher statistical moments affect the PDF's shape, while the size of the PDF tail is a measure of frequency of climate extremes. The variance of a distribution is directly related to the frequency of the extremes. The statistical parameters of variance, skewness, and kurtosis describe the rough shape of a PDF that is not too far away from being normal, and, thus, these statistical moments and the mean can describe the basic statistical behavior of a climate system of the daily mean, maximum, minimum SATs. Knowledge beyond the mean is required in the studies of climate change, particularly the changes of climate extremes. Thus, finding whether the climate is getting more or less variable is relevant to the trends in the higher moments of the statistical distribution of climatic variables. Vinnikov and Robock (2002) already 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. They used the higher power of the residual series obtained by first removing the linear trend from a time series. Parker et al. (1994) compared the variances from the seasonal global SAT anomalies for two 20-year (1954-1973 and 1974-1993) periods 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, "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 temperatures for each January and July. The results showed that the intra-monthly variance of the both minimum and maximum temperatures had decreasing trends. Houghton et al. (2001) found some evidences showing 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.

When exploring climate extremes and higher order variability in a credible fashion, the highest possible temporal resolution data are required. When using the station data and examining the long time period of several decades, the highest resolution is a day. The GDCN (Global Daily Climatology Network) of the NCDC (US National Climatic Data Center) is such as dataset. Its version 1.0 (Gleason 2002) is used here. However, most of the past studies on SAT changes used either monthly or annual climatic series. Studies on the variations of the climatic parameters in the daily resolution were very few. Although a few works have considered the higher statistical moments, their conclusions should be consolidated, compared, and verified. Our study is therefore aimed at understanding the behaviors of the higher moments and the probability distribution functions of the daily mean, maximum, and minimum SAT during different periods to detect and quantify the changes in a SAT climate system.

To ensure the correctness of our calculations, we have paid much attention to the accuracy of the first moments by comparing our results with the widely existing results. The good agreement of the results on the first moment demonstrates the reliability of our

higher statistical moments and PDF results of the daily mean, maximum, and minimum SAT. The changes of the higher moments and PDFs are thus assessed with a high level of confidence.

The rest of this report contains research method, data description, results and their analyses, conclusions and discussion, and references.

2 Methodology

2.1 Climatology and anomalies

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 (Jones et al., 1982). It is usually to use the climate mean of the period 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

$$S_i(t) = T_i(t) - c_i(m), \tag{1}$$

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 climate anomaly method (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 station 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 Equation (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 (the 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 years 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

(i) it overlaps with station A for at least one year,

(ii) its first difference series overlaps with 1961-1990 for at least 21 years (with 1961 inclusive), and

(iii) 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 1000 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:

A: $\delta_A(1892), \delta_A(1893), \cdots, \delta_A(1944)$,

B: $\delta_B(1939), \delta_B(1940), \cdots, \delta_B(1998)$.

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

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

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:

$$\begin{aligned}
&\tilde{T}_A(1891) \quad \text{the beginning year temperature data,} \\
&\tilde{T}_A(1892) = \tilde{T}_A(1891) + \hat{\delta}_A(1892), \\
&\tilde{T}_A(1893) = \tilde{T}_A(1891) + \hat{\delta}_A(1892) + \hat{\delta}_A(1893), \\
&\qquad \cdots\cdots\cdots\cdots \\
&\tilde{T}_A(1961) = \tilde{T}_A(1891) + \hat{\delta}_A(1892) + \cdots + \hat{\delta}_A(1961), \\
&\tilde{T}_A(1962) = \tilde{T}_A(1891) + \hat{\delta}_A(1892) + \cdots + \hat{\delta}_A(1961) + \hat{\delta}_A(1962), \\
&\qquad \cdots\cdots\cdots\cdots \\
&\tilde{T}_A(1990) = \tilde{T}_A(1891) + \hat{\delta}_A(1892) + \cdots + \hat{\delta}_A(1961) + \hat{\delta}_A(1962) + \cdots + \hat{\delta}_A(1990).
\end{aligned}$$

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:

$$\tilde{T}_A(1891) = -\left[\hat{\delta}_A(1892) + \hat{\delta}_A(1893) + \cdots + \hat{\delta}_A(1960)\right] - \frac{30\hat{\delta}_A(1961) + 29\hat{\delta}_A(1962) + \cdots + \hat{\delta}_A(1990)}{30},$$

$$\tilde{T}_A(1892) = \tilde{T}_A(1891) + \hat{\delta}_A(1892),$$

.....,

$$\tilde{T}_A(1944) = \tilde{T}_A(1943) + \hat{\delta}_A(1944).$$

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

$$\tilde{T}_A(1891) = -\left[\hat{\delta}_A(1892) + \hat{\delta}_A(1893) + \cdots + \hat{\delta}_A(1960)\right] - \frac{22\hat{\delta}_A(1961) + 21\hat{\delta}_A(1962) + \cdots + \hat{\delta}_A(1982)}{22},$$

$$\tilde{T}_A(1892) = \tilde{T}_A(1891) + \hat{\delta}_A(1892),$$

.....,

$$\tilde{T}_A(1944) = \tilde{T}_A(1943) + \hat{\delta}_A(1944).$$

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.

2.2 The mean and higher moments

The mean, variance, skewness and kurtosis are computed. The higher order statistics are related to the higher statistical moments and estimated with respect to the mean. It is easy to understand that a higher mean stands for an increase of the SAT. However, the climate change is not just a simple shift of the mean, rather a complex change of probability distribution shape which is reflected in the higher statistical moments.

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

$$\overline{S}_i = \frac{1}{\tau}\sum_{t=1}^{\tau} S_i(t), \qquad (2)$$

where $S_i(t)$ is the SAT anomaly at station i and time t. We purposely use this biased estimate of the first statistical moment, since the unbiased estimate is not physically meaningful when τ is small and hence the unbiased estimate of mean is quite different from the arithmetic mean.

The standard deviation is estimated as follows

$$\sigma_i = \left(\frac{1}{\tau} \sum_{t=1}^{\tau} \left[S_i(t) - \overline{S}_i \right]^2 \right)^{1/2}. \quad (3)$$

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, which imply an increase of extreme temperature events on both the warm and cold (higher and lower) sides or, at least, on one side.

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

$$M_i^k = \frac{1}{\tau} \sum_{t=1}^{\tau} \left[\frac{S_i(t) - \overline{S}_i}{\sigma_i} \right]^k, \quad (4)$$

where M_i^3 is here defined as the skewness and M_i^4 is the kurtosis.

Positive skewness means that a distribution has its asymmetric tail extending out toward the right (implying 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. An asymmetric tail extending out toward the left implies more extreme cold events, and 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.,

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

The leptokurtic distribution implies a sharp and lesser dispersion of the measurement, while the platykurtic indicates a flattened or wider dispersion of the measurement and a greater probability of both extremes.

In the present context, we assign M_i^1 and M_i^2 to be $\bar{S}_i$ and σ_i^2, respectively. Finally, we assign a cosine latitude weight to all the grid values to obtain the area-weighted average moment $\bar{M}^k$ by

$$\bar{M}^k = \frac{\sum_{i=1}^{N} M_i^k \cos(\phi_i)}{\sum_{i=1}^{N} \cos(\phi_i)}, \tag{5}$$

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

2.3 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. There is no proved reliable method to fill the missing data, although various kinds of filling methods have been proposed in statistics and used in practice. Cautious climatologists often choose to retain the most reliable station data when assessing climate changes in terms of monthly data or the data

of longer time scales. 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, Hansen et al. (1999) used only the station monthly 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 processes like the station data include in the GDCN dataset is useful in retaining the spatial and temporal variance. Thus, an effort should be made to utilize all the station data. Our method is to interpolate the station 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 objective-analysis-type of optimal interpolations, which assumes stationary statistics, is not applicable here due to the strong non-stationarity of daily climate processes (Shen et al., 2001).

The interpolation method is as follows. Create the 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 during the studying period. For each grid point, a distance table is created in the ascending order of distances between the grid point and 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 those of 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 this happens, the amplification of the variance is negligible in the case of contiguous US's area-weight 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 **Data**

The essential data used in this study are from the dataset of the Global Daily Climatology Network (GDCN) v1.0 (Gleason, 2002), compiled by the United States NCDC. The global data set contains the daily maximum and minimum temperatures and daily total precipitation data of 32,857 stations around the world within the time period 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 the statistical analysis of sets of observations, to locate and identify potential outliers and/or erroneous data (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 GDCN source data were mostly collected by the National Climatic Data Center scientists through cooperative research exchanges. In addition, 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 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. The user can then make the decision to include the datum based upon their particular application. For our study of higher moments, we intend to use all the data that, of course, will include some false outliers, and also falsely smoothed data. We thus regard the current GDCN data actually reflect the real climate variability at daily scales in terms of higher statistical moments and PDFs.

Other data are also used to verify our results by comparing our results with others when applicable. The comparison is mainly made for the 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.

A) The raw United States Historical Climatological Network's (USHCN) data (Karl et al., 1990 and 2000), denoted by RAW. The USHCN contains the raw, time-of-observation-bias (TOB) corrected data. The raw data, also known as areal edited data, are the original data. The TOB indicates that the time-of-

observation biases have been removed so that the data are consistent in reflecting a midnight-to-midnight temperature variation.

B) The FILNET (Fill Missing Original Data in the Network) annual mean temperature time series from 1930 to 2000 (Karl et al., 2000). The FILNET data are the time-of-observation data that have been adjusted for the maximum/minimum temperature system bias, and the station moves/changes bias and that contain the estimated values for the missing/outlier data. The FILNET contains the maximum and minimum SATs of 1,221 USHCN stations.

C) The urban adjusted data, which adjusted the urbanization effects among the FILNET stations by regression (Karl et al., 2000), denoted by URB-ADJ.

D) The IPCC (Intergovernmental Panel for Climate Change) temperature series originally developed by Jones (1994).

4 Results and Their Analysis

The GDCN dataset includes both SAT and precipitation. Only the daily data of maximum and minimum SAT from January 1, 1901 to December 31, 2000 are extracted from the dataset and reformatted. Each station corresponds to a data file and the 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 corresponds to 2451910. The daily mean temperature is defined as the mean 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 hours temperature of a day, which should be defined by an integration of the temperature function of time.

Calculation of the mean of the maximum and minimum SAT requires the existence of both records of the maximum and minimum SAT. Almost all the stations recorded both maximum and minimum temperatures, but there were some exceptions. The contiguous US during January 1, 1901-December 31, 2000 had 8,620 stations with records of both 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.

4.1 Monthly temperature anomalies and their spatial average

We will first examine the first statistical moment, i.e., the ensemble mean. Since many results are available for comparison for the ensemble mean, we can develop confidence of our other results based on the good agreement on the first moment. The main issue about the first moment is the use of datasets. Various datasets exist for the calculations of ensemble mean in the monthly and annual time scales. The four datasets listed in Section 3 are examples. The common characteristics of the four datasets and the other datasets of similar type are high quality of the data. They selected only the long-term and reliable stations into their network and excluded those stations that may cause various errors. The monthly and annual mean temperatures have a large spatial scale and hence do not need many stations to get a correct spatial average of the first moment. 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 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.

In our present analysis, because of the need to calculate higher moments and to pay attention to variances and climate extremes, we have retained as many stations as possible. Retaining a large number of stations will certainly introduce some stations of low quality. Whether these low quality stations will affect the first moment is important to the credibility of our higher moments, although there are no better practical ways available of computing higher moments and their spatial average besides our present method. Fortunately, our first-moment results calculated from the 8,620 stations agree with the results from the networks of fewer but higher quality stations.

Our comparison procedure is below. Our stations are classified into seven subsets based on different conditions. The first six subsets are 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 station subsets 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. Climatology and anomalies are computed for each station by using the anomaly calculation method reported in Section 2. Despite the flexibility of the method, the conditions for computing the climatology still exclude around 2400 stations in the analysis. It thus leaves only 6196, 6236, and 6234 stations to calculate the climatology, anomalies, and the statistical moments of the anomalies of mean, maximum, and minimum temperatures, respectively.

4.1.1 Simple average

From the station monthly temperature anomalies, we compute the simple arithmetic average of the mean, maximum, and minimum temperatures of all the stations. The result

is the contiguous US average monthly mean anomalies of the daily mean, maximum, and minimum SAT. The 12-month mean anomalies are then averaged into the annual anomalies, i.e., the contiguous US average annual mean anomalies. The results for the annual anomalies are shown in Figures 1, 2 and 3 for the daily mean, maximum, and minimum SATs, respectively. Each figure shows seven time series according to the seven subsets of data stated above. Although all the seven time series agree well with each other and the linear trends are about the same, differences still exists among the data sets G50, …, G100 of different minimum record lengths (50, 60, 70, 80, 90, and 100 years), and all the stations' data GALL. The differences are analyzed.

The data with longer record lengths are of better quality, but do not have a good spatial coverage. The data of shorter record lengths contain more noise, but have better spatial coverage. Despite the general agreement of the results, particularly in terms of trend, of the seven different GDCN sub-networks in Figures 1-3, it is thus needed to examine carefully the differences among these results and debate the balance of the data quality and spatial coverage. The differences are computed between the US average annual mean temperatures corresponding to the sub-network G50 of minimum record length of 50 years and other GDCN sub-networks. The differences are shown in Figures 4, 5 and 6 for the mean, maximum, and minimum daily SATs, respectively.

Two features are noticeable. First, the differences show a slight positive bias toward positive side before the 1950s and a slight negative bias from the 1950s. These might be due to the dramatic increase of the number of stations right after the World War II (see Figure 7). Second, the peaks are larger for the data of longer minimum record lengths; i.e., higher values for both the positive and negative peaks are observed while

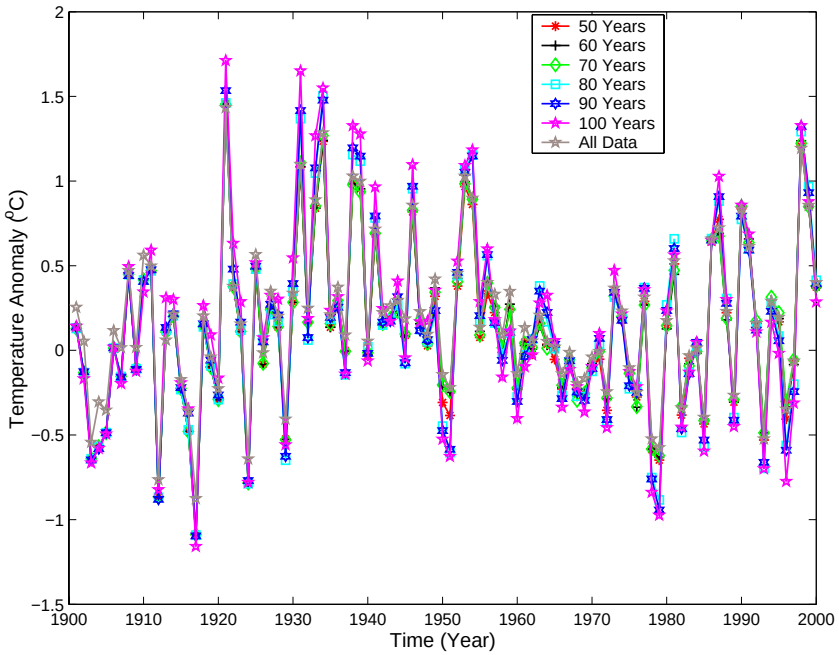

Figure 1. The US average annual mean anomalies of the daily mean SAT computed from the GDCN sub-networks of stations with different minimum record lengths. The US average monthly anomalies are obtained by the simple average from stations. The US annual mean is calculated from the US average monthly anomalies.

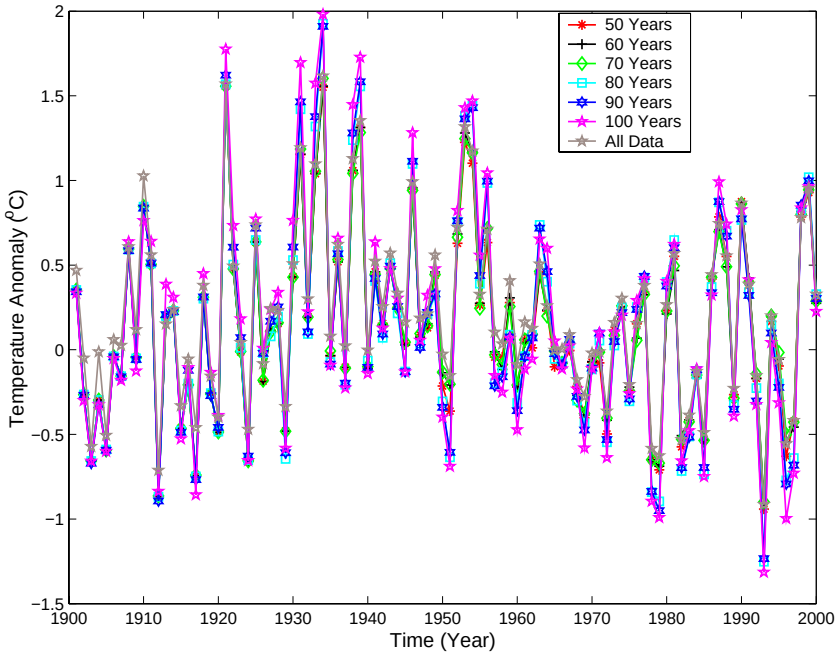

Figure 2. The US average annual mean anomalies of the daily maximum SAT computed from the GDCN sub-networks of stations with different minimum record lengths. The US averages are obtained by the simple average method from stations.

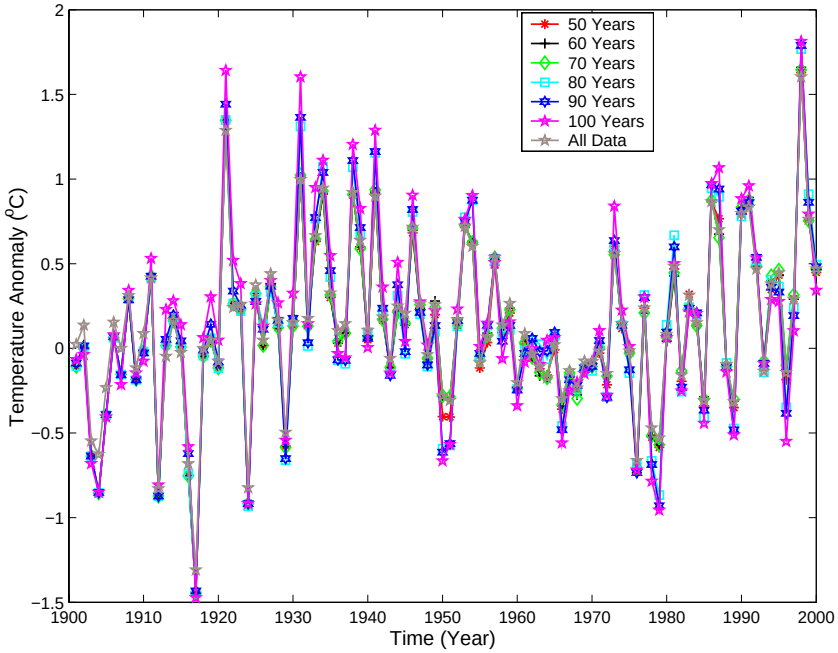

Figure 3. The US average annual mean anomalies of the daily minimum SAT computed from the GDCN sub-networks of stations with different minimum record lengths. The US averages are obtained by the simple average method from stations.

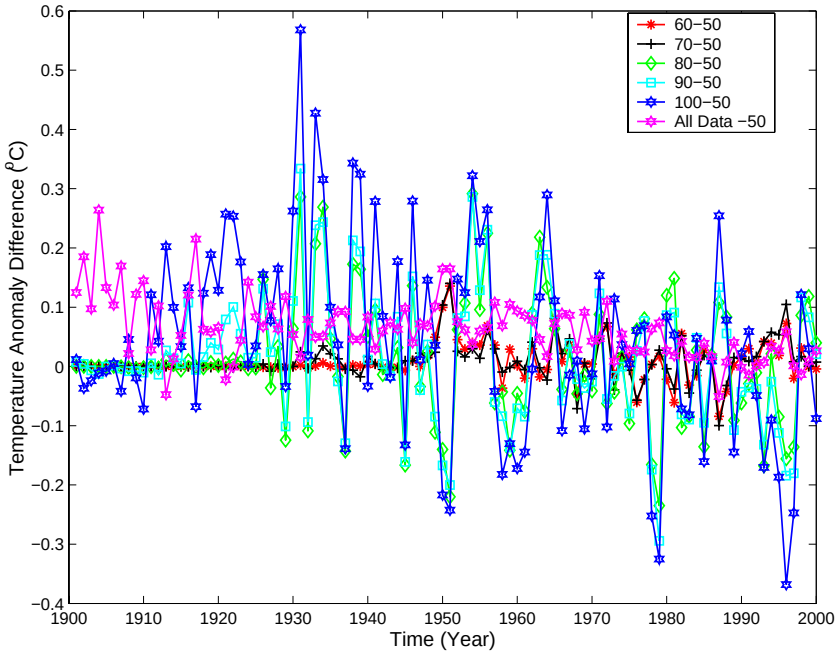

Figure 4. Differences of the US average annual mean of the daily mean SAT computed from the GDCN sub-network of a minimum 50 years record length and other GDCN sub-networks of different minimum record lengths.

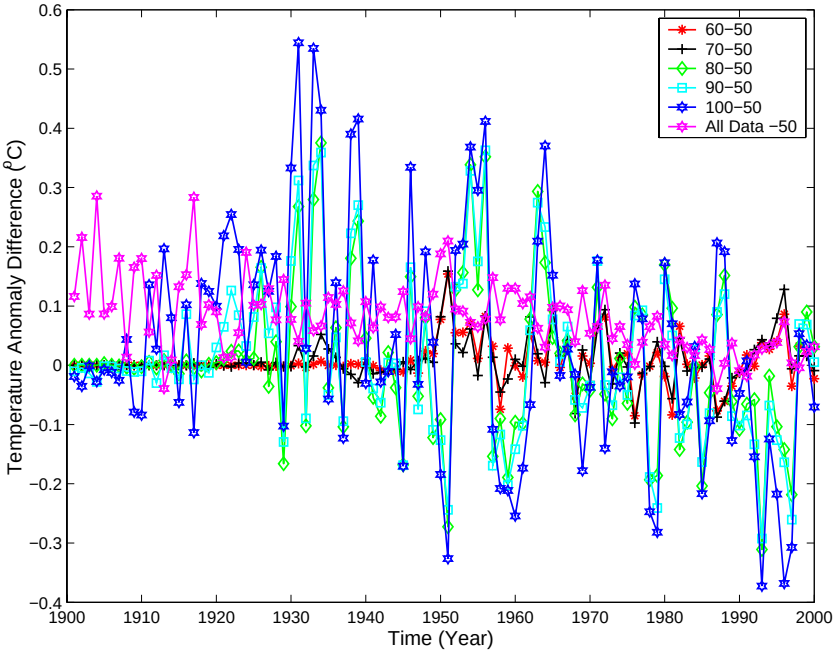

Figure 5. Differences of the US average annual mean of the daily maximum SAT computed from the GDCN sub-network of a minimum 50 years record length and other GDCN sub-networks of different minimum record lengths.

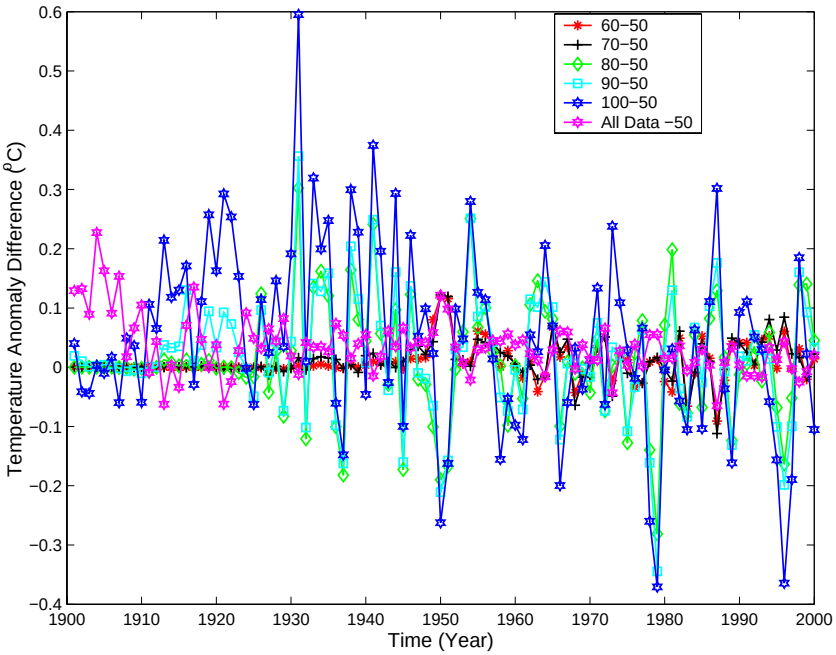

Figure 6. Differences of the US average annual mean of the daily minimum SAT computed from the GDCN sub-network of a minimum 50 years record length and other GDCN sub-networks of different minimum record lengths.

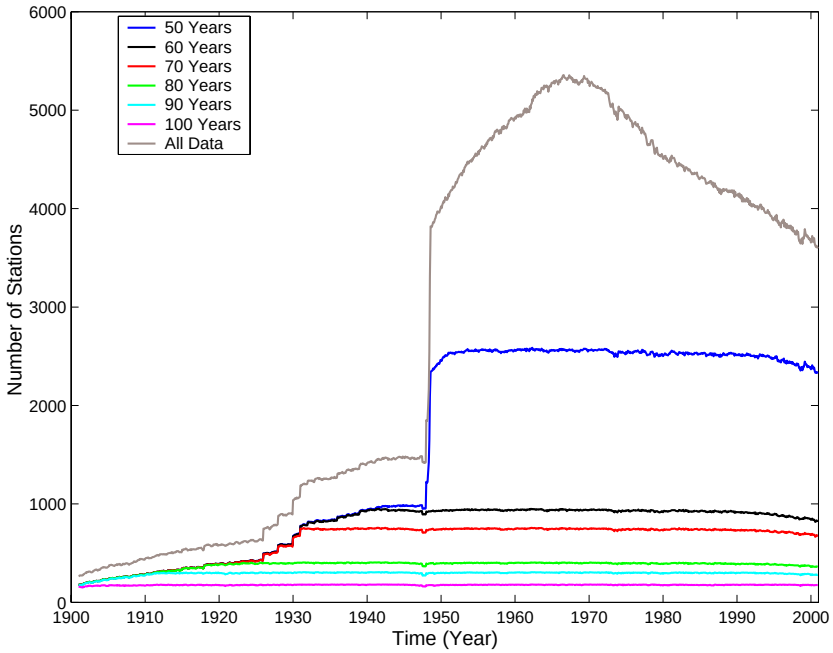

Figure 7. Number of stations with respect to time for the different record lengths in the GDCN dataset.

considering the data sets with minimum record lengths of 100, 90, and 80 years of records, as compared to those of 50 and 60 years. Figure 7 shows the number of stations versus time. With different minimum length of records, different number of stations was available at a given year. Of course, the dataset that requires the longest minimum record length, i.e., 100 years, has the fewest stations. The simple average tends to smooth out the temporal peaks if the peaks of all the stations are not synchronized. The climate inhomogeneity of the US implies the frequent occurrence of the non-synchronized peaks, e.g., a very cold anomaly in the northeast US happens with a very warm anomaly in the southwest US. Thus, the fewer the number of stations, the more pronounced the peaks from simple spatial average. Statistically speaking for a white noise field, variance is inversely proportional to the square root of sample size. The more stations (corresponding to shorter minimum record lengths), the smaller the variances of the simple average.

Of course, it can happen that the high quality stations of minimum record length equal to 100 years are located at places of low climate variability. In that case, these high quality stations will lead to lower variance and hence smaller peaks. However, this seems not the situation in the US climate observation history.

4.1.2 Area-weighted average

The simple average has the advantage of being simple, but the representative area of a station is not reflected. Some stations are necessarily redundant in the station-dense regions, and hence these stations should be weighted less in the spatial average. In the station sparse regions, there are few or no redundant stations, thus these stations should be weighted more. Our next step is to compute the area-weighted average anomalies of different temperatures. Our approach is equivalent to the Thiessen polygon method but is implemented by the fine-grid-method suggested by Shen et al. (2001).

We interpolate the monthly anomalies (as used for simple average) of each station to the grids of 0.1° latitude × 0.1° longitude, 0.5° latitude × 0.5° longitude, and 1.0° latitude × 1.0° longitude resolutions by the nearest-station-assignment method. Grid point masks are made to approximately cover the contiguous US for all the three resolutions (Figures 8 to 10). Since the areas for the uniform latitude and longitude boxes are linearly proportional to the cosine of the latitude, the cosine of the latitude of a grid box's centroid is assigned to the grid box' weight. The corresponding mathematical formula is Equation (5) in Sub-section 2.2.

The results obtained by using the area-weighted average technique with the 0.5° latitude × 0.5° longitude grid with different minimum record lengths are shown in Figures 11-13 for the US average annual mean anomalies of the daily mean, maximum,

and minimum SATs, respectively. The differences between the annual average temperatures corresponding to the minimum record length of 50 years and others are also computed, and the results are shown in Figures 14-16 for the daily mean, maximum, and minimum temperatures, respectively. Compared to the results obtained by using the simple average technique, the differences in the area-weighted results from the different data sets are smaller. Comparing Figures 4-6 and 14-16, the clear transition from positive biases to negative ones around the late 1940s and the 1950s in the simple average do not exist in the area-weighted averages.

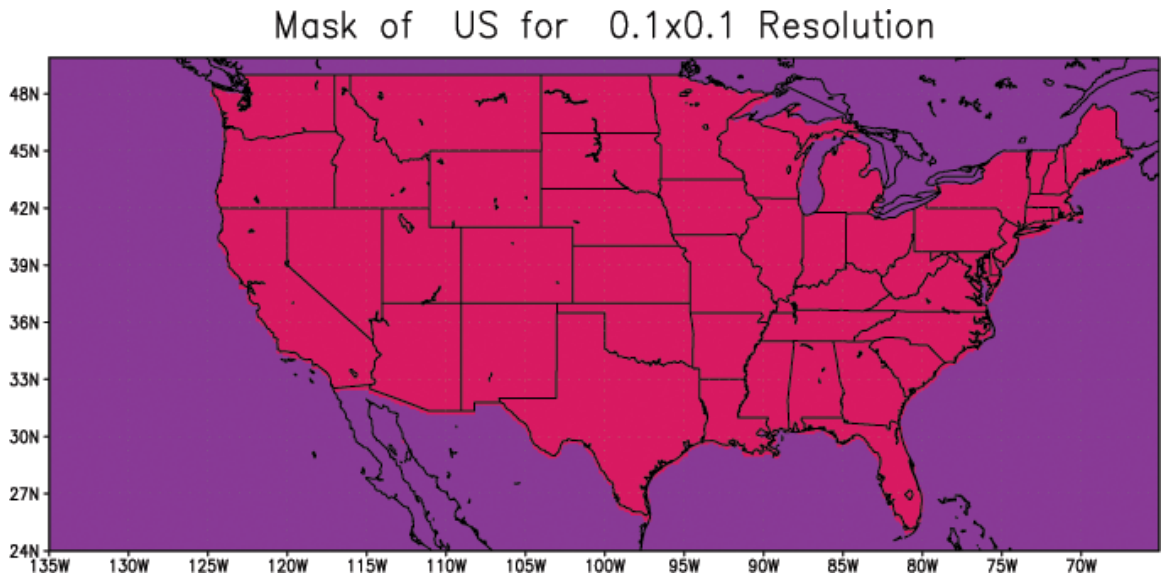

Figure 8. Mask of the contiguous US with a 0.1° latitude x 0.1° longitude grid.

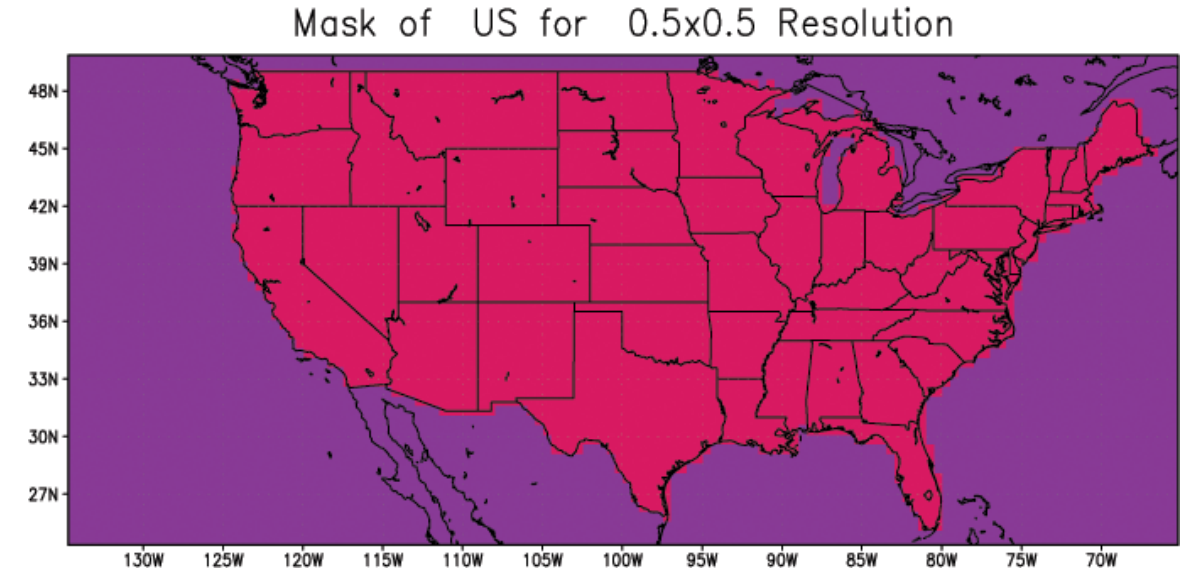

Figure 9. Mask of the contiguous US with a 0.5° latitude x 0.5° longitude grid.

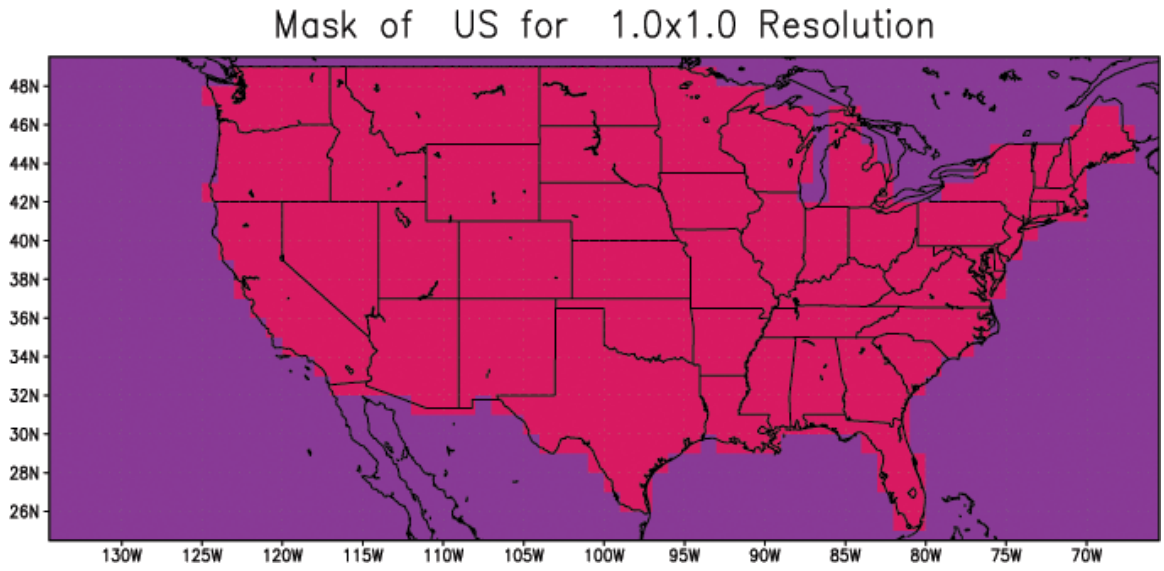

Figure 10. Mask of the contiguous US with 1.0° latitude x 1.0° longitude grid.

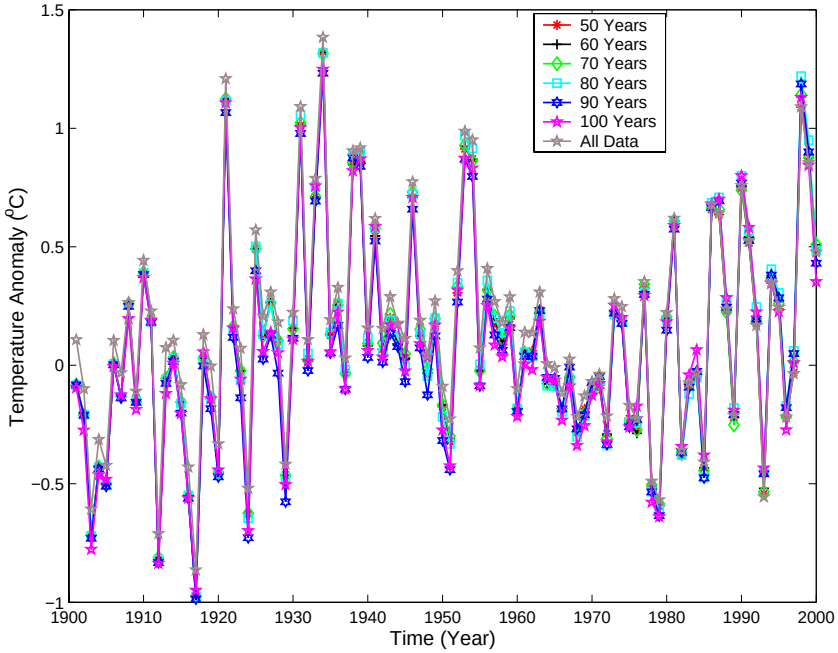

Figure 11. The US average annual mean of the daily mean temperature from the GDCN's data subsets with different minimum record lengths by using the area-weighted average method.

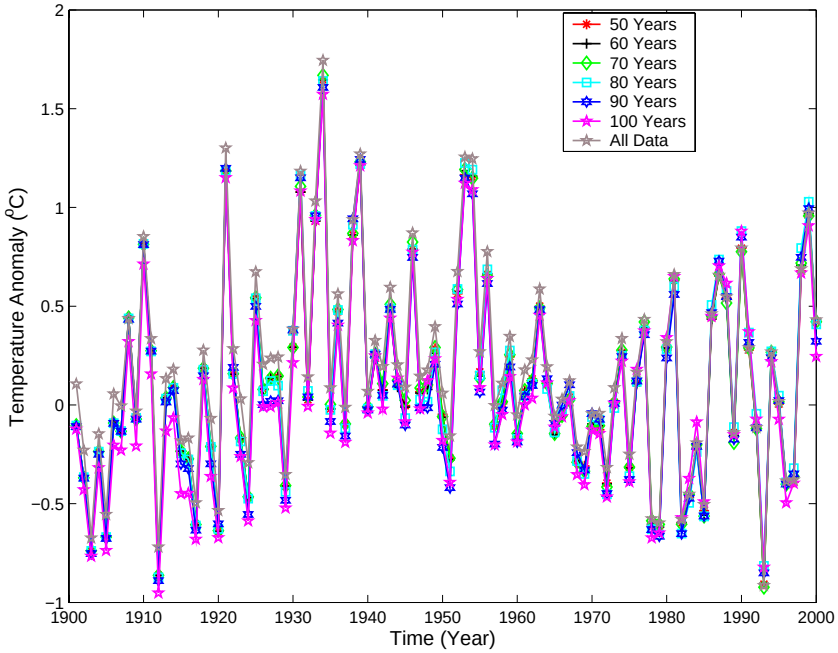

Figure 12. The US average annual mean of the daily maximum temperature from the GDCN's data subsets with different minimum record lengths by using the area-weighted average method.

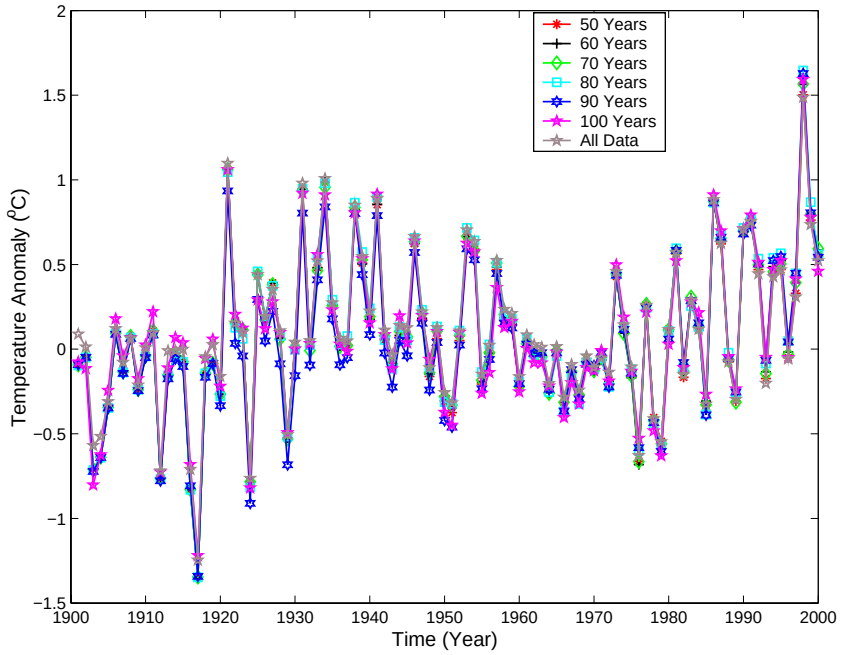

Figure 13. The US average annual mean of the daily minimum temperature from the GDCN's data subsets with different minimum record lengths by using the area-weighted average method.

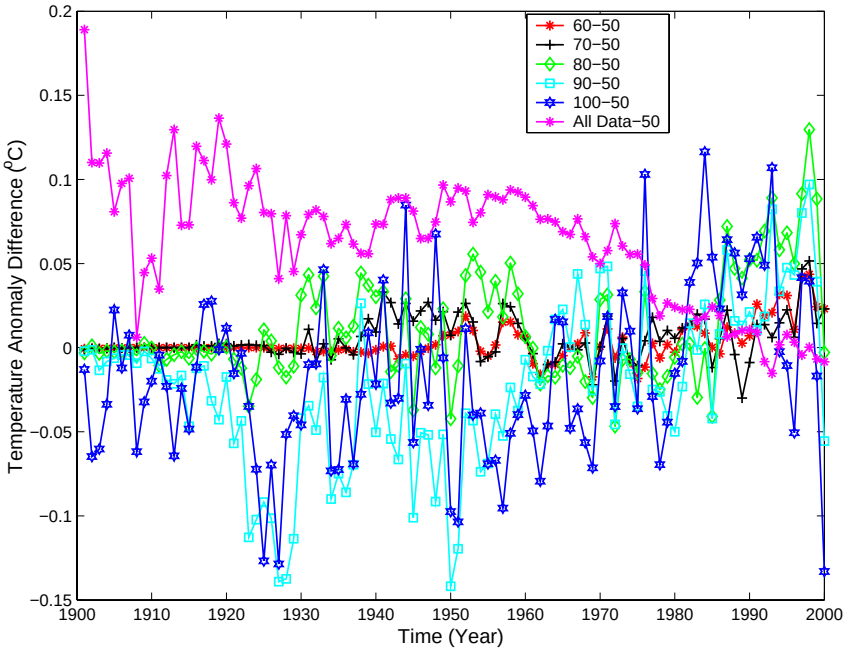

Figure 14. Differences between the US average annual mean of the daily mean temperature computed from the GDCN's data of minimum record length of 50 years and the same temperature computed from the six other GDCN sub-datasets.

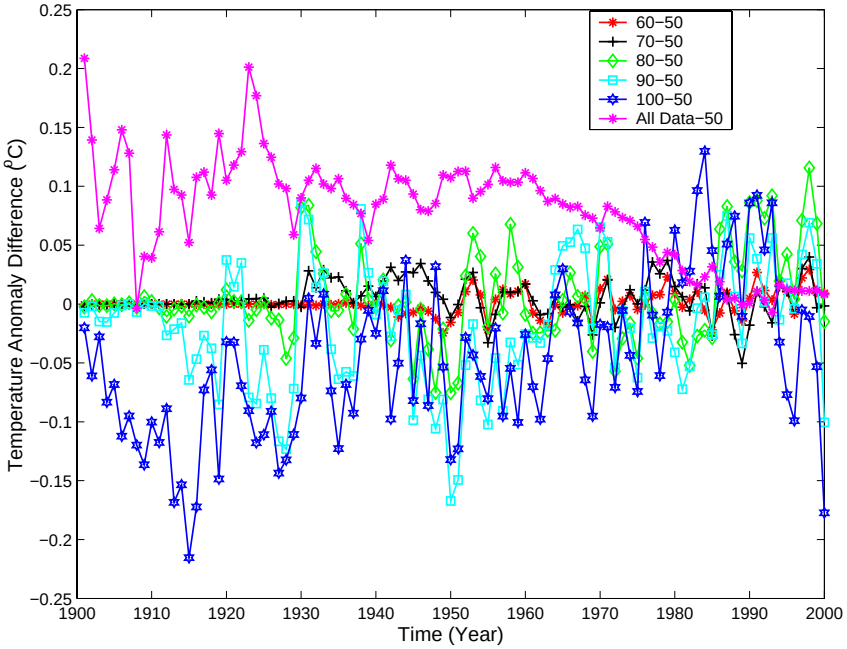

Figure 15. Differences between the US average annual mean of the daily maximum temperature computed from the GDCN's data of minimum record length of 50 years and the same temperature computed from the six other GDCN sub-datasets.

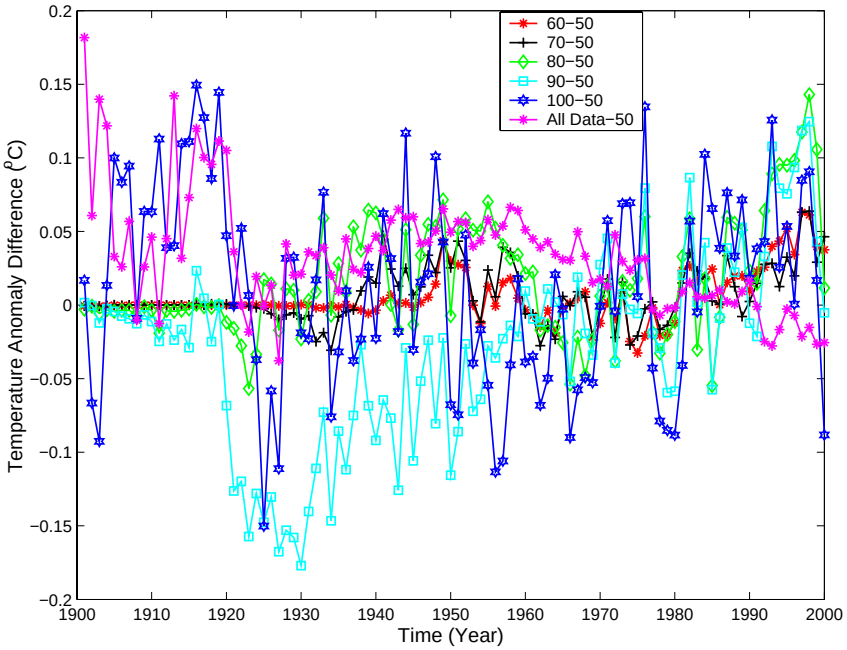

Figure 16. Differences between the US average annual mean of the daily minimum temperature computed from the GDCN's data of minimum record length of 50 years and the same temperature computed from the six other GDCN sub-datasets.

Figures 14-16 are the plots of the differences between the US average annual mean of the daily SAT computed from the GDCN's data of minimum record length of 50 years and the same temperature computed from the six other GDCN sub-datasets of different minimum record lengths. While most differences are centered around zero particularly after the 1970s, some differences show apparent trends. For example, most of the differences between the results from the data of minimum record lengths of 50 and 100 years were negatives before 1970, but in the same period, the difference between the results obtained from the dataset of the minimum record lengths of 50 years and that of all stations were positive. The apparent trend in the differences compels a probe into the causes. The results are supposed to represent spatial averages. Two possible causes for the difference in trends: the first is the quality of data, and the other is the distribution of the stations. Since the GDCN data have gone through the quality assessment and quality control procedures, the biases in the station data toward one side for the sub-networks of different minimum record lengths appear unlikely. Our focus is thus on the second cause: station distribution. Figure 17 shows the station distribution of minimum record length of 50 years, and Figure 18 that of 100 years. Figure 18 includes many fewer stations than Figure 17 due to the strict 100-year length requirement on stations. Only 188 stations satisfy this requirement, while 2,713 stations satisfy the 50-year requirement. The corresponding numbers of stations for the minimum record of 60, 70, 80 and 90 years are 1006, 785, 412, and 311, respectively. See Figure 7 again for the variations of the number of stations of different minimum record lengths. Apparently, the sub-network of the stations of minimum record length of 100 years leaves many states in the west and east coasts uncovered, plus several station-void states in the middle. It is a reasonable

inference that the bias is caused by the incomplete cover of the stations of minimum record lengths of 100 years. To further justify this inference, we create an auxiliary network from the sub-network of stations of minimum record length of 50 years, so that this auxiliary network has the same spatial coverage. If a $0.5° \times 0.5°$ grid box contains no station of minimum record length of 100 years, then the stations of minimum record length of 50 years are deleted. Thus, the leftover stations from the sub-network of minimum network of 50 years forms the auxiliary network and is referred to as the network of "50 Year Partial Data" in Figure 19. This auxiliary network and the sub-network of minimum record length of 100 years have the same spatial coverage. Figure 19 shows that the difference between these two networks of the same spatial coverage are centered around zero and has no bias toward one side.

Therefore, we can conclude biases in the spatial average of annual or monthly means of temperatures are caused by the incomplete spatial coverage, if the arithmetic or area-weighted averaging methods are used. However, if optimal averaging method with consideration of climate inhomogeneities, the spatial coverage may become a less serious problem. The inhomogeneous patterns may be reflected by empirical orthogonal functions (EOF) and optimal weights can be calculated based upon the EOFs. See Shen et al. (1994 and 1998) and Folland et al. (2001) for the method.

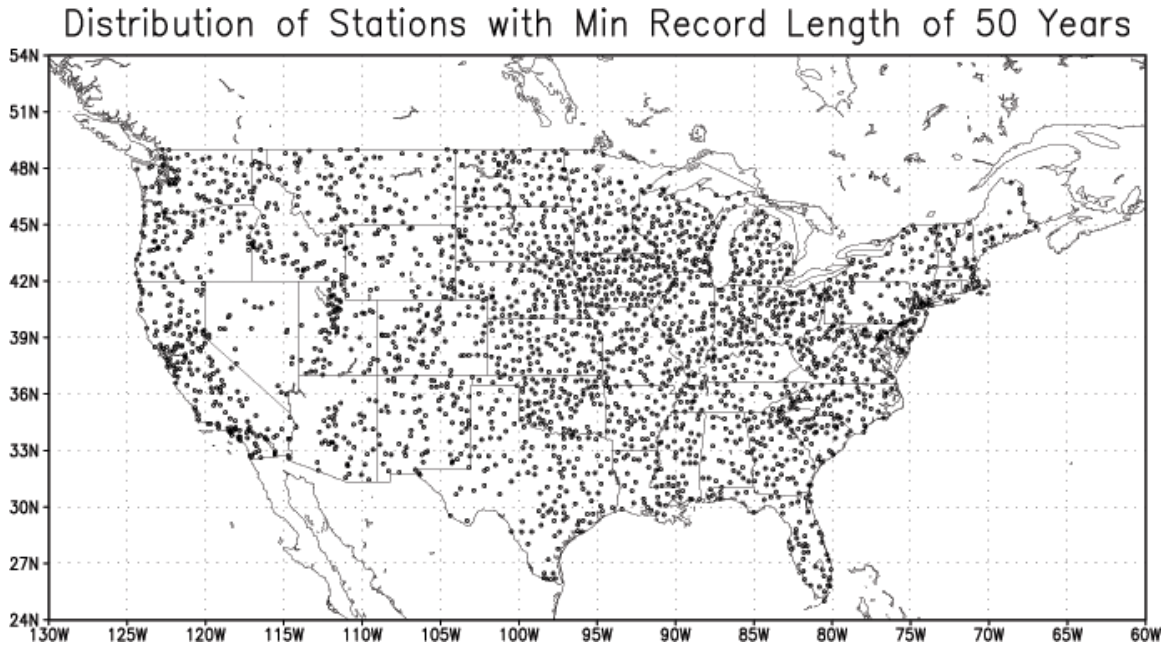

Figure 17. Distribution of stations corresponding to a minimum record length of 50 years.

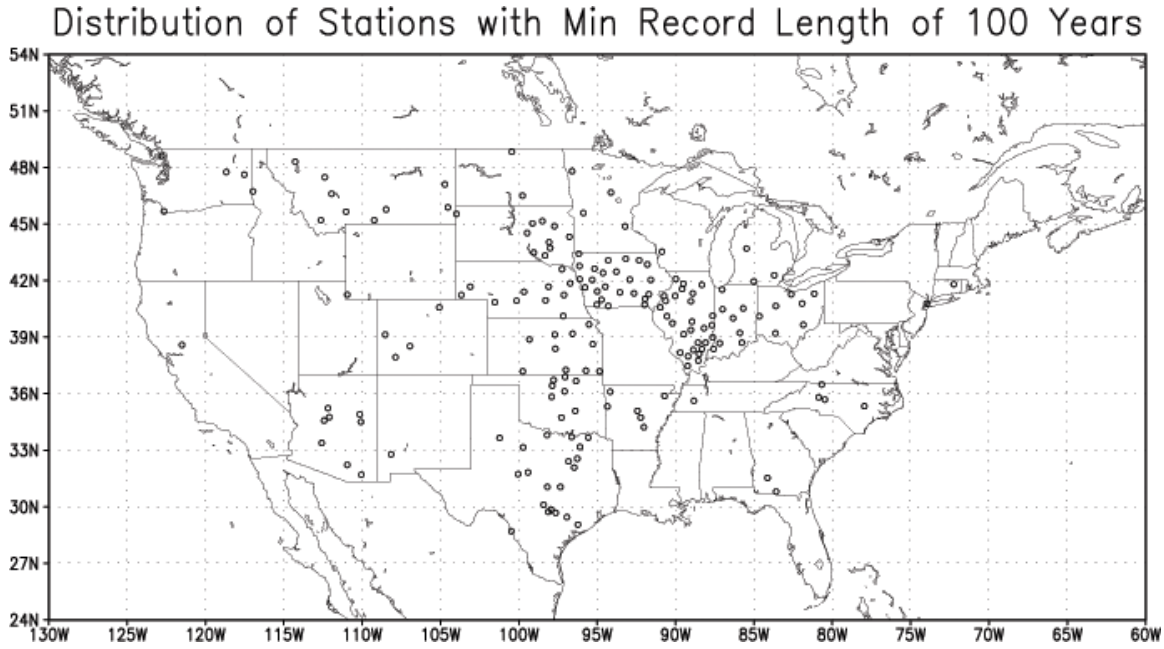

Figure 18. Distribution of stations corresponding to a minimum record length of 100 years.

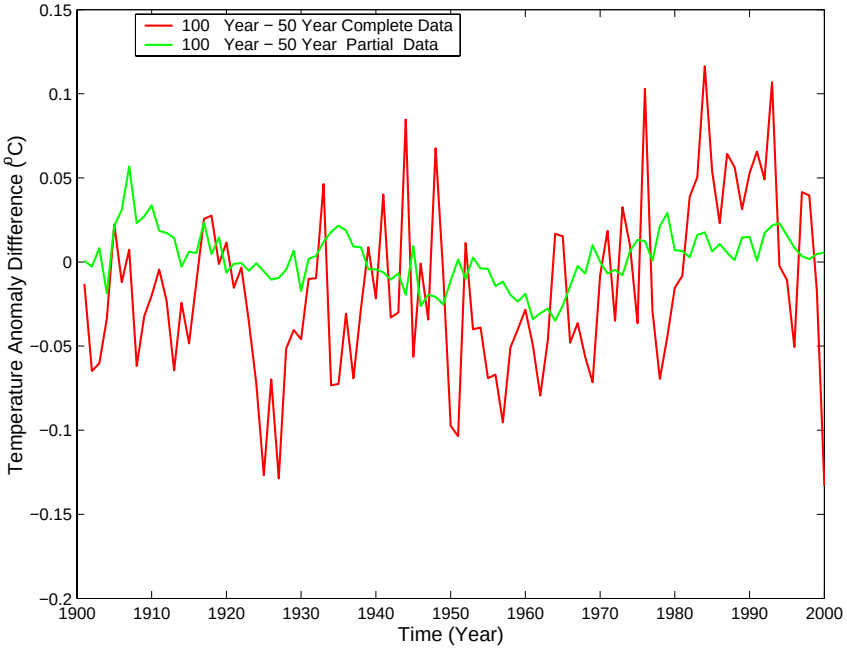

Figure 19. Difference of the annual mean of the average temperatures computed from a minimum record length of 100 years with complete and partial data sets with a minimum record length of 50 years.

4.2 Daily temperature anomalies and their annual means

The spatial and temporal variances are inversely proportional to the time scale. The shorter the time scale, the larger the variances. 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 stress index in environmental health, and heating degree days in utility management. If climate variances have changed in the last century, they should have been shown more likely in the daily data than in the monthly or annual data.

Similar to the monthly data analysis above, the daily data analysis is also carried out for anomalies. The thirty-year-mean climatology for daily data is not stable and varies from one thirty-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 thirty-year-mean monthly climatology according to Huang and Shen (2005), but not to use the nonlinear climatology for comparison reasons. Thus, the daily anomaly data are generated by subtracting the daily SAT from the corresponding monthly climatology defined in the period of 1961-1990. Our analysis of the daily anomalies for the US will be through simple and area-weighted averages. For comparison reason, we first investigate the annual mean from the daily anomalies (this sub-section 4.2). The good agreement with the existing results will help us probe into the higher statistical moments (See sub-section 4.4). When making the simple average, the station's annual mean requires a minimum of 300 days of anomaly data. This number 300 is of course an ad hoc value and can be adjusted if appropriate reasons exist. For the area-weighted average, the station anomalies are gridded onto a regular grid and the annual mean is calculated for the gridded data that has 365 or 366 values a year.

4.2.1 Simple averages

The simple average of the annual mean of the daily SAT anomalies are shown in Figures 20, 21, and 22 for the daily mean, maximum, and minimum SATs, respectively. Again, the averages from seven sub-networks are shown. Like the peaks of the monthly anomalies, the peaks of the longer record lengths (100 years of minimum record length is the longest) for the daily anomalies are more pronounced than those of the shorter ones.

The differences between the annual mean from the sub-network of the minimum record length of 50 years and other sub-networks are shown in Figures 23, 24, and 25 for the daily mean, maximum, and minimum SATs, respectively.

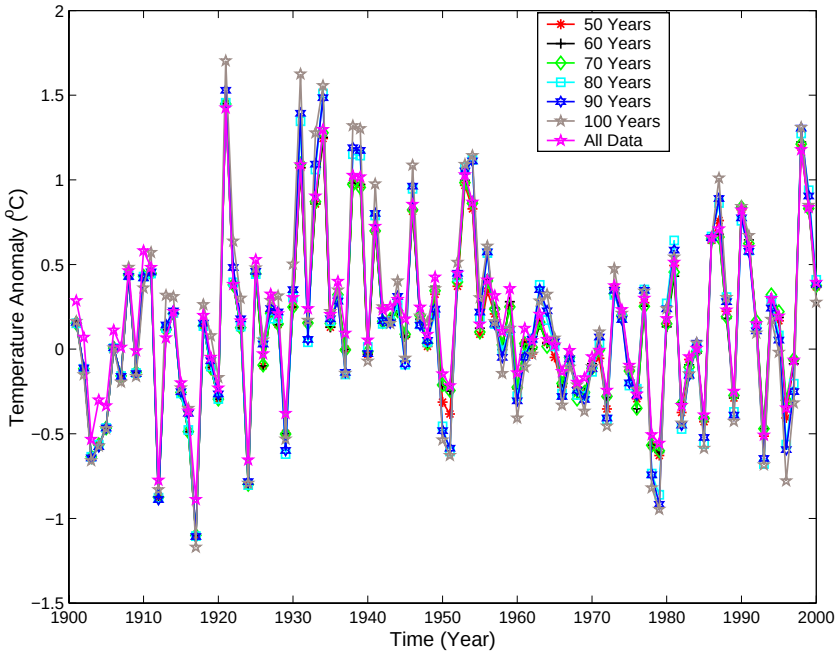

Figure 20. The US average annual mean of the daily mean temperature computed from the GDCN's data sets with different minimum record lengths by using the simple average method. The annual mean here is computed directly from the daily anomalies.

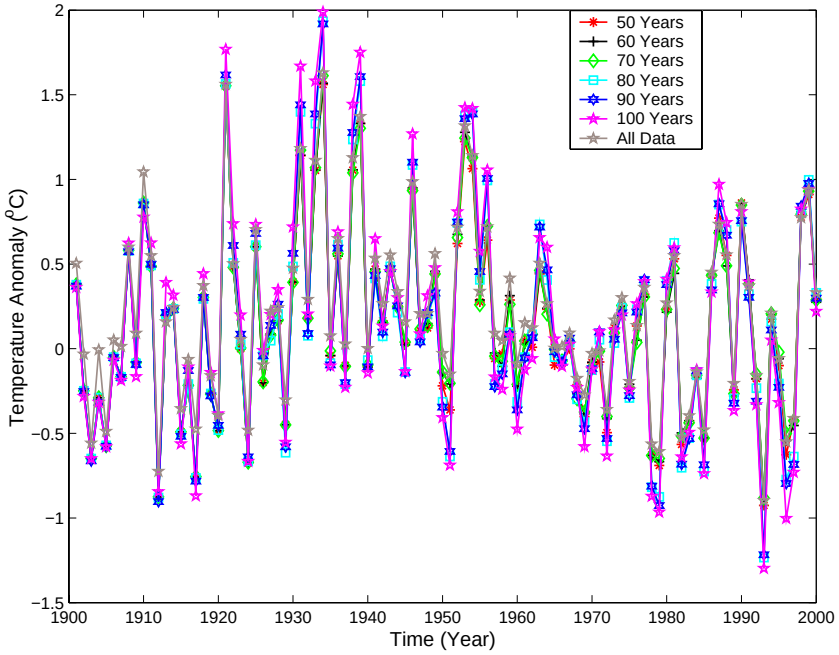

Figure 21. The US average annual mean of the daily maximum temperature computed from the GDCN's data sets with different minimum record lengths by using the simple average method. The annual mean here is computed directly from the daily anomalies.

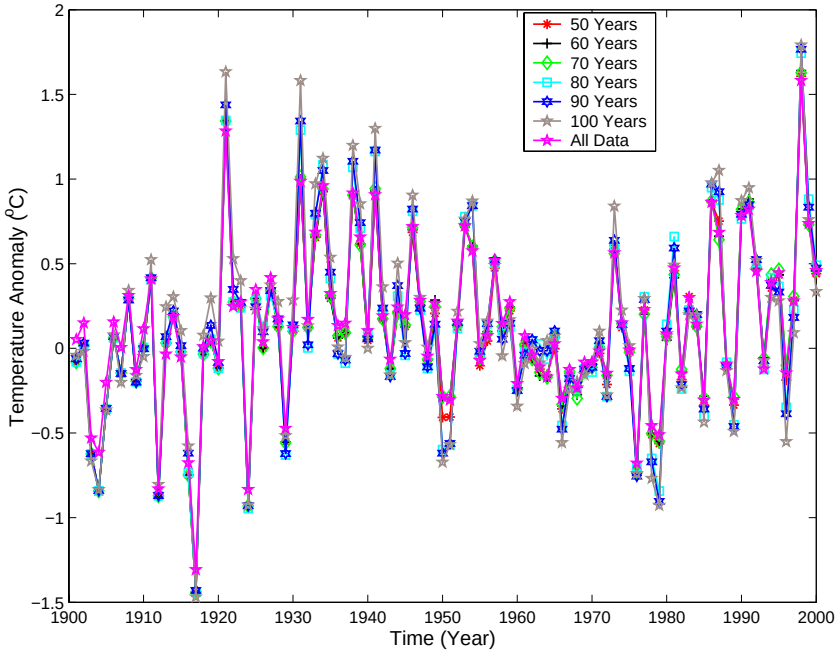

Figure 22. The US average annual mean of the daily minimum temperature computed from the GDCN's data sets with different minimum record lengths by using the simple average method. The annual mean here is computed directly from the daily anomalies.

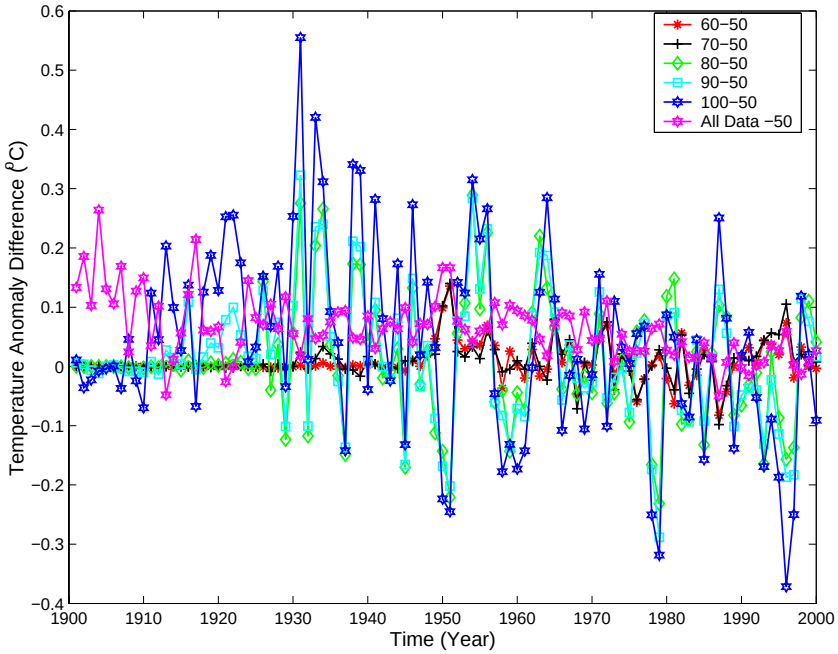

Figure 23. Differences of the US average annual mean of the daily mean temperature computed from the dataset of a minimum 50 years record length and others by using the simple average method. The annual mean here is computed directly from the daily anomalies.

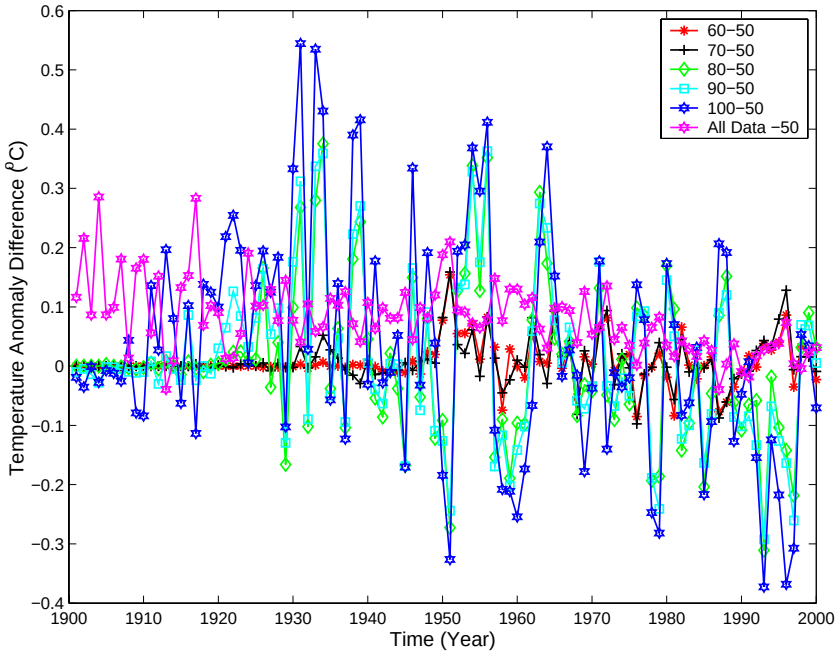

Figure 24. Differences of the US average annual maximum of the daily mean temperature computed from the dataset of a minimum 50 years record length and others by using the simple average method. The annual mean here is computed directly from the daily anomalies.

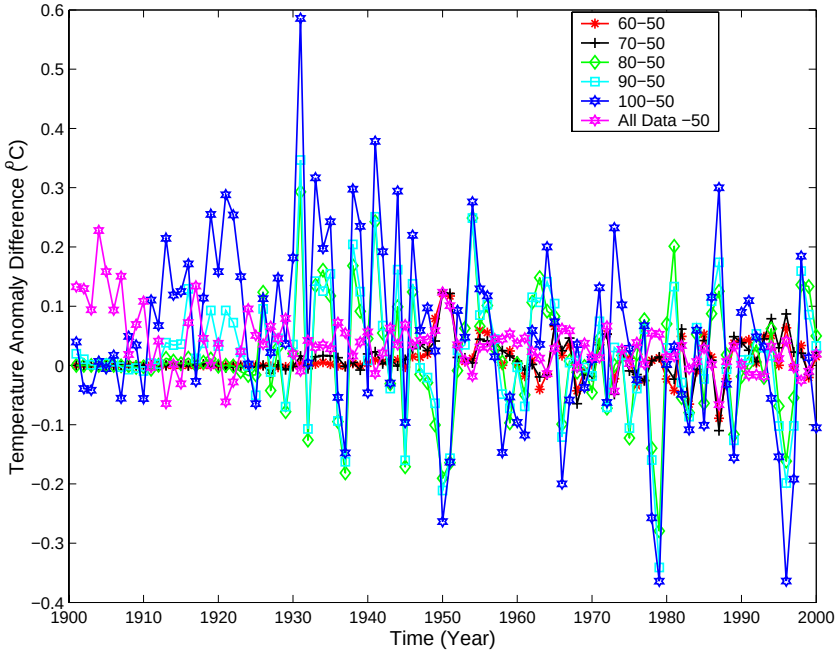

Figure 25. Differences of the US average annual mean of the daily minumum temperature computed from the dataset of a minimum 50 years record length and others by using the simple average method. The annual mean here is computed directly from the daily anomalies.

Area-weighted averages

The area-weighted average for the contiguous US is calculated by using dense grids as has been done for the monthly anomalies. Again, the grids of 0.1° latitude × 0.1° longitude, 0.5° latitude × 0.5° longitude, and 1.0° latitude × 1.0° longitude are tested. The results from these three grids are almost the same. The ones corresponding to the 0.5° latitude × 0.5° longitude grid with the GDCN data from the seven different sub-networks of minimum record lengths are shown in Figures 26, 27, and 28, which are for the daily mean, maximum, and minimum SATs, respectively. The fluctuations and trends are in good agreement with the known results (see Figures 32-34 for the comparison).

The differences between the annual mean SATs corresponding to the sub-network of minimum record length of 50 years and those to other six sub-networks are shown in Figures 29, 30, and 31 for the daily mean, maximum, and minimum SATs, respectively. Compared to the results obtained by using the simple average (Figures 23-25), the differences among the area-weighted average results are smaller. Most of the differences are within the range of (-0.15, 0.15), while the corresponding range for the simple average is (-0.30, 0.30), twice as large. The reason for this consequence is the non-uniform station coverage. The area-weighted average can handle the non-uniform station coverage better than the simple average.

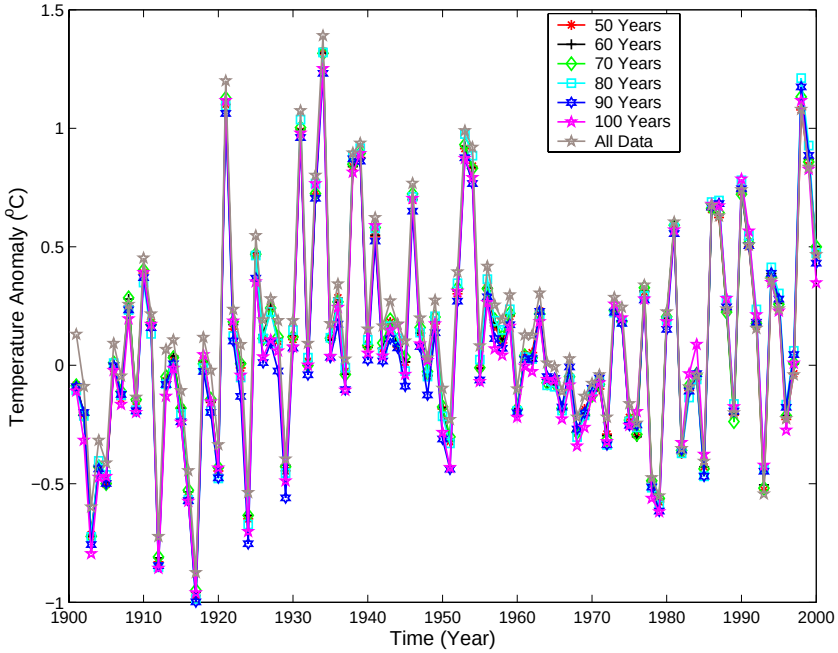

Figure 26. The US average annual mean of the daily mean SAT calculated from the GDCN's data sets with different minimum record lengths by using the area-weighted average method. The annual data are computed directly from the daily anomalies.

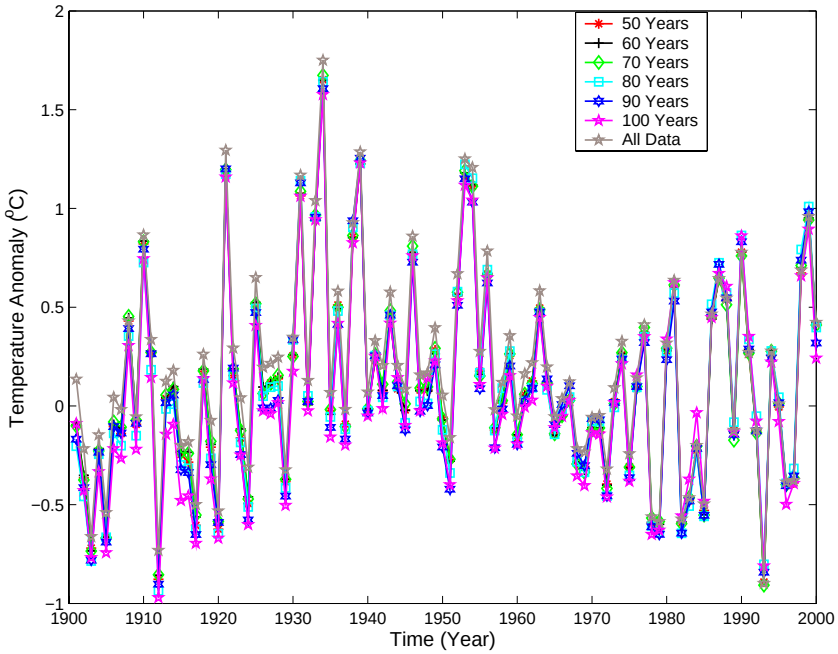

Figure 27. The US average annual mean of the daily maximum SAT calculated from the GDCN's data sets with different minimum record lengths by using the area-weighted average method. The annual data are computed directly from the daily anomalies.

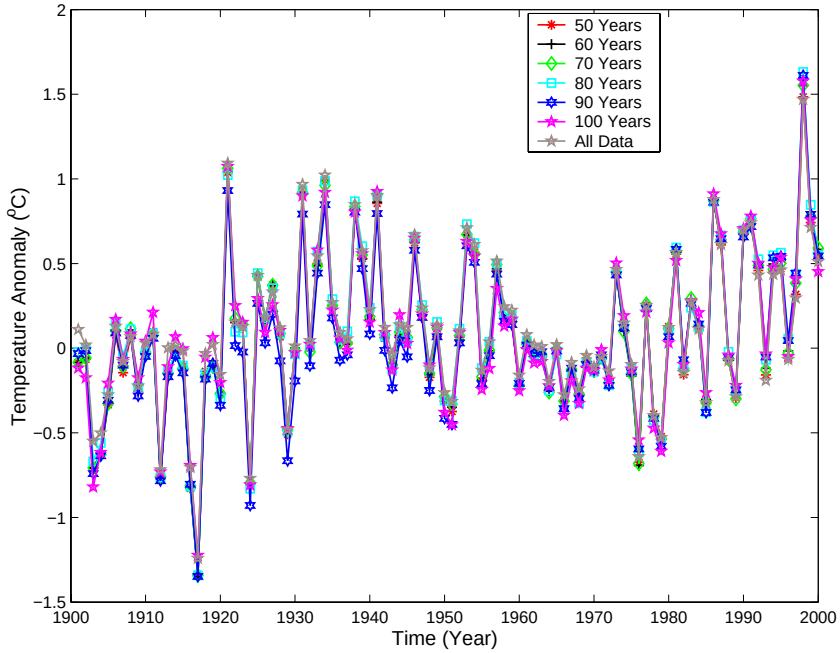

Figure 28. The US average annual mean of the daily minimum SAT calculated from the GDCN's data sets with different minimum record lengths by using the area-weighted average method. The annual data are computed directly from the daily anomalies.

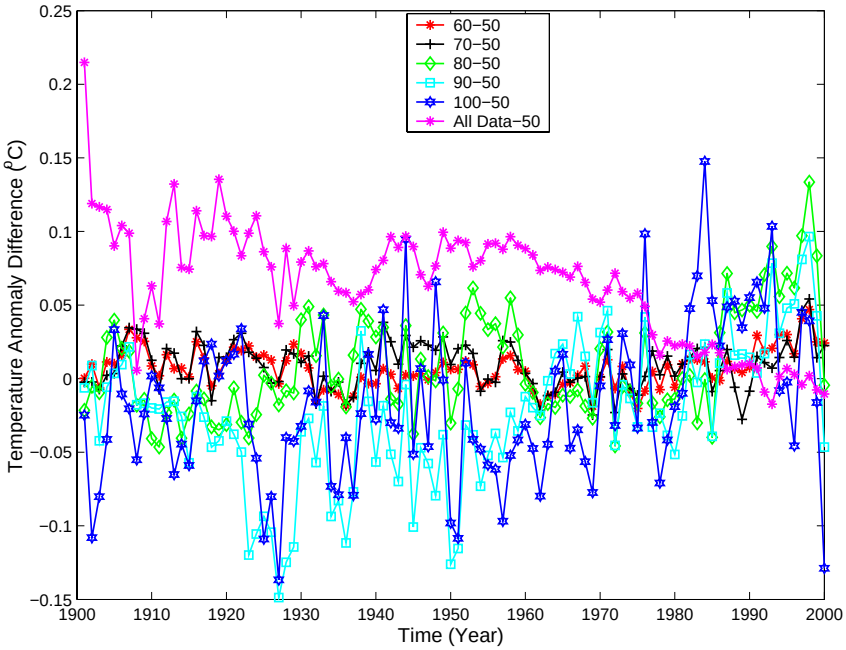

Figure 29. Differences between the US average annual mean of daily mean SAT computed from the sub-network G50 and from other sub-networks by using the area weighted average method.

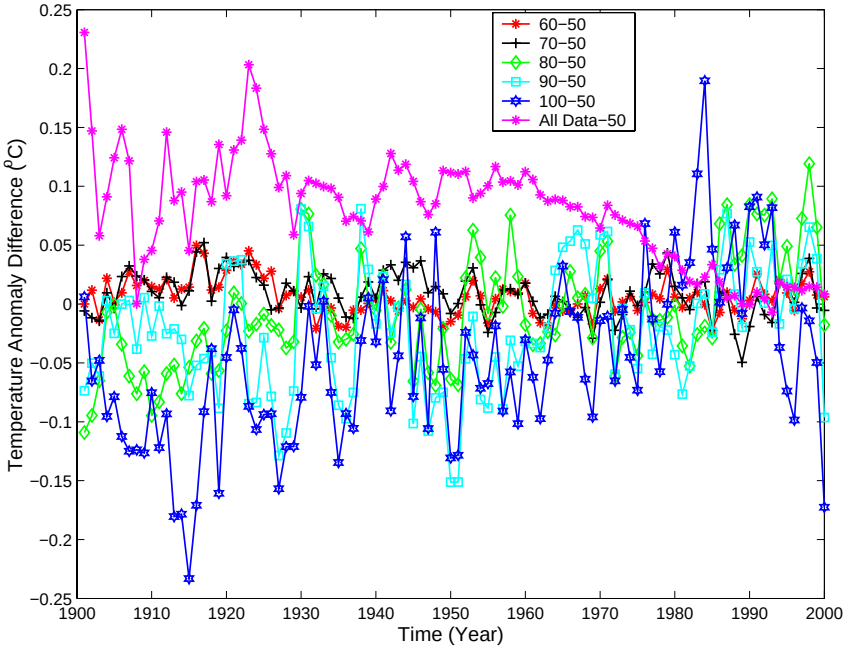

Figure 30. Differences between the US average annual mean of daily maximum SAT computed from the sub-network G50 and from other sub-networks by using the area weighted average method.

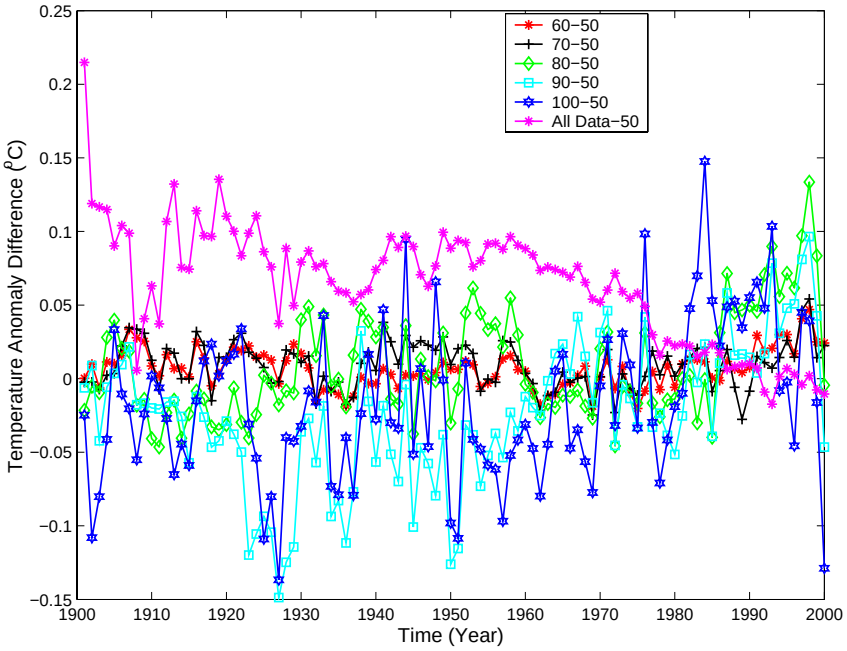

Figure 31. Differences between the US average annual mean of daily minimum SAT computed from the sub-network G50 and from other sub-networks by using the area weighted average method.

4.3 Comparing the daily mean temperature results with existing ones

Balling and Idso (2002) computed and compared the average of the mean temperature of different climatological databases: the United States Historical Climatological Network's (USHCN)'s raw, FILNET, and the urban adjusted data, and the IPCC'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. More recently 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° latitude × 5.0° longitude, and all these data sets are at monthly resolution. Their Figure 1 results are compared with the results obtained from the GDCN's data based on daily anomalies. Our two annual anomalies are considered: the first is computed from the monthly anomalies, and the second is computed directly from the daily anomalies. 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 results from the sub-network with minimum record of 70 years (i.e., G70) is shown here. This G70 sub-network includes 785 stations and has a reasonable spatial coverage of the contiguous US.

First, we compare the results of the annual means derived from the monthly anomalies. Balling and Idso (2002)'s four curves computed by area-weighted average are compared with our simple average of the annual mean from the monthly anomalies of

the GDCN sub-network of a minimum record length of 70 years (see Figure 32). The plot shows some differences of the five curves. Our GDCN results and the USHCN's raw data results are most similar. Table 1 lists the maximum differences between the Balling and Idso's four results and the GDCN data of different sub-networks. The occurrence years of the maximum differences are in the parentheses. The difference between the GDCN's and the USHCN's raw data sets is the lowest in all cases, perhaps because these two datasets include almost the same station networks and because no adjustments have been made in either of the data sets, particularly the sub-GDCN network of 70-year minimum record length. Among the sub-networks of different record lengths, the smallest value of the maximum differences is with the sub-network of 70 years of minimum record length, and the value is 0.107°C with the USHCN's raw data and happened in the year 1931. The results obtained from the GDCN's data with a minimum record length of 50 and 60 years are also comparable with those of 70 years. The largest maximum difference belongs to G100 and the value is 0.690°C with the FILNET data in the year 1933. The main cause of this large difference is likely due to the different spatial coverage of stations and simple average method.

Again for the monthly anomalies, the area-weighted averages are considered. Figure 33 shows five curves: four from Balling and Idso (2002) and one of ours obtained from the GDCN sub-network of a minimum record length of 70 years. Our spatial average is calculated by using the Thiessen polygon type of area-weighted average implemented by a $0.1° \times 0.1°$ grid. Once again, the results from the GDCN and the USHCN's raw data sets are the most similar, and their maximum differences are the smallest, around $0.20 \sim 0.22$°C occurred in year 1963. The maximum differences

between our results and the four results reported in Balling and Idso (2002) are listed in Table 2. To check whether the grid resolution alter the results, the grid resolutions of $0.5° \times 0.5°$ and $1.0° \times 1.0°$ are also implemented and similar results have been obtained (see Tables 3 and 4). The values of the maximum differences do not vary much either among the sub-network of different minimum record lengths or among the grid resolutions. The occurring years of the maximum differences are exactly the same in the three grids $0.1° \times 0.1°$, $0.5° \times 0.5°$ and $1.0° \times 1.0°$. These results show that the area-weighted average is important to reduce the errors due to the station coverage, but the area-weighted average itself is insensitive to the area a station represents. In other words, the area-weighted average algorithm is robust. Thus, it is important to take the area-weights into account, but it is not crucial that one calculates the area for each station with high accuracy.

Table 1. Maximum absolute difference between the simple US average annual mean of daily mean SAT computed for the GDCN data with different minimum record lengths and the results of Balling and Idso (2002). The number inside the parentheses refers to the year of observation of the maximum difference.

Minimum record length of GDCN's data sets (years)	Maximum absolute difference between the GDCN results and other USHCN results, and the year of the maximum difference (°C)			
	IPCC	Raw	FILNET	Urban Adjusted
50	0.291 (1992)	0.131 (1931)	0.263 (1933)	0.241 (1933)
60	0.289 (1992)	0.129 (1931)	0.263 (1933)	0.241 (1933)
70	0.275 (1992)	0.107 (1931)	0.276 (1933)	0.254 (1933)
80	0.376 (1931)	0.329 (1954)	0.508 (1934)	0.486 (1934)
90	0.424 (1931)	0.379 (1979)	0.501 (1933)	0.479 (1933)
100	0.658 (1931)	0.491 (1933)	0.690 (1933)	0.668 (1933)
All Data	0.290 (1992)	0.138 (1945)	0.320 (1936)	0.296 (1935)

Table 2. Maximum absolute difference between the area-weighted US average (with 0.1°x0.1° grid) annual mean of daily mean SAT computed for the GDCN monthly data with different minimum record lengths and the results of Balling and Idso (2002). The number inside the parentheses refers to the year of observation of the maximum difference.

Minimum record length of GDCN's data sets (years)	Maximum absolute difference between the GDCN results and other USHCN results, and the year of the maximum difference (°C)			
	IPCC	Raw	FILNET	Urban Adjusted
50	0.260 (1992)	0.239 (1949)	0.346 (1934)	0.325 (1998)
60	0.243 (1992)	0.228 (1949)	0.344 (1934)	0.322 (1934)
70	0.249 (1992)	0.214 (1949)	0.343 (1934)	0.321 (1934)
80	0.199 (1992)	0.237 (1996)	0.345 (1934)	0.323 (1934)
90	0.244 (1992)	0.288 (1949)	0.288 (1991)	0.289 (1991)
100	0.215 (1992)	0.246 (1949)	0.293 (1998)	0.300 (1998)
All Data	0.273 (1992)	0.221 (1934)	0.421 (1934)	0.399 (1934)

Table 3. Maximum absolute difference between the area-weighted US average (with 0.5°x0.5° grid) annual mean of daily mean SAT computed for the GDCN monthly data with different minimum record lengths and the results of Balling and Idso (2002). The number inside the parentheses refers to the year of observation of the maximum difference.

Minimum record length of GDCN's data sets (years)	Maximum absolute difference between the GDCN results and other USHCN results, and the year of the maximum difference (°C)			
	IPCC	Raw	Filnet	Urban adjusted
50	0.273 (1992)	0.226 (1949)	0.327 (1934)	0.322 (1998)
60	0.254 (1992)	0.219 (1949)	0.325 (1934)	0.303 (1934)
70	0.259 (1992)	0.205 (1949)	0.32(1934)	0.298 (1934)
80	0.204 (1992)	0.223 (1996)	0.322 (1934)	0.30(1934)
90	0.255 (1992)	0.278 (1949)	0.284 (1991)	0.285 (1991)
100	0.224 (1992)	0.232 (1949)	0.275 (1998)	0.282 (1998)
All Data	0.281 (1992)	0.200 (1963)	0.389 (1934)	0.367 (1934)

Table 4. Maximum absolute difference between the area-weighted US average (with 1.0°x1.0° grid) annual mean of daily mean SAT computed for the GDCN monthly data with different minimum record lengths and the results of Balling and Idso (2002). The number inside the parentheses refers to the year of observation of the maximum difference.

Minimum record length of GDCN's data sets (years)	Maximum absolute difference between the GDCN results and other USHCN results, and the year of the maximum difference (°C)			
	IPCC	Raw	Filnet	Urban adjusted
50	0.250 (1992)	0.238 (1949)	0.329 (1934)	0.312 (1998)
60	0.246 (1992)	0.224 (1949)	0.329 (1934)	0.307 (1934)
70	0.251 (1992)	0.210 (1949)	0.325 (1934)	0.303 (1934)
80	0.210 (1992)	0.231 (1996)	0.321 (1934)	0.299 (1934)
90	0.256 (1992)	0.283 (1949)	0.288 (1991)	0.289 (1991)
100	0.236 (1992)	0.227 (1949)	0.287 (1998)	0.294 (1998)
All Data	0.271 (1992)	0.211 (1934)	0.411 (1934)	0.389 (1934)

Next, we compare our annual mean computed from the daily anomalies with the results reported in Balling and Idso (2002). Figure 34 shows five curves: four reported in Balling and Idso and one from our GDCN sub-network of a minimum record length of 70 years by using the daily anomalies and simple spatial average. The differences between the GDCN's and the USHCN's raw data sets are again the lowest in all cases. The maximum differences similar to Tables 2-4 are listed in Table 5. The smallest value of the maximum difference is 0.119°C occurred in 1931. The largest maximum difference belongs to the sub-network G100 and the value is 0.701°C with FILNET in the year 1933. Again, the differences are mainly due to the different spatial coverage of station networks. The good agreement is of course not a surprise since the GDCN and the USHCN share almost all the stations. The purpose of this comparison is to show the reliability of our GDCN results rather than to make new discoveries.

Table 5. Maximum absolute difference between the simple US average annual mean of daily mean SAT computed for the GDCN data with different minimum record lengths and the results of Balling and Idso (2002). The GDCN station's annual mean is computed directly from the daily anomalies. The number inside the parentheses refers to the year of observation of the maximum difference.

Minimum record length of GDCN's data sets (years)	Maximum absolute difference between the GDCN results and other USHCN results, and the year of the maximum difference (°C)			
	IPCC	Raw	FILNET	Urban Adjusted
50	0.302 (1992)	0.143 (1931)	0.280(1933)	0.258 (1933)
60	0.299 (1992)	0.141 (1931)	0.281 (1933)	0.259 (1933)
70	0.285 (1992)	0.119 (1931)	0.294 (1933)	0.272 (1933)
80	0.365 (1950)	0.315 (1934)	0.515 (1934)	0.493 (1934)
90	0.400(1931)	0.353 (1979)	0.516 (1933)	0.494 (1933)
100	0.632 (1931)	0.502 (1933)	0.701 (1933)	0.679 (1933)
All Data	0.300(1992)	0.151 (1936)	0.348 (1936)	0.323 (1936)

Again for the daily anomalies, consider the area-weighted average of annual means. Figure 35 shows the five curves: four from Balling and Idso (2002) and one from the GDCN sub-network of a minimum record length of 70 years and a $0.5° \times 0.5°$. The GDCN results agree with Balling and Idso's results better because all of the results are computed from area-weighted average. Particularly in the case of USHCN, that network and the GDCN sub-network of 70 years of minimum record length share almost the same stations. Table 6 shows the maximum differences for the grid $0.1° \times 0.1°$ like those in Tables 2-5. Tables 7 and 8 list the same maximum differences for grids $0.5° \times 0.5°$ and $1.0° \times 1.0°$, respectively. The maximum differences between the GDCN's and the USHCN's raw data sets remain low in most cases (see Tables 6-8), and the lowest value is 0.196°C occurred in the year 1934.

Table 6. Maximum absolute difference between the area-weighted US average (with the 0.1°x0.1° grid) annual mean of daily mean SAT computed for the GDCN data with different minimum record lengths and the results of Balling and Idso (2002). The GDCN station's annual mean is computed directly from the daily anomalies. The number inside the parentheses refers to the year of observation of the maximum difference.

Minimum record length of GDCN's data sets (years)	Maximum absolute difference between the GDCN results and other USHCN results, and the year of the maximum difference (°C)			
	IPCC	Raw	FILNET	Urban Adjusted
50	0.272 (1992)	0.254 (1931)	0.344 (1934)	0.335 (1998)
60	0.257 (1992)	0.238 (1931)	0.349 (1934)	0.327 (1934)
70	0.265 (1992)	0.221 (1931)	0.346 (1934)	0.324 (1934)
80	0.208 (1992)	0.243 (1996)	0.349 (1934)	0.327 (1934)
90	0.256 (1992)	0.275 (1949)	0.310 (1991)	0.311 (1991)
100	0.224 (1992)	0.241 (1931)	0.303 (1998)	0.310 (1998)
All Data	0.285 (1992)	0.225 (1934)	0.425 (1934)	0.403 (1934)

Table 7. Maximum absolute difference between the area-weighted US average (with the 0.5°x0.5° grid) annual mean of daily mean SAT computed for the GDCN data with different minimum record lengths and the results of Balling and Idso (2002). The GDCN station's annual mean is computed directly from the daily anomalies. The number inside the parentheses refers to the year of observation of the maximum difference.

Minimum record length of GDCN's data sets (years)	Maximum absolute difference between the GDCN results and other USHCN results, and the year of the maximum difference (°C)			
	IPCC	Raw	FILNET	Urban Adjusted
50	0.284 (1992)	0.226 (1931)	0.330 (1934)	0.334 (1998)
60	0.265 (1992)	0.219 (1931)	0.322 (1934)	0.300 (1934)
70	0.273 (1992)	0.211 (1931)	0.321 (1934)	0.299 (1934)
80	0.213 (1992)	0.229 (1996)	0.325 (1934)	0.303 (1934)
90	0.267 (1992)	0.264 (1949)	0.305 (1991)	0.306 (1991)
100	0.236 (1992)	0.235 (1931)	0.288 (1998)	0.295 (1998)
All Data	0.293 (1992)	0.196 (1934)	0.396 (1934)	0.374 (1934)

Table 8. Maximum absolute difference between the area-weighted US average (with the 1.0°x1.0° grid) annual mean of daily mean SAT computed for the GDCN data with different minimum record lengths and the results of Balling and Idso (2002). The GDCN station's annual mean is computed directly from the daily anomalies. The number inside the parentheses refers to the year of observation of the maximum difference.

Minimum record length of GDCN's data sets (years)	Maximum absolute difference between the GDCN results and other USHCN results, and the year of the maximum difference (°C)			
	IPCC	Raw	FILNET	Urban Adjusted
50	0.262 (1992)	0.237 (1949)	0.333 (1934)	0.323 (1998)
60	0.258 (1992)	0.224 (1931)	0.332 (1934)	0.310 (1934)
70	0.265 (1992)	0.213 (1931)	0.328 (1934)	0.306 (1934)
80	0.228 (1992)	0.231 (1996)	0.322 (1934)	0.300 (1934)
90	0.267 (1992)	0.278 (1949)	0.307 (1991)	0.308 (1991)
100	0.244 (1992)	0.236 (1931)	0.302 (1998)	0.309 (1998)
All Data	0.283 (1992)	0.211 (1934)	0.411 (1934)	0.389 (1934)

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 1930 - 2000, a statistically significant trend of 0.05°C/decade at 5% level of significance is obtained in the GDCN's data sets, which is exactly 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 anomalies to assess the climate changes in terms of higher statistical moments.

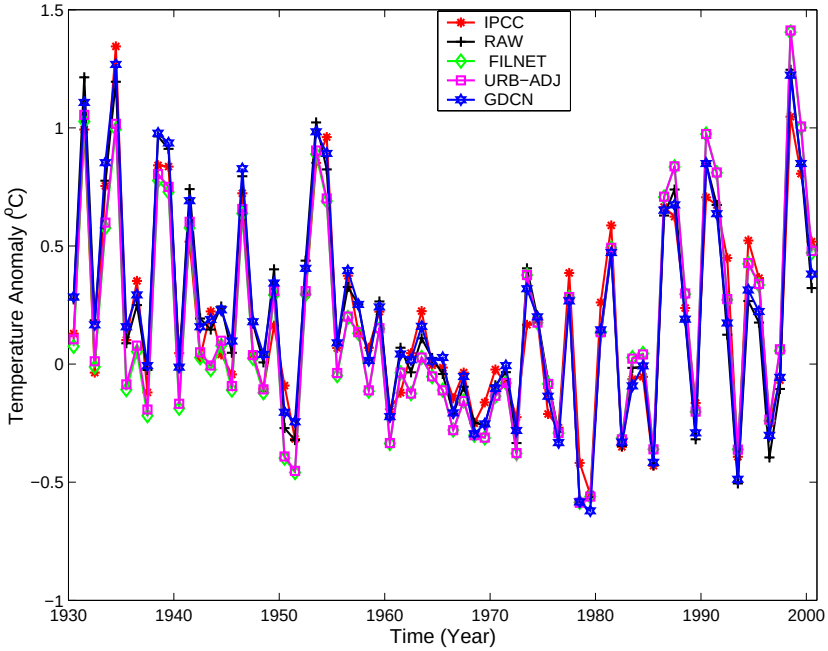

Figure 32. The US average annual mean of the daily mean SAT computed from different data sets: the United States Historical Climatological Network's (USHCN) raw data, FILNET, and the urban adjusted data set, the IPCC, and the GDCN. The GDCN result is obtained from the simple average of the station's annual anomalies computed from the monthly anomalies. The other results are from Balling and Idso (2002).

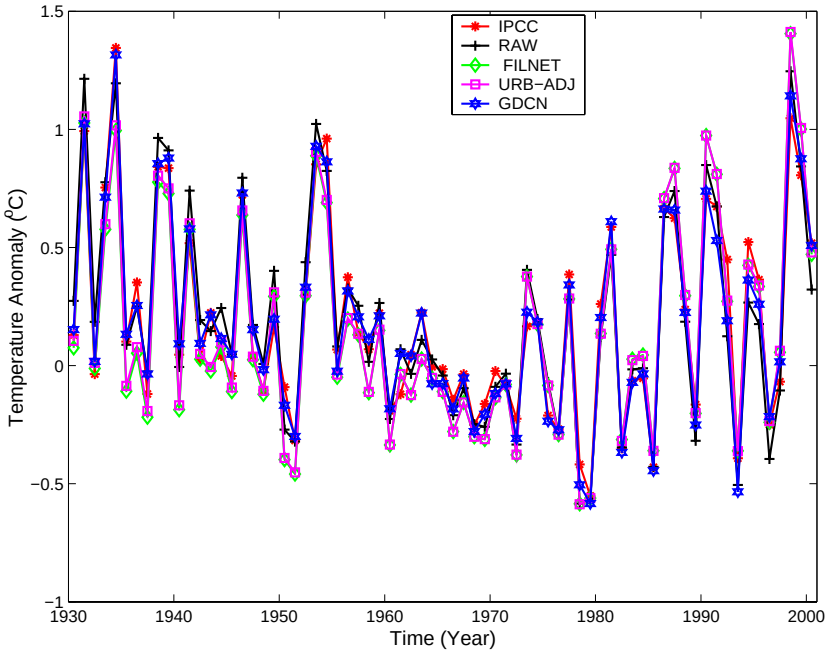

Figure 33. The US average annual mean of the daily mean SAT computed from different data sets: the United States Historical Climatological Network's (USHCN) raw data, FILNET, and the urban adjusted data set, the IPCC, and the GDCN. The GDCN result is obtained from the area-weighted average with the 0.5°x0.5° grid and the station's annual anomalies computed from the monthly anomalies. The other results are from Balling and Idso (2002).

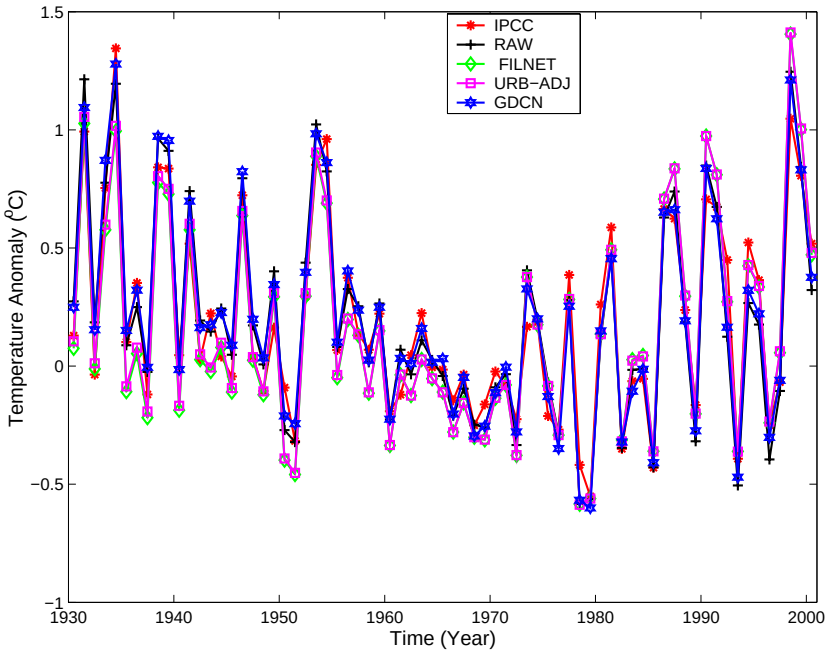

Figure 34. The US average annual mean of the daily mean SAT computed from different data sets: the United States Historical Climatological Network's (USHCN) raw data, FILNET, and the urban adjusted data set, the IPCC, and the GDCN. The GDCN result is obtained from the simple average of the station's annual anomalies computed directly from the daily anomalies. The other results are from Balling and Idso (2002).

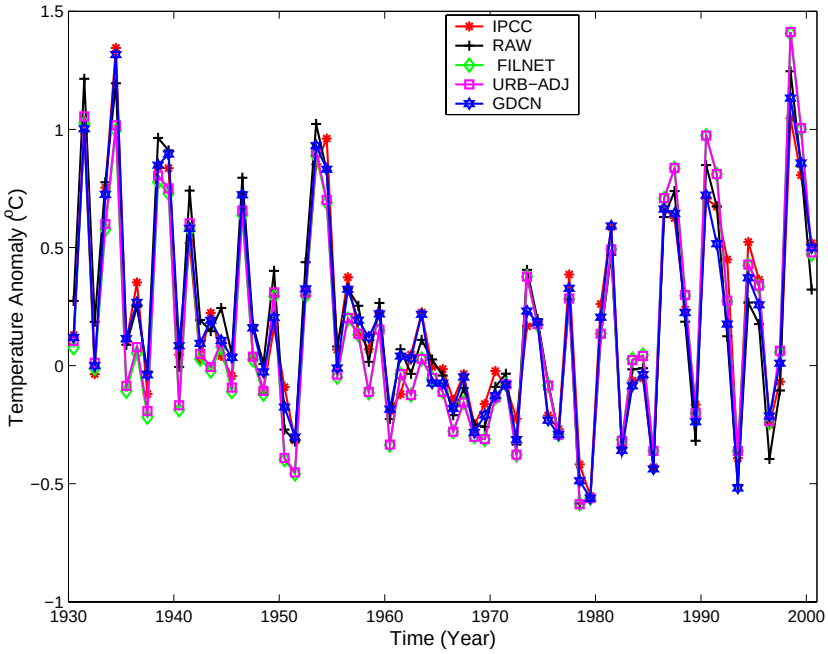

Figure 35. The US average annual mean of the daily mean SAT computed from different data sets: the United States Historical Climatological Network's (USHCN) raw data, FILNET, and the urban adjusted data set, the IPCC, and the GDCN. The GDCN result is obtained from the area-weighted-average (with of the 0.1°x0.1°) the station's annual anomalies computed directly from the daily anomalies. The other results are from Balling and Idso (2002).

4.4 The higher statistical moments

After having justified the reliability of the area-weighted average method and the GDCN daily mean SAT, we have gained the confidence in our monthly and daily anomalies, particularly the latter, which will be used to analyze the higher moments of the mean, maximum, and minimum daily SATs. Equation (3) is used to compute the standard deviation and its square is the station or grid point variance, and Equation (4) with the value of k equal to 3 and 4 are used for skewness and kurtosis, respectively. Equation (5) is used to compute the area-weighted average.

4.4.1 Higher statistical moments calculated from the USHCN data

We first make a test on the USHCN monthly data and try to find what are the possible variations and trends contained in this dataset in terms of the first four statistical moments. Again, the three kinds of USHCN's monthly data are used: the raw, time-of-observation-bias-corrected (TOB), and FILNET networks for the maximum and minimum daily SATs. Altogether, the USHCN has1,221 stations in the period of 1901 to 2000. Figures 36-43 show the first four orders of statistical moments of the maximum and minimum SATs. The moments are computed on annual basis, and hence $\tau = 12$ in Equations (3) and (4). The spatial average is made by area-weighted average using the $0.1° \times 0.1°$ grid. Namely, anomalies are gridded onto the $0.1° \times 0.1°$ grid by the nearest station assignment method, and the statistical moments are first computed for each grid point, and Equation (5) for the area-weighted average is applied to the moments. The first SAT moment has comparatively larger differences than the higher moments (the variance, skewness, and kurtosis). This result infers that no large error would be introduced into the final results for the higher moments even if correction and

homogenization procedures are not thoroughly carried out in the USHCN raw data set, as is the GDCN case now.

Figures 32-35 are for the daily mean SATs for the period of 1930 to 2000. Figures 36-39 are for the daily maximum SATs for the period of 1901 to 2000, while Figures 40-43 are the daily minimum SATs for the period of 1901 to 2000. They show the common characteristics of unusually low volatility of SATs in the 1960s and relatively higher volatility in the periods of 1930-1950 and 1980-2000. Namely, the variances are consistently small in the 1960s and can be large in the periods of 1930-1950 and 1980-2000. More comments on the properties of the higher moments will be given in the next sub-section that contains the results computed from the daily anomalies data from the GDCN. Supposedly, the GDCN results on the higher moments are more conclusive than the monthly USHCN data since the USHCN annual higher moments are computed from only 12 samples (i.e., $\tau = 12$). Normally, the sample size should be at least 30 by some statisticians or 50 by others.

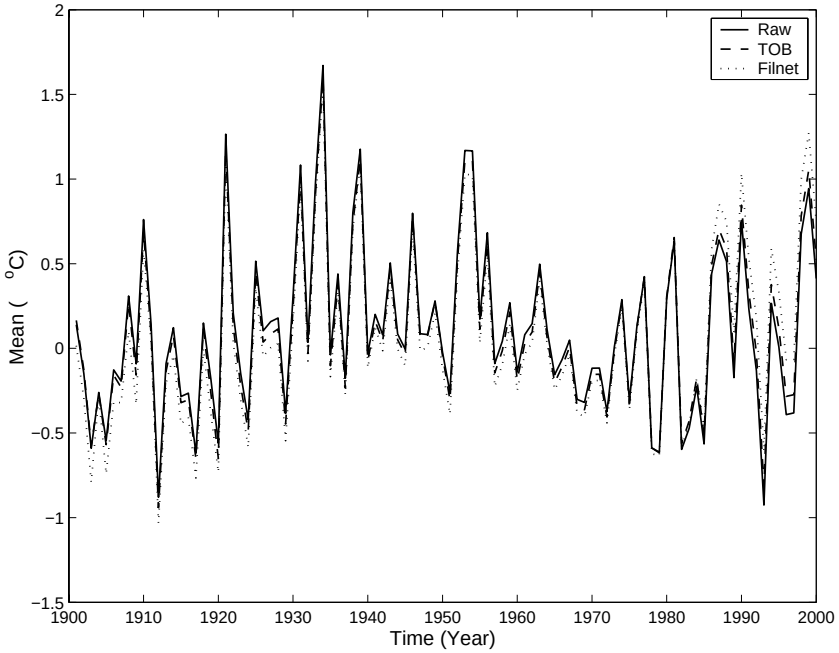

Figure 36. The US area-weighted average annual mean of the daily maximum SAT from the USHCN raw, time-of-observation-bias-corrected, and FILNET data sets.

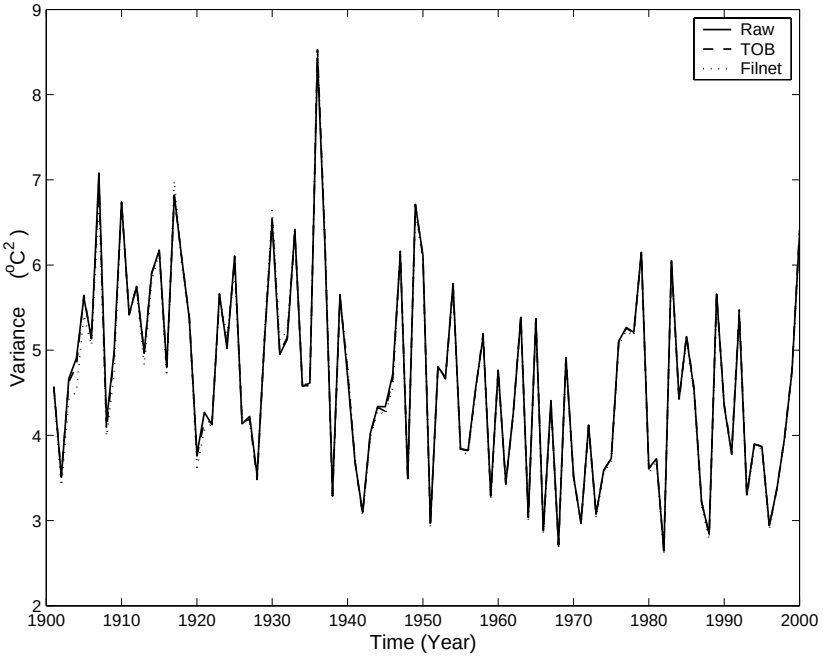

Figure 37. The US area-weighted average annual variance of the daily maximum SAT from the USHCN raw, time-of-observation-bias-corrected, and FILNET data sets.

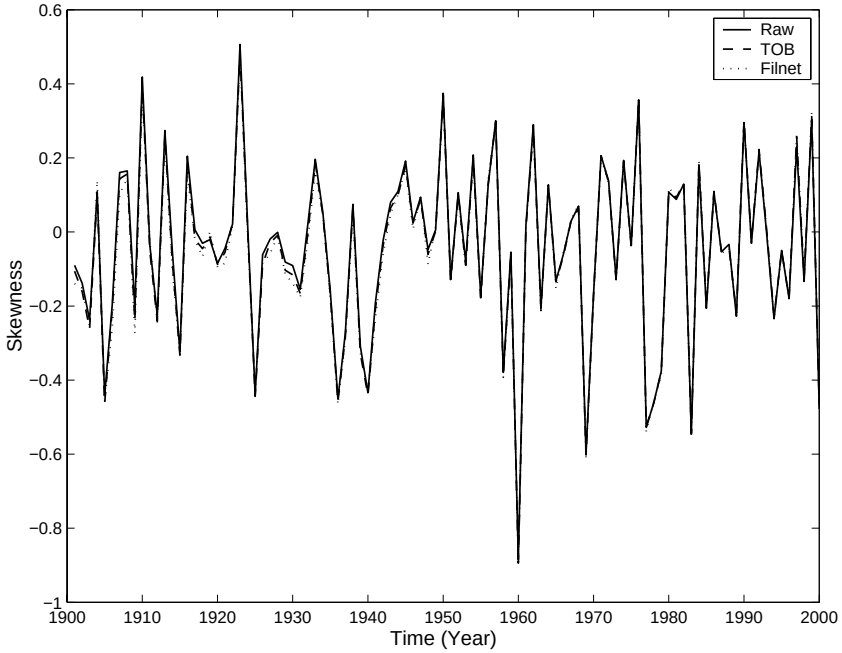

Figure 38. The US area-weighted average annual skewness of the daily maximum SAT from the USHCN raw, time-of-observation-bias-corrected, and FILNET data sets.

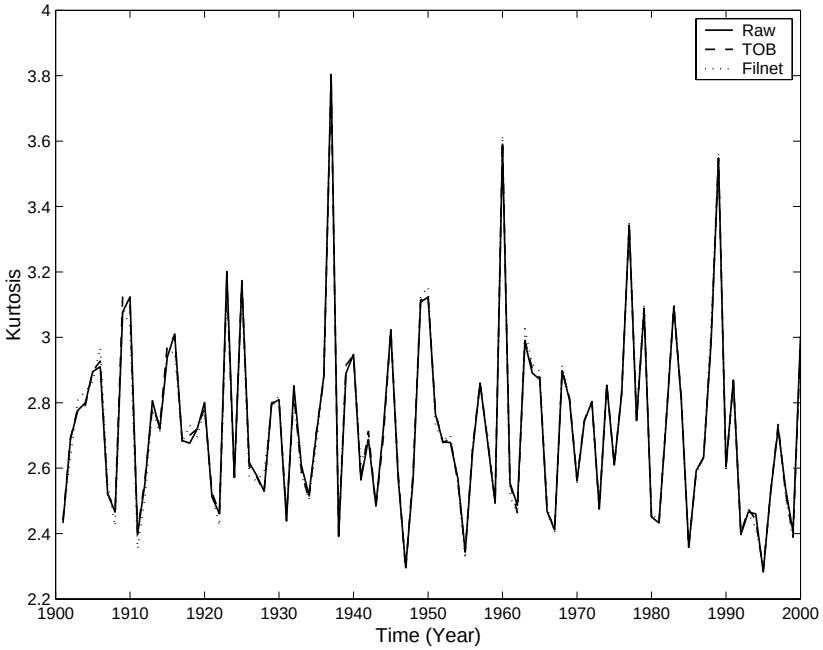

Figure 39. The US area-weighted average annual kurtosis of the daily maximum SAT from the USHCN raw, time-of-observation-bias-corrected, and FILNET data sets.

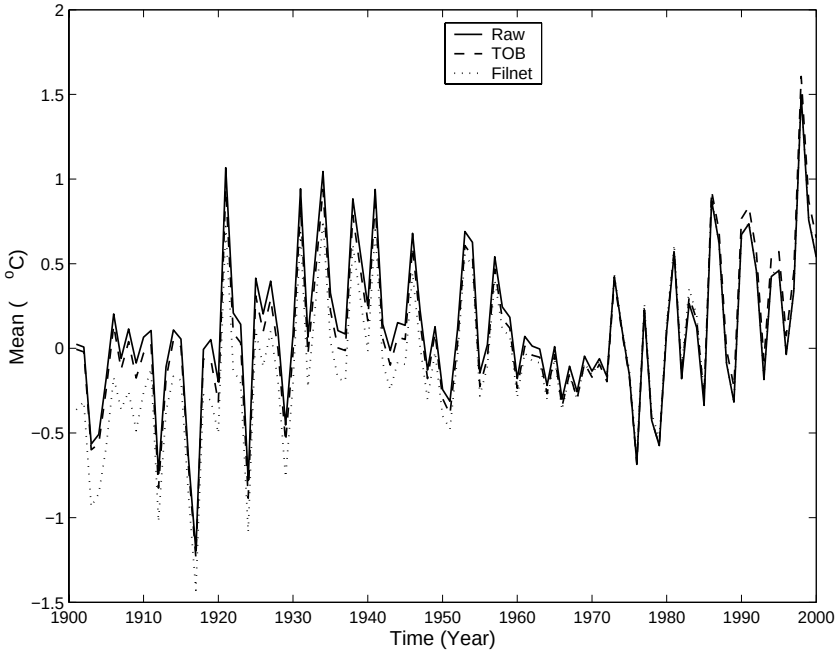

Figure 40. The US area-weighted average annual mean of the daily minimum SAT from the USHCN raw, time-of-observation-bias-corrected, and FILNET data sets.

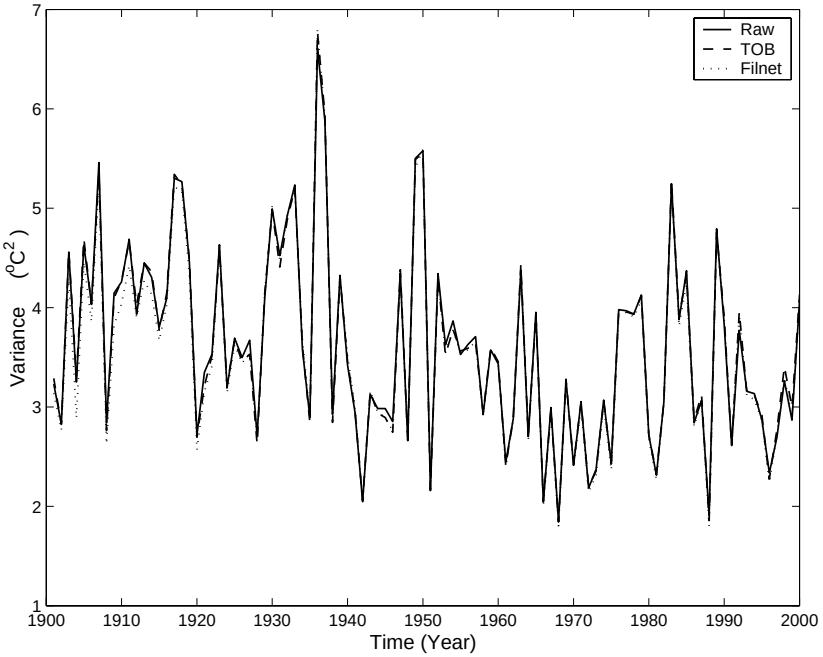

Figure 41. The US area-weighted average annual variance of the daily minimum SAT from the USHCN raw, time-of-observation-bias-corrected, and FILNET data sets.

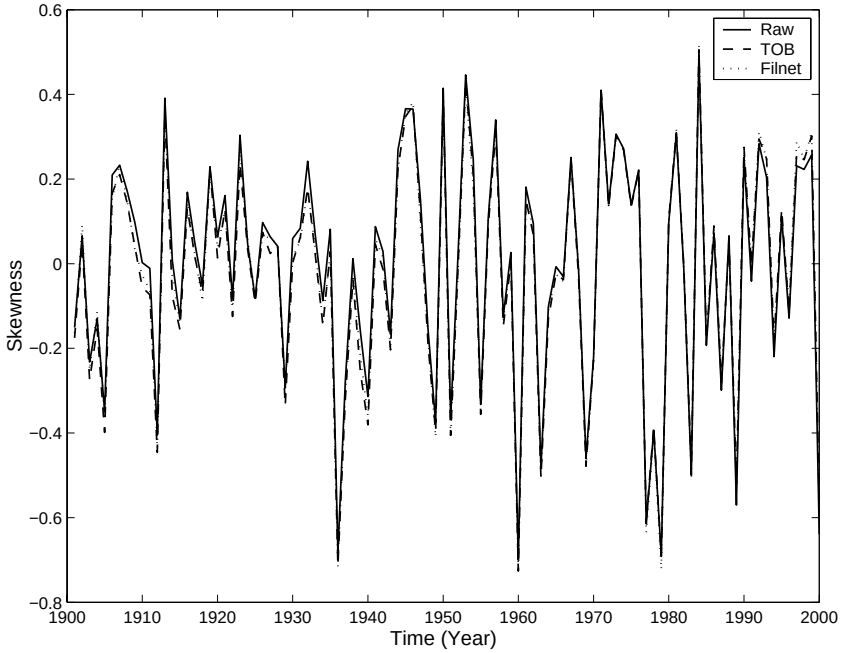

Figure 42. The US area-weighted average annual skewness of the daily minimum SAT from the USHCN raw, time-of-observation-bias-corrected, and FILNET data sets.

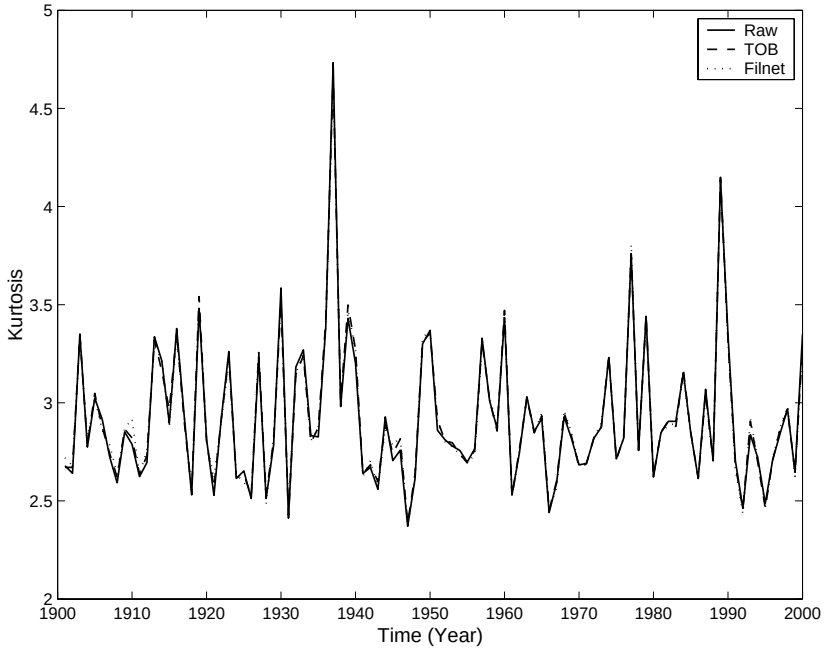

Figure 43. The US area-weighted average annual kurtosis of the daily minimum SAT from the USHCN raw, time-of-observation-bias-corrected, and FILNET data sets.

4.4.2 Higher moments calculated from the GDCN daily anomaly data

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° latitude $\times$ 0.1° longitude grid. The higher order moments are computed at each grid point for each year according to Equations (3) and (4) 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 Equation (5). Figures 44, 45, 46, and 47 show the results of the annual mean, variance, skewness, and kurtosis for the mean, maximum, and minimum SATs, respectively. Figure 44 roughly shows an increasing trend of the annual mean of the daily mean SAT from 1901 to 1934, a decreasing trend from 1935 to 1975, and an increasing trend from 1976 onward. Of course, this is a rough assessment of the change points. There are various kinds of statistical methods for detecting change points. We choose not to care too much about the rigorous detection in this report. The characteristics of the trends for annual mean of the daily maximum and minimum SAT are similar to those of the mean SAT in the entire period of analysis, but the magnitudes of the changes show some differences. 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 from 1917 to 1934 and that from 1976 to 1998 is noticeable. These conclusions are well known and are well reflected in our results.

Figure 45 shows that the general trend of variance is decreasing, although the rate is very slow. This is in contrast to the misconception of many people (perhaps due to the increase of media coverage and the increased effectiveness of mass propagation) that the climate has become more violent and hence more variable. 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 variance. Gong and Ho (2004), Balling (1997), and Michael 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 many more data, and has gone through an extensive comparison.

Figure 45 shows two conspicuous large peaks at 1936 and 1989, which might be caused by strong La Nina events. The El Nino oscillations supposedly imply large variance, but 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 45 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 our life experience that the daily maximum temperature fluctuates a lot more than the daily minimum temperature, particularly in the winter.

Figure 46 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 un-symmetric) 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.

The annual skewness computed from the monthly data is mostly positive (Figures 38 and 42). 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 47 shows the area-weighted US average of annual kurtosis of the daily mean, maximum, and minimum SAT computed from the GDCN daily anomalies. The figure shows two peaks at 1933 and 1989. 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. Thus, the maximum temperature has stronger

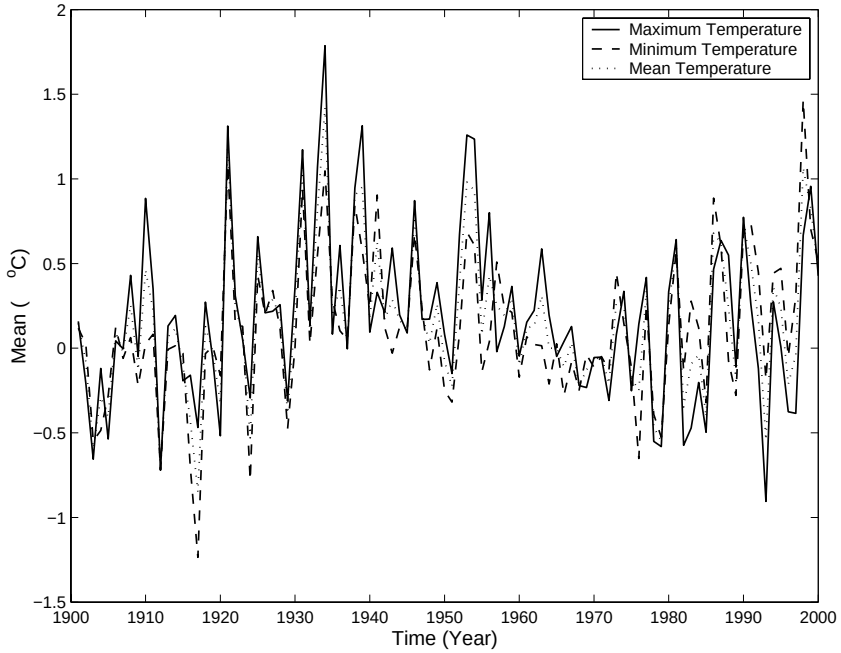

Figure 44. The area-weighted US average of annual mean of the daily mean, maximum, and minimum SAT computed from the GDCN daily anomalies.

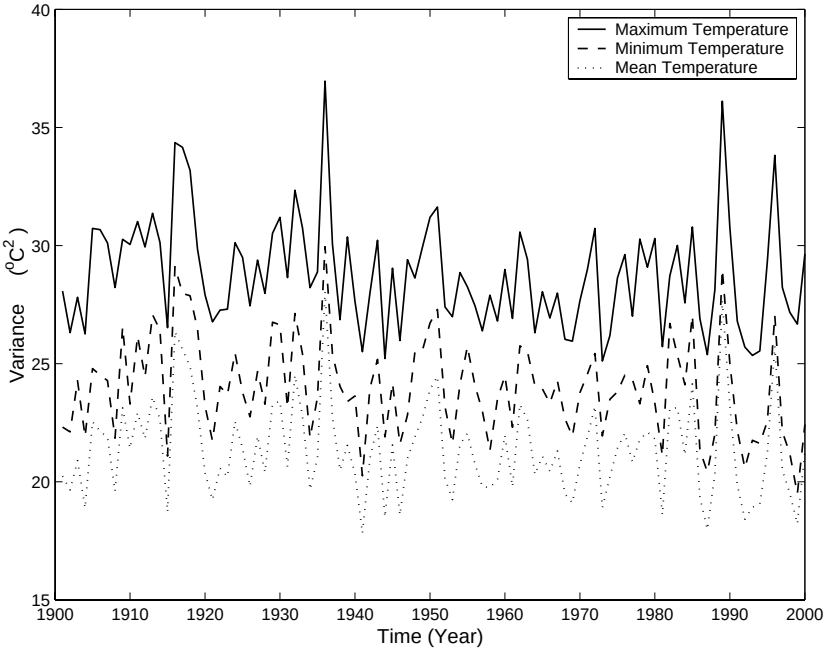

Figure 45. The area-weighted US average of annual variance of the daily mean, maximum, and minimum SAT computed from the GDCN daily anomalies.

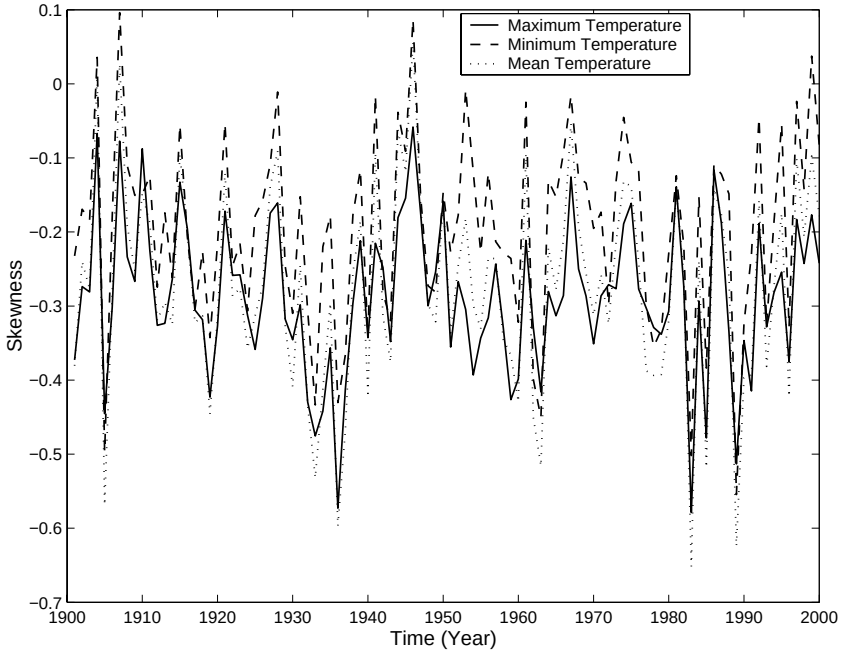

Figure 46. The area-weighted US average of annual skewness of the daily mean, maximum, and minimum SAT computed from the GDCN daily anomalies.

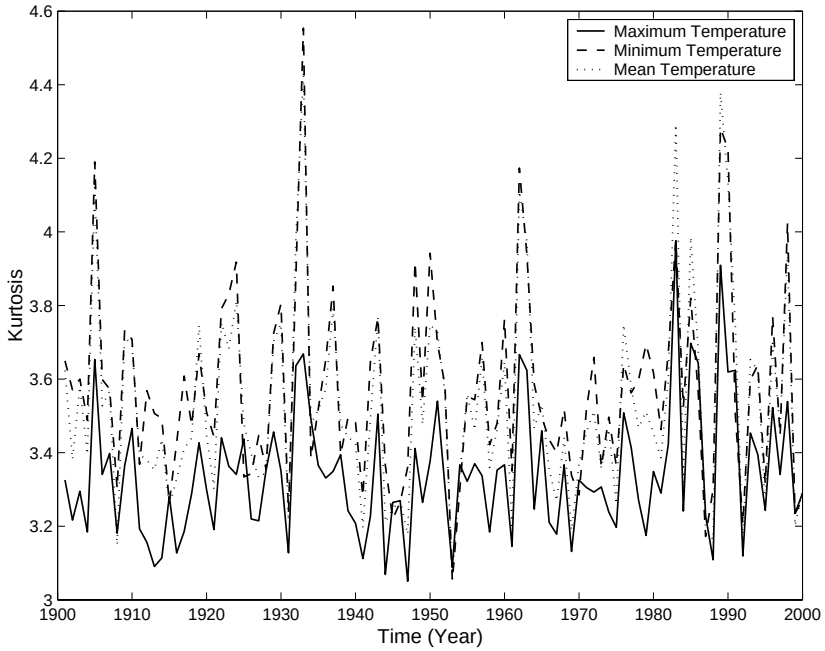

Figure 47. The area-weighted US average of annual kurtosis of the daily mean, maximum, and minimum SAT computed from the GDCN daily anomalies.

dispersion capability than the other two temperatures. This finding is consistent with the larger variance of the maximum temperature than the other two temperatures.

Compared with the kurtosis computed from the monthly data (see Figures 39 and 43), the ones computed from the daily data are larger. 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. This implies that the daily anomalies have stronger dispersion and are more widely distributed than the monthly anomalies. Our intuition supports this conclusion.

Houghton et al. (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 in these 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

$$M^k(t) = A + Bt + \varepsilon, \tag{6}$$

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 unit of the moment, and T is the time, which can be expressed either in days, years, decades, or centuries.

Table 9 lists the slope B values of the different moments for the mean temperature. 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 in1946-1975. These trends are comparable to that of the global average temperature.

Table 9. Trends of different moments for the daily mean temperature.

Period	Trends of Moments (Per Century)			
	Average (°C)	Variance ($°C^2$)	Skewness	Kurtosis
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

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/century; i.e., the mean temperature of the contiguous US become less variable, about a degree Celsius ^{2}, in the past century.

The skewness also showed 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 clearly observed due to that the absolute value of the decreasing slope was very small, only -0.021/century. Further, the increasing trend 0.645/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 and1976-2000. The relatively large magnitude of decreasing slope of kurtosis in the 1976-2000 period implies stronger dispersion in the last two decades of the 20th century.

Table 10 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 thirty 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 -1.684°C^2/century. The prevailed change was from 1910 to 1945 and the net change was about 2.0°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 10. Trends of different moments for the daily maximum surface air temperature.

Period	Trends of Moments (Per Century)			
	Average (°C)	Variance (°C^2)	Skewness	Kurtosis
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

Table 11 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 20th 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°C^2/century. The large decreasing slope -11.23°C^2/century from 1976 to 2000 is a reflection of the sharp decrease of the diurnal cycle's amplitude (or called the diurnal temperature range) (Houghton et al., 2001). 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/century). The decrease implies that the minimum temperature's dispersion became stronger from 1976 to 2000.

Table 11. Trends of different moments for the daily minimum surface air temperature.

Period	Trends of Moments (Per Century)			
	Average (°C)	Variance (°C^2)	Skewness	Kurtosis
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

4.5 Probability distribution functions

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 probability distribution functions (PDF). The assessment of the PDF changes, however, requires a lot more data, which we do not have. Thus, 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 remained stations will not completely cover the entire contiguous US.

The SAT histograms of in the periods 1910-1945, 1946-1975, and 1976-2000 are calculated and compared. The trend of the different temperatures of the individual stations with a minimum record length of 70 years (i.e., equal to 25,567 daily data records) in the twentieth century 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 fifty stations where the highest increasing trend in the period of 1910-2000 is detected are used to construct the histograms. Locations of the fifty influential stations of largest climate change for the maximum and minimum surface air temperatures are shown in Figures 48 and 49, respectively. These stations are mostly located in the western US.

Since the selected most influential stations do not cover the entire contiguous US, the results do not reflect the spatial average statistics of the US. However, our results raise some interesting questions on the issues in the studies of climate change. For example, how are the moments' changes reflected in the PDF changes? Can the PDF change reflect some changes of climate dynamics or atmospheric circulation patterns? Questions like these may not be apparent from the histogram constructed by considering all sets of data at once. The use of daily data in different periods and different seasons for the most influential stations may help answer questions of this kind. Our computed histogram results and statistical moments are shown in Figures 50-55 and Tables 12-14.

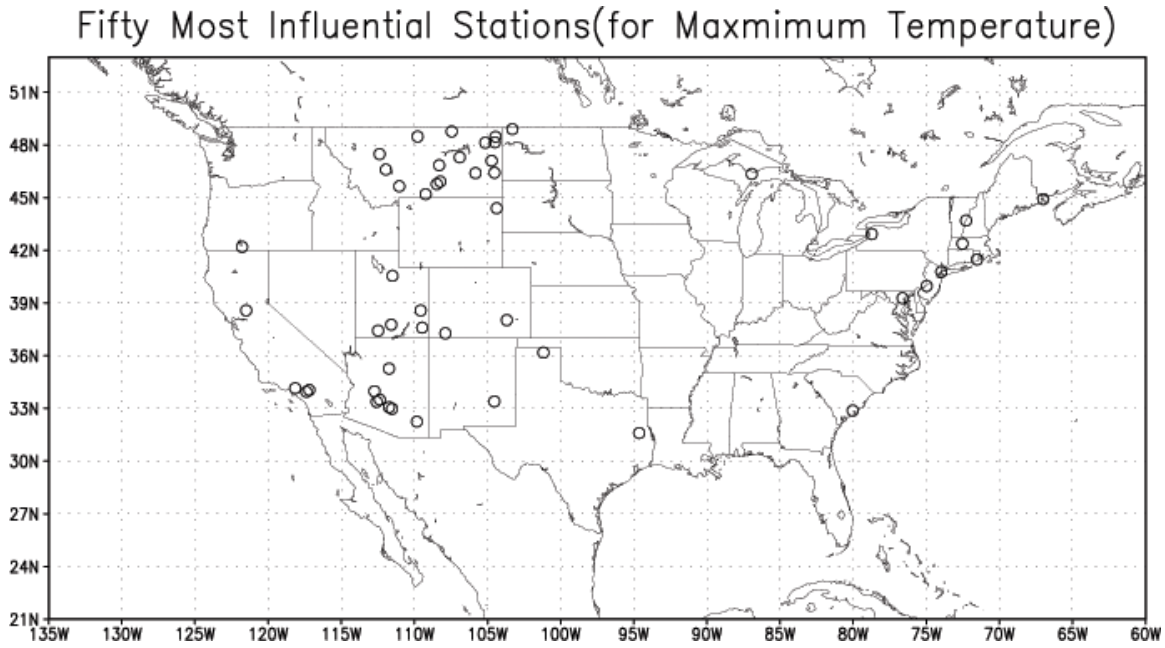

Figure 48. Locations of the fifty most influential stations for the maximum surface air temperature.

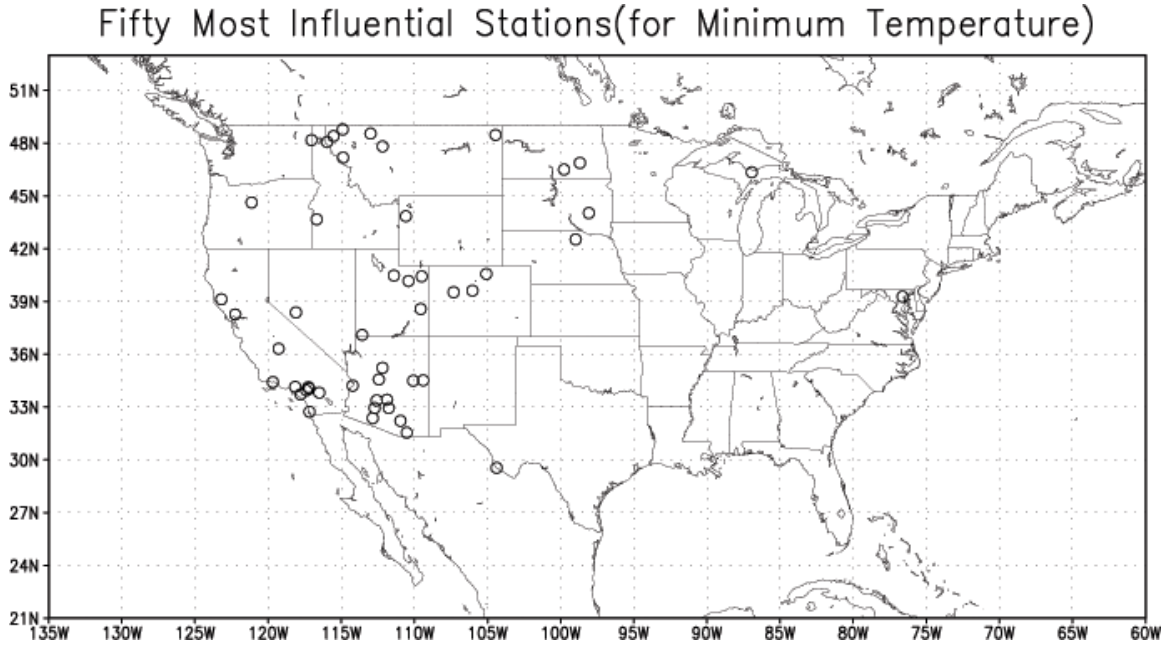

Figure 49. Locations of the fifty most influential stations for the minimum surface air temperature.

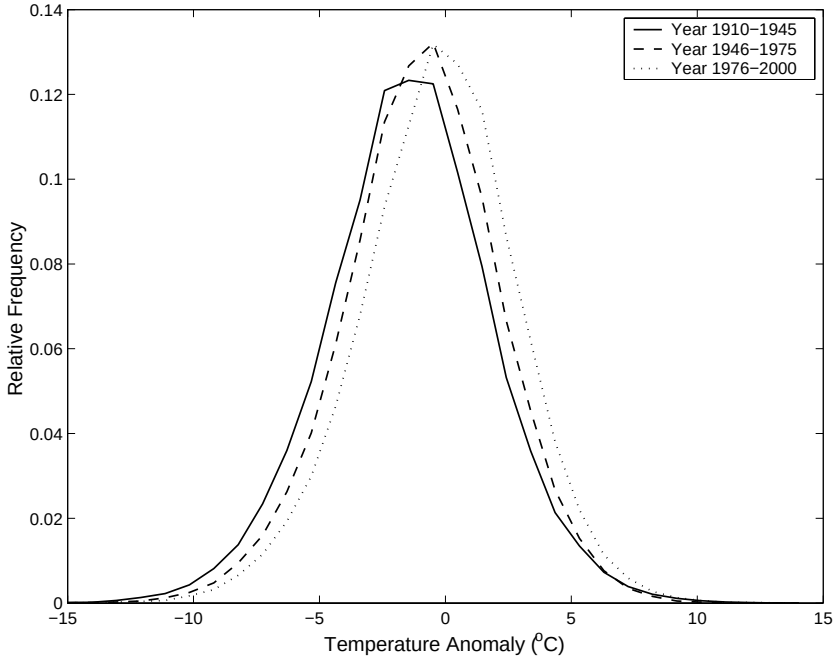

Figure 50. Three histograms of the daily mean SAT for the summer season of fifty stations where the highest increasing trend occurred.

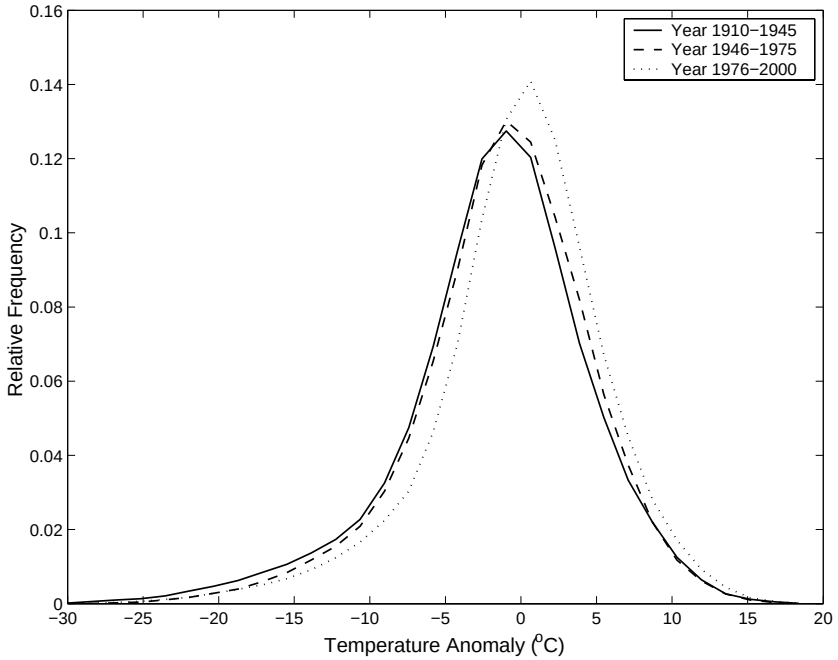

Figure 51. Three histograms of the daily mean SAT for the winter season of fifty stations where the highest increasing trend occurred.

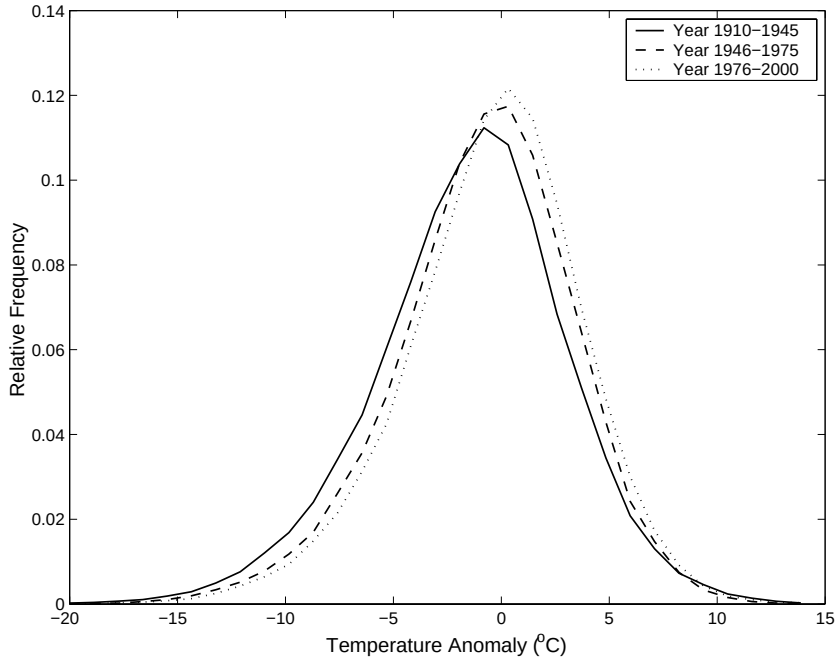

Figure 52. Three histograms of the daily maximum SAT for the summer season of fifty stations where the highest increasing trend occurred.

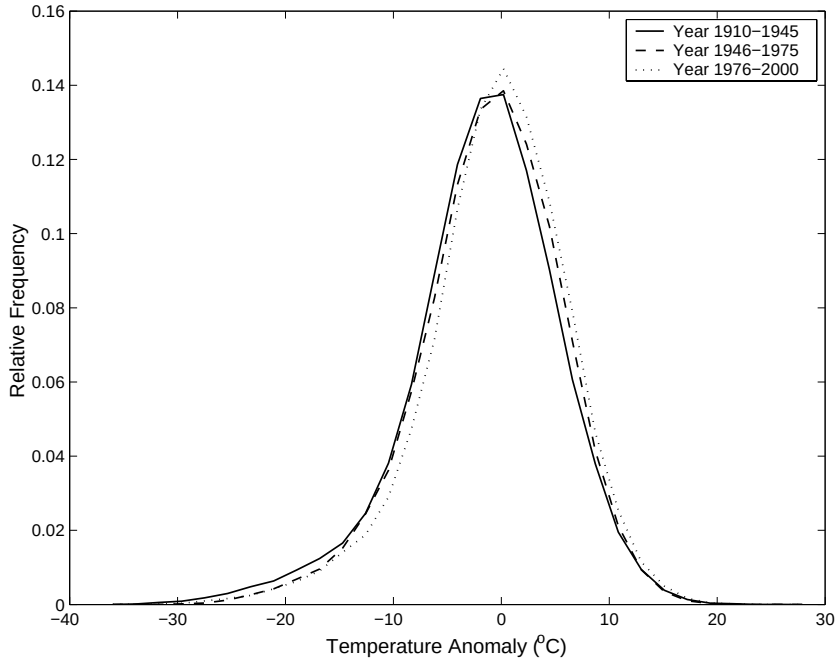

Figure 53. Three histograms of the daily maximum SAT for the winter season of fifty stations where the highest increasing trend occurred.

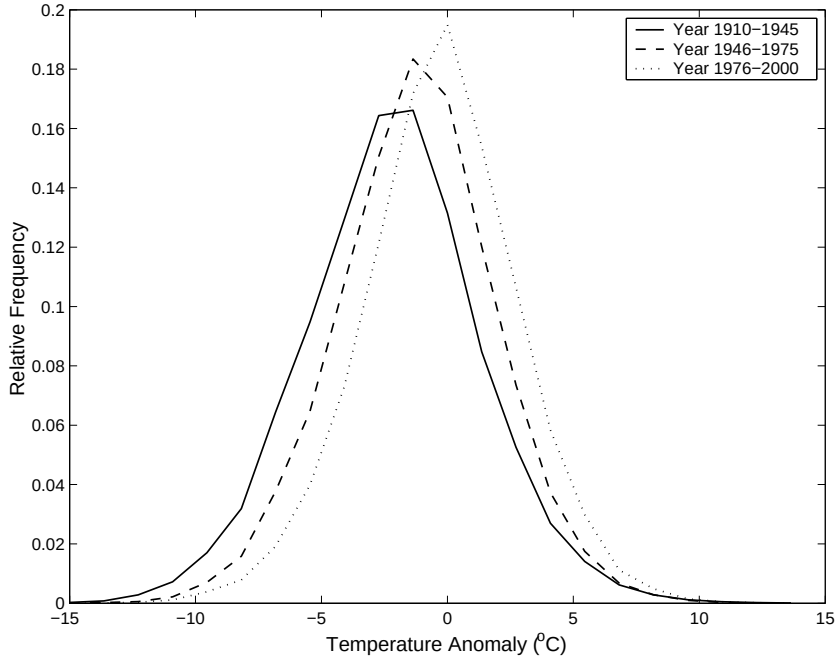

Figure 54. Three histograms of the daily minimum SAT for the summer season of fifty stations where the highest increasing trend occurred.

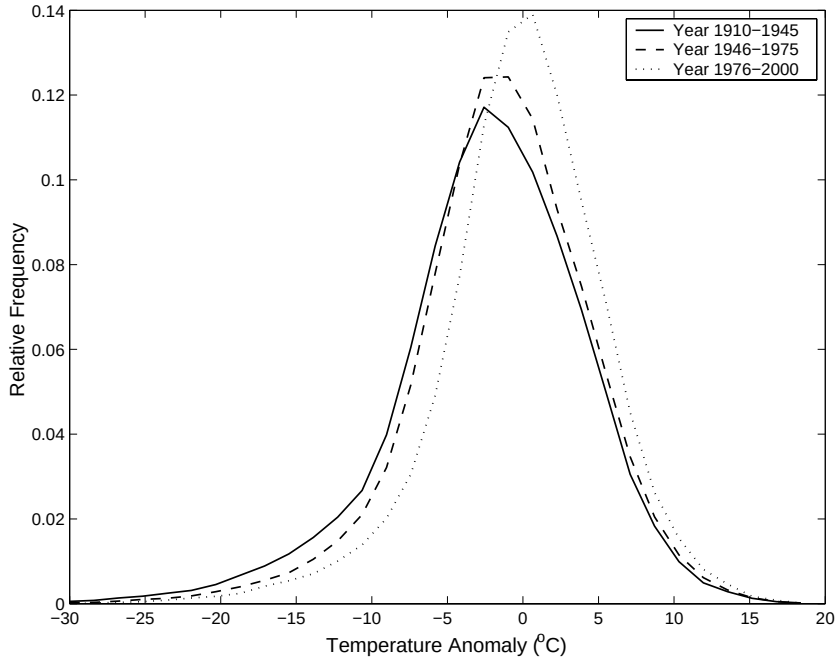

Figure 55. Three histograms of the daily minimum SAT for the winter season of fifty stations where the highest increasing trend occurred.

The histogram Figures 50-55 and statistical moments Tables 12-14 imply the following.

A) 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 marched from 1910-1945 to 1946-1975 to 1976-2000. However, the magnitudes of the right shift are not the same for the six cases.

B) 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 the daily minimum SAT as concluded in many studies.

C) 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 fifty 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.

D) No significant changes of skewness and kurtosis: The histograms do not show a clear change of shapes or symmetry, except Figures 54 and 55. Thus, the changes of skewness and kurtosis are only noticeable in the minimum SAT in winter. The net changes of skewness and kurtosis are 0.05 and 0.23,

respectively, implying that the histogram had become more peaky and more symmetric.

Table 12. Moments of the average temperature in the winter and summer seasons for different periods of time in the fifty stations where the highest increasing trends occurred.

Season	Period	Moments			
		Average (°C)	Variance ($°C^2$)	Skewness	Kurtosis
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

Table 13. Moments of the maximum surface air temperature in the winter and summer seasons for different periods of time in the fifty stations where the highest increasing trends occurred.

Season	Period	Moments			
		Average (°C)	Variance ($°C^2$)	Skewness	Kurtosis
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

Table 14. Moments of the minimum surface air temperature in the winter and summer seasons for different periods of time in the fifty stations where the highest increasing trends occurred.

Season	Period	Moments			
		Average (°C)	Variance ($°C^2$)	Skewness	Kurtosis
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

The above findings suggest that the 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 the three temperatures, the statistics of the minimum temperature of winter had the largest changes.

5 Conclusions and Discussions

We have analyzed the GDCN daily minimum, maximum, and mean temperatures from January 1, 1901 to December 31, 2000 over the contiguous US. The daily SAT anomalies are with respect to the 1961-1990 monthly climatology. The mean, variance, skewness, and kurtosis in different seasons, different sub-periods, and different sub-networks of stations have been calculated. The statistical moments at station locations have been spatially averaged by both simple arithmetic average and area-weighted average. The US averages of the mean SATs have been compared with the existing results obtained from other station networks and analysis methods, but the variance, skewness, and kurtosis have no existing results for comparison. 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 fifty 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 Houghton et al. (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 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 unintuitive since many reports state that climate has become warmer and more violent. However, a growing number of researchers, including those performing computer simulations of the climate with increased green house gas conditions, have also found a reduction in the variability. Understanding the behavior of variation 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/support the findings of this study and other similar ones.

Several points are worth discussions. The first is the data homogenization problem. 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 and Coauthors, 1998). Therefore, such data sets must be homogenized before they can be used for climatological analysis (Vincent et al., 1999). For homogenization, a base/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 GDCN is a global data set. Quality tests have been carried out on each set of data, and the data sets of the contiguous US are considered to be better than those of other parts of the world. Within the contiguous US, different databases are available, and some of them are derived ones. Lately, scientists have been debating on whether the adjustments made in the United States Historical Climatological Network's (USHCN)'s temperature data sets are justified (Balling and Idso, 2002; Vose et al., 2003; Keim et al., 2003). Balling and Idso (2002) argued that these adjustments had introduced a warm bias into the US temperature database and that the adjustments applied to the temperature data may result in bias errors greater than any errors in the raw time series. However, Vose et al. (2003), Keim et al. (2003), and Hansen et al. (2001) argued that the adjustments made were essential and justified. This dispute is unlikely to be resolved in the near future, so the adjustments to be carried out must be justified individually. Any corrections/adjustments required or applied depend on the objective of the data use (Smith, 2004).

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 space optimal averaging (RSOA) by Folland et al. (2001) or spectral optimal averaging (SOA) by Shen et al. (1994 and 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 go to 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 or 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).

Acknowledgments

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 x 0.1° longitude, 0.5° latitude x 0.5° longitude, and 1.0° latitude x 1.0° longitude resolutions were generated by Xiadong 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.

References

Balling, R. C. (1997). Analysis of daily and monthly spatial variance components in historical temperature records. *Physical Geography* **18**, 544-552.

Balling, R. C. and C. D. Idso (2002). Analysis of adjustments to the United States historical climatological network (USHCN) temperature database. *Geophys. Res. Lett.* **29**, 25-1 - 25-3.

Cunderlik, J. M. and D. H. Burn (2003). Non-stationary pooled flood frequency analysis. *J. Hydrol.* **276**, 210-223.

Folland, C. K., N. A. Rayner, S. J. Brown, T. M. Smith, S. S. P. Shen, D. E. Parker, I. Macadam, P. D. Jones, R. N. Jones, N. Nicholls, and D. M. H. Sexton (2001). Global temperature change and its uncertainties since 1861, *Geophys. Res. Lett.* **28**, 2621-2624.

Gleason, B. E. (2002). Data documentation for data set 9101, Global Daily Climatology Net- work, V1.0. Technical report, National Climatic Data Center, 151 Patton Ave, Asheville, NC 28801-5001, USA.

Gong, D. Y. and C. H. Ho (2004). Intra-seasonal variability of wintertime temperature over East Asia. *Intl. J. Climatol*. **24**, 131-144.

Hansen, J., R. Ruedy, J. Glascoe, and M. Sato (1999). GISS analysis of surface temperature change. *J. Geophys. Res.* **104**, 30997-31022.

Hansen, J., 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.* **106**, 23947-23963.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson (2001). Climate change 2001: The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. New York: Cambridge University Press.

Huang, N. E. and S. S. P. Shen (2005), Hlibert-Huang Transform and Its Applications, Singapore: World Scientific Press.

Jones, P.D. (1994). Hemispheric surface air temperature variations: a reanalysis and an update to 1993. *J. Climate* **7**, 1794-1802.

Jones, P.D., T.J. Osborn, and K.R. Briffa (1997). Estimating sampling errors in large-scale temperature averages. *J. Climate* **10**, 2548-2568.

Jones, P. D., T. M. L. Wigley, and P. M. Kelly (1982). Variations in surface air temperature Part I, northern hemisphere 1881 - 1980. *Mon. Weath. Rev.* **110**, 59-70.

Karl, T.R., H.F. Diaz, and G. Kukla (1988). Urbanization: its detection and effect in the United States climate record. *J. Climate*, **1**, 1099-1123.

Karl, T. R., D. R. Easterling, P. Y. Hughes, E. H. Mason, and J. Lawrimore (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.

Karl, T. R., P. D. Jones, R. W. Knight, G. Kukla, N. Plummer, V. Razuvayev, K. P. Gallo, J. Lindseay, R. J. Charlson, and T. C. Peterson (1993). A new perspective on recent global warming: Asymmetric trends of daily maximum and minimum temperature change. *Bull. Amer. Meteor. Soc.* **74**, 1007-1023.

Karl, T. R., R. W. Knight, and N. Plummer (1995). Trends in high-frequency climate variability in the twentieth century. *Nature* **377**, 217-220.

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.

Keim, B. D., A. M. Wilson, C. P. Wake, and T. G. Huntington (2003). Are there spurious temperature trends in the United States climate division database? *Geophys. Res. Lett.* **30**(7),1404. doi:10.1029/2003GL016295.

King'uyu, S. M., L. A. Gallo, and E. K. Anima (2000). Recent trends of minimum and maximum surface temperature over eastern Africa. *J. Climate* **13**, 2876-2886.

Meehl, G. A. and Coauthors (2000). An introduction to trends in extreme weather and climate events: Observations, socioeconomic impacts, terrestrial ecological impacts, and model projections. *Bull. Amer. Meteor. Soc.* **81**, 413-416.

Michaels, P. J., R. C. Balling, R. S. Vose, and P. C. Knappernberger (1998). Analysis of trends in the variability of daily and monthly historical temperature measurements. *Climate Res.* **10**, 27-33.

Parker, D. E., P. D. Jones, C. K. Folland, and A. Bevan (1994). Interdecadal change of surface temperature since the late nineteenth century. *J. Geophys. Res.* **99**, 14373-14399.

Peterson, T. C. and Coauthors (1998). Homogeneity adjustments of in situ atmosphere climate data: a review. *Intl. J. Climatol.* **18**, 1493-1517.

Peterson, T. C. and R. S. Vose (1997). An overview of the Global Historical Climatology Network temperature data base. *Bull. Amer. Meteor. Soc.* **78**, 2837-2849.

Shen, S. S. P., P. Dzikowski, G. Li, and D. Griffth (2001). Interpolation of 1961-97 daily temperature and precipitation data onto Alberta polygons of ecodistrict and soil landscapes of Canada. *J. App. Metero.* **40**, 2162-2177.

Shen, S. S. P., G. Li, C. Williams, and T. R. Karl (2004). Prediction of sea surface temperature from the Global Historical Climatology Network data. *Environmetrics*. **49** manuscript pages, accepted for publication.

Shen, S. S. P., G. R. North and K. Y. Kim (1994).Spectral approach to optimal estimation of the global average temperature, *J. Climate* **7**, 1999-2007 (1994).

Shen, S. S. P., T. M. Smith, C. F. Ropelewski and R. E. Livezey (1998). An optimal regional averaging method with error estimates and a test using tropical Pacific SST data, *J. Climate* **11**, 2340-2350.

Smith, T. (2004). Senior Scientist:National Climatic Data Center/NOAA. Personal Communications.

Vincent, L. A., X. Zhang, B. R. Bonsal, and W. D. Hogg (1999). Homogenization of daily temperatures over Canada. *J. Climate* **15**, 1322-1334.

Vinnikov, K. Y. and A. Robock (2002). Trends in moments of climatic indices. *Geophys. Res. Lett.* **29**, 14-1 - 14-4.

Vose, R. S., C. N. Williams, T. C. Peterson, T. R. Karl, and D. R. Easterling (2003). An evaluation of the time of observation bias adjustment in the U. S. Historical Climatology Network. *Geophys. Res. Lett.* **30**(20),2046. doi:10.1029/2003GL018111.

Wang, X. and S. S. P. Shen (1999). Estimation of spatial degrees of freedom of a climate field, *J. Climate* **12**, 1280-1291.