4-2003

Are there spurious temperature trends in the United States Climate Division database

Barry D. Keim
Louisiana State University at Baton Rouge

Adam M. Wilson
University of New Hampshire, Durham

Cameron P. Wake
University of New Hampshire - Main Campus, cameron.wake@unh.edu

Thomas G. Huntington
U.S. Geological Survey

Recommended Citation

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, 1404, doi:10.1029/2002GL016295, 7.

Are there spurious temperature trends in the United States Climate Division database?

Barry D. Keim
Southern Regional Climate Center, Louisiana State University, Baton Rouge, Louisiana, USA

Adam M. Wilson and Cameron P. Wake
Climate Change Research Center, University of New Hampshire, Durham, New Hampshire, USA

Thomas G. Huntington
United States Geological Survey, Augusta, Maine, USA

Received 17 September 2002; revised 11 October 2002; accepted 22 October 2003; published 11 April 2003.

[1] The United States (U.S.) Climate Division data set is commonly used in applied climatic studies in the United States. The divisional averages are calculated by including all available stations within a division at any given time. The averages are therefore vulnerable to shifts in average station location or elevation over time, which may introduce spurious trends within these data. This paper examines temperature trends within the 15 climate divisions of New England, comparing the NCDC's U.S. Divisional Data to the U.S. Historical Climate Network (USHCN) data. Correlation and multiple regression revealed that shifts in latitude, longitude, and elevation have affected the quality of the NCDC divisional data with respect to the USHCN. As a result, there may be issues with regard to their use in decadal- to century-scale climate change studies. *INDEX TERMS:* 3309 Meteorology and Atmospheric Dynamics: Climatology (1620); 1610 Global Change: Atmosphere (0315, 0325); 1803 Hydrology: Anthropogenic effects; 6620 Public Issues: Science policy; 6309 Policy Sciences: Decision making under uncertainty. **Citation:** Keim, B. D., A. M. Wilson, C. P. Wake, and T. G. Huntington, Are there spurious temperature trends in the United States Climate Division database?, *Geophys. Res. Lett.*, *30*(7), 1404, doi:10.1029/2002GL016295, 2003.

1. Introduction

[2] The National Climatic Data Center (NCDC) has subdivided the contiguous United States (U.S.) into 344 climate divisions, which represent nearly homogenous climatic regions [*Guttman and Quayle*, 1996]. For each division, mean monthly temperature, monthly water equivalent precipitation, Palmer Drought Severity Index, and Palmer Hydrological Drought Index values have been generated back to 1895. The data set is maintained by the NCDC.

[3] These climate divisional data have numerous applications, e.g., they are used to monitor the U.S. climate by the NCDC, the Climate Prediction Center, the National Drought Mitigation Center, and others. These divisional data sets are also used frequently in applied research [e.g., *Keim et al.*, 1996; *Leathers et al.*, 2000].

[4] Although these data are useful in many applied contexts, there may be issues with regard to their use in decadal- to century-scale climate change studies. Long-term trends in these data may be spuriously generated by the methods in which these data sets are calculated, i.e., whatever data are available at a given time are used in the average even though the total number of stations and their locations change through time. To date, the influence of these changes has not been fully documented. To illustrate the problems, temperature trends over the past 70 years (1931–2000) in the 15 climate divisions of New England (Figure 1) are analyzed and compared with United States Historical Climate Network (USHCN) data. The focus of this paper is on the potential impacts of the changing distribution of stations through time, as this alters the mean latitude and longitude, and average elevation of the stations used in these averages.

2. Data and Methods

[5] Monthly temperature from 1931–2000 for New England climate divisions are averaged by calendar year to produce annual time series. This analysis avoids the NCDC divisional data from 1895–1930, which are synthesized from statewide data as described by *Guttman and Quayle* [1996] and are therefore not true averages of data from within a climate division. Annual time series from the NCDC climate divisional data are compared to the USHCN station data [*Easterling et al.*, 1996], which are also available through the NCDC. The USHCN FILNET data are used as the control to evaluate the performance of the NCDC divisional data sets. USHCN data are excellent for this purpose since the stations were selected based on length and quality of data, which includes limiting the number of station changes. In addition, the FILNET data have undergone numerous quality assurances and adjustments to best characterize the actual variability in climate. These adjustments take into consideration the validity of extreme outliers, time of observation bias [*Karl et al.*, 1986], changes in instrumentation [*Quayle et al.*, 1991], random relocations of stations [*Karl and Williams*, 1987], and urban warming biases [*Karl et al.*, 1988]. Furthermore, missing data are estimated from surrounding stations to produce a nearly continuous data set for each station. Monthly averages from the USHCN stations within each climate division of New

0094-8276/03/2002GL016295

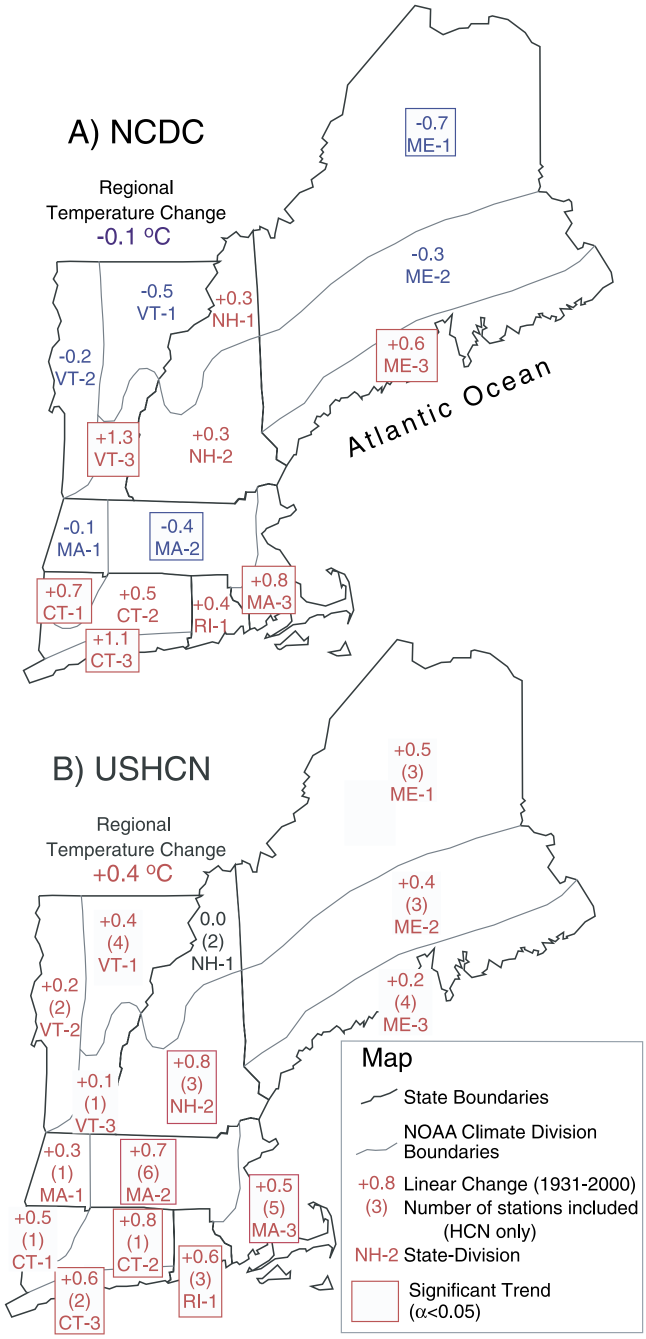

Figure 1. Climate divisions of the New England states, linear temperature trends from 1931–2000 by division (red = positive, blue = negative, boxed values = significant at $\alpha < .05$) for A) NCDC divisional data, and B) USHCN divisional data. Number of USHCN stations included by division, and regional average temperature trend.

England are then averaged annually, thereby constructing a comparative "divisional data" annual time series. Only USHCN FILNET stations with a continuous monthly record of temperature from January 1931 through December 2000 were included in the analysis. This eliminates biases introduced into the USHCN data set through changing the number and relative locations of the USHCN stations within each climate division. The number of USHCN stations used in these divisions range from 1 to 6 (Figure 1). The USHCN-derived divisional data set, hereafter referred to as USHCN data, are then used as a control to analyze 1931–2000 trends in the NCDC U.S. Climate Divisional data, hereafter referred to as NCDC data. These data sets are analyzed using linear regression, Pearson correlation, and multiple regression.

3. Comparison of Trends in Temperature

[6] Linear temperature trends for the 15 climate divisions in Figure 1 show that in all cases, the magnitudes of the NCDC data trends are different than the USHCN data, and in several cases the sign of the trends differ. Connecticut Division 2 (CT-2) and Massachusetts Division 3 (MA-3) are cases where the two trend slopes closely parallel one another (Figure 2) (for time series of all New England divisions, see e-supplement 1). These two NCDC divisional data sets have limited contamination introduced through time. However, the mean location of stations within the NCDC data for CT-2 are more southerly and at lower elevations than the USHCN data. This leads to a temperature series with a higher temperature, but similar trend. The opposite is the case for MA-3.

[7] General agreement between the two data sets, however, is more the exception than the rule. In several cases, both time series are showing increases in temperature, but at differing rates (Figure 1). For example, the New Hampshire Division 2 (NH-2) trend shows that the NCDC data

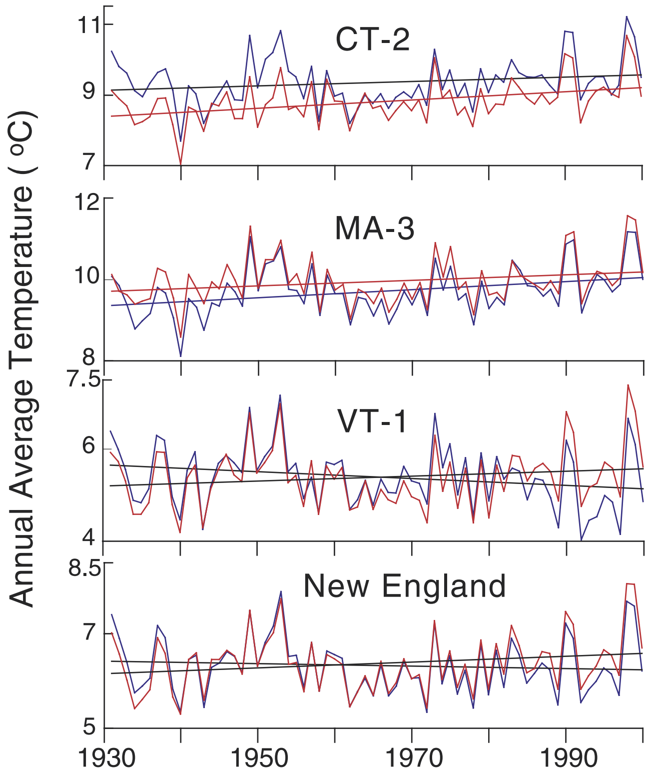

Figure 2. Time series and linear trends of annual temperature comparing NCDC (blue) and USHCN (red) divisional data for selected New England climate divisions, 1931–2000. Slopes significant at $\alpha < .05$ are presented in colored trend lines.

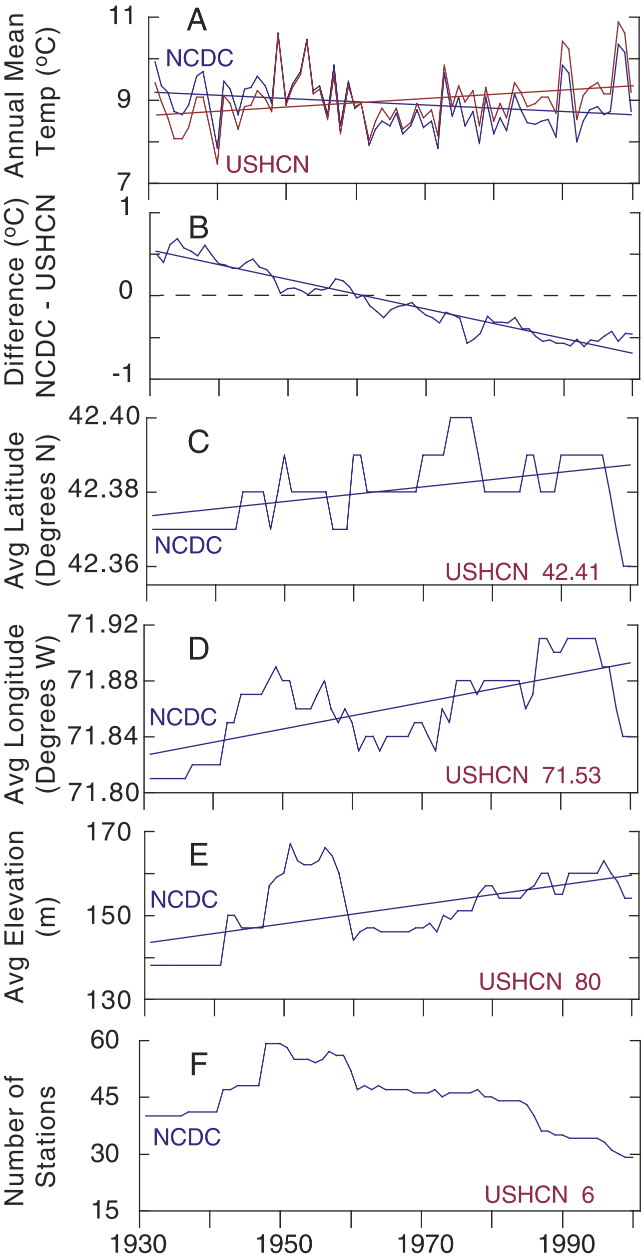

Figure 3. Massachusetts Division 2 (MA-2) annual time series of A) NCDC and USHCN divisional temperature; B) temperature difference between NCDC and USHCN; C) mean latitude of NCDC stations; D) mean longitude of NCDC stations; E) mean elevation of NCDC stations, and F) number of stations included in the NCDC divisional average. Constant USHCN values are presented in red in lower right of panels C-F.

increased in temperature by 0.3°C from 1931–2000, while the USHCN data increased by 0.8°C. Even more surprising are the climate divisions with opposing linear trends, like that of Vermont Division 1 (VT-1) (Figure 2). This also holds true in Maine Divisions 1 (ME-1) and 2 (ME-2), Vermont Division 2 (VT-2), and Massachusetts Division 2 (MA-2), indicating regional cooling in the NCDC data and warming in the USHCN data.

[8] We propose that temporal variations in the mean latitude, longitude, and elevation of stations within the NCDC data are responsible for the deviations in Figure 1. Temporal changes in these variables are examined in detail for MA-2 (Figure 3). This division was selected because of the opposing 1931–2000 regression slopes of the NCDC and USHCN data, which show an overall −0.4°C (cooling) and 0.7°C (warming) trend over the 70 year period, respectively, both of which are significant at $\alpha < .05$ (Figure 1). In addition, this climate division has the highest number (6) of temporally consistent USHCN stations comprising the USHCN divisional average in the region. Before 1960, USHCN temperatures are consistently lower than NCDC, then after 1960, USHCN data are consistently higher (Figure 3A). The difference between these annual temperatures (NCDC - USHCN) shows a declining trend indicating that the NCDC data have gradually drifted with a progressive lowering of the temperatures relative to USHCN data (Figure 3B).

Table 1. Pearson Cross Correlation Statistics for Massachusetts Climate Division 2 (MA-2)

	Latitude	Elevation	Longitude
Elevation	.29		
Longitude	.51	.80	
Difference	−.57	−.58	−.68

Difference is NCDC Divisional Temperature subtracted from the USHCN Divisional Average. Each R value is significant at $\alpha < .02$.

[9] A total of 87 stations were, at one time or another, included in the NCDC divisional average for MA-2, with a low of 29 (2000) and high of 59 (1948–1950) (Figure 3F). The average latitude, longitude, and elevation of these stations also underwent gradual change with significant trends ($\alpha < .01$) toward a more northerly, westward location and a higher elevation (Figure 3C–E). In contrast, the averaged USHCN data remained constant with a mean location of 42.41°N, 71.53°W, and at an elevation of 80 m (Figure 3). In this case, the mean location of stations used to calculate the USHCN climatic division averages fall outside the mean range of latitude and longitude of the NCDC data. However, because the USHCN locations are temporally stable, the only influences to be expected from these spatial differences would be a consistent difference in mean temperature through time, e.g., USHCN consistently colder if located more northwest and/or at a higher elevation. The observed cooling in the NCDC data may therefore represent a spurious trend, whereas the actual temperature, according to the stable USHCN record, increased over the same time period within this climate division. One could argue that these differences may be spuriously induced by the adjustments made to the USHCN data as proposed by *Balling and Idso* [2002]. However, these adjustments all have a sound empirical foundation as noted above.

[10] To test the hypothesis that station variability is inducing spurious trends, Pearson correlations were deter-

Table 2. Results of Stepwise Multiple Regression Where Mean Latitude, Longitude, and Elevation are Used to Explain Deviations of the NCDC Divisional Data Set From USHCN Data for Massachusetts Climate Division 2 (MA-2)

	Variable	r^2	Probability
Step 1	Longitude	.46	<.04
Step 2	Latitude	.53	<.01
Step 3	Elevation	.54	.14

Table 3. Pearson Cross Correlation Statistics for Vermont Climate Division 3 (VT-3)

	Latitude	Elevation	Longitude
Elevation	−.29		
Longitude	−.91	.57	
Difference	.33	−.67	−.50

Difference is NCDC Divisional Temperature subtracted from the USHCN Divisional Average. Each R value is significant at $\alpha < .01$.

mined between the variables in question (Table 1). The matrix indicates that all three variables (latitude, longitude, and elevation) are negatively associated with the difference between the two data sets (NCDC minus USHCN). Hence, increasing mean latitude, longitude, and elevation are all associated with an overall decline in annual temperatures in the NCDC data relative to USHCN. The difference between the two data sets is best explained by longitude (r^2 = .46), followed by elevation (r^2 = .33), and latitude (r^2 = .32), respectively. A multiple stepwise regression was performed where annual differences between the NCDC and USHCN data are regressed against the annual mean latitude, longitude, and elevation of the NCDC data (Table 2). Results show that longitude and latitude are significant variables, together explaining 53 percent of the variance. Adding elevation to the model only increases the r^2 to .54 and is considered insignificant. We believe elevation to be the key, but as the annual mean location of stations migrated westward (increased in longitude), the stations also climbed in elevation so that these two variables largely share the same explanatory power. Remaining unexplained variance may be attributable to the lack of adjustments in the NCDC data, e.g., changes in instrumentation, and urban warming biases and could also be related to changes in the aspect of the stations. Other climate divisions may have different associations between these variables and temperature, that relate to the physical geography of the specific climate division. For example, findings from MA-2 are largely validated with results from VT-3, which has an overall increase (opposite that of MA-2) in the difference between the two datasets (NCDC minus USHCN) (e-supplement 2). Here, NCDC data had significant temporal shifts (at $\alpha < .01$) in mean station location northward, eastward, and, most importantly, downward in elevation, which artificially increased temperatures through time in the NCDC data. The decrease in mean elevation of the NCDC stations alone explains 45 percent of the difference between NCDC and USHCN in VT-3 (Table 3).

[11] Figures 1 and 2 also show a regional temperature change for New England, based on the two areally weighted divisional data sets. The NCDC data show a regional decline in temperature by 0.1°C, whereas the USHCN data shows an increase of 0.4°C. The USHCN results are consistent with the *Intergovernmental Panel on Climate Change* [2001] and *Karl et al.* [1996] for New England, which strongly suggests that the region has warmed to an even larger extent than that documented by the *New England Regional Assessment Group* [2001] who used the NCDC climate divisional data to analyze statewide and regional trends. Even at the climate divisional level, the USHCN pattern is more geographically cohesive in that no division has cooled and the region of significant warming are all contiguous divisions in the southeastern portion of the region (Figure 1). This seems more logical than the NCDC data pattern where adjacent divisions have significant trends, but in opposing directions, e.g., MA-1 and MA-2.

4. Summary and Conclusions

[12] Given the method of construction of the frequently used NCDC climate divisional data sets, we examined potential impacts of the systematic redistribution of stations within a division on trends in temperature. Comparison of the NCDC divisional data to USHCN-derived divisional data shows that the two data sets do not always agree. Changing the annual mean latitude, longitude, and perhaps most importantly, the elevation of stations within a division, can have significant affects when analyzing trends. It is likely that similar biases exist regarding precipitation trends. All divisions may not behave exactly like MA-2, but the mechanistic explanation of the difference between datasets shows one clear problem, which is sufficient to cast doubt on use of all NCDC divisional data. Detecting and eliminating problems in other climate divisions now becomes the obligation of the investigator using these NCDC divisional data for time series analysis.

[13] **Acknowledgments.** This research was supported by NOAA Grant No. NA17RP1488.

References

Balling, R. C., Jr., and C. D. Idso, Analysis of adjustments to the United States Historical Climatology Network (USHCN) temperature database, *Geophys. Res. Lett.*, *29*(10), 1387, 10.1029/2002GL014825, 2002.

Easterling, D. R., et al., (Eds.), United States Historical Climatology Network (U.S. HCN) monthly temperature and precipitation data, *Rep. ORNL/CDIAC-87*, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., Oak Ridge, Tenn., 1996.

Guttman, N. B., and R. G. Quayle, A historical perspective of U.S. climate divisions, *Bull. Am. Meteorol. Soc.*, *77*, 293–303, 1996.

Intergovernmental Panel on Climate Change, *The Third Assessment Report of Working Group I of the Intergovernmental Panel on Climate Change* (*IPCC*), edited by J. T. Houghton et al., 881 pp., Cambridge Univ. Press, New York. 2001.

Karl, T. R., and C. W. Williams Jr., An approach to adjusting climatological time series for discontinuous inhomogeneities, *J. Clim. Appl. Meteorol.*, *26*, 1744–1763, 1987.

Karl, T. R., et al., A model to estimate the time of observation bias with monthly mean maximum, minimum, and mean temperatures for the United States, *J. Clim. Appl. Meteorol.*, *25*, 145–160, 1986.

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

Karl, T. R., et al., Indicies of climate change for the United States, *Bull. Am. Meteorol. Soc.*, *77*, 279–292, 1996.

Keim, B. D., et al., Long-term trends in precipitation and runoff in Louisiana, USA, *Int J. Climatol.*, *15*, 531–541, 1996.

Leathers, D. J., A. J. Grundstein, and A. W. Ellis, Growing season moisture deficits across the northeastern United States, *Clim. Res.*, *14*, 43–55, 2000.

New England Regional Assessment Group, *Preparing for a Changing Climate: The Potential Consequences of Climate Variability and Change. The New England Regional Assessment Overview*, 96 pp., U.S. Global Change Res. Program, Univ. of N.H., Durham, 2001.

Quayle, R. G., D. R. Easterling, T. R. Karl, and P. Y. Hughes, Effects of recent thermometer changes in the cooperative station network, *Bull. Am. Meteorol. Soc.*, *72*, 1718–1724, 1991.

T. G. Huntington, United States Geological Survey, Augusta, ME, USA.

B. D. Keim, Southern Regional Climate Center, Louisiana State University, Baton Rouge, LA 70803, USA.

A. M. Wilson and C. P. Wake, Climate Change Research Center, University of New Hampshire, Durham, NH 03824, USA.