DRAFT

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

Barry D. Keim and Michelle Russo Fischer
Louisiana State University
Louisiana Office of State Climatology

Adam M. Wilson
University of New Hampshire

Abstract

The United States Climate Division datasets may contain spurious precipitation trends, induced by the method in which climate divisional data are created. This study investigated how the redistribution of NCDC stations within a climate division affected annual trends in precipitation. Divisional data from NCDC were compared to USHCN data for the 15 climate divisions of New England. Results showed that trends in the datasets are not always the same. We conclude that changing mean latitude, longitude, and elevation of stations within a division through time can affect, and even induce, trends in the time series. Trends found here were mostly increasing through time, but the rates of change differed dramatically in some cases. Two climate divisions did have opposing trends between datasets. It appears that other factors, such as distance from the coast and changing aspect of stations, may also influence these trends.

Introduction

The National Climatic Data Center (NCDC) divided the 48 contiguous United States into 344 climate divisions, which are generally considered climatically homogeneous regions. For each division, the NCDC calculated monthly means of temperature and precipitation, as well as the Palmer Drought Severity Index and Palmer Hydrological Drought Index back to 1895 (Guttman and Quayle, 1996). Although these data have many useful applications, some argue that they should be used with caution (Muller *et al*., 1990; Guttman and Quayle, 1996), especially with decadal- to century-scale climate change studies (Keim *et al*., 2003).

This research is an extension of *Keim et al*. (2003) who investigated differences in annual temperature trends between the NCDC climate division database and the United States Historical Climate Network (USHCN) for the climate divisions of New England. That study found spurious trends in annual temperatures in some climate divisions, relative to the USHCN, by virtue of the systematic redistribution of stations within these divisions. As a result, this study examines potential spurious trends in climate division annual precipitation in New England, since the redistribution of stations within a climate division could also affect precipitation averages. We postulate that redistributing stations to varying elevations, and/or from or toward a coastline, will have impacts on the climate division precipitation series.

Data and Methods

Annual times series are generated by averaging monthly precipitation data for each of the 15 New England climate divisions from 1931 to 2000. Data prior to 1931 are not examined as a result of being estimated from statewide averages (Guttman and

Quayle, 1996). The NCDC climate divisional annual averages (hereafter referred to NCDC) are then compared to USHCN precipitation data, which serves as the control group in this analysis. USHCN data have undergone intensive quality control, have long station histories and limited station relocations. Annual time series of USHCN "divisional data" are calculated by averaging monthly averages from stations within a climate division, as in *Keim et al*. (2003) for temperature. Only stations with continuous monthly records from January 1931 through December 2000 were included. There were fewer USHCN precipitation stations that met these requirements than for the temperature analysis. This resulted in two climate divisions having no USHCN precipitation data for comparison, Massachusetts Division 1 (MA-1) and Connecticut Division 3 (CT-3), and several divisions only had one USHCN station within its boundaries. Despite the small number of USHCN stations available, they should provide more reliable and accurate accounting of long-term climate change in the region (Trombulak and Wolfson, 2004). The NCDC and the USHCN datasets are analyzed using linear regression, Pearson correlation, and multiple stepwise regression.

Comparison of Trends in Precipitation

Linear precipitation trends for the 15 New England climate divisions show that NCDC and USHCN divisional datasets do not always agree (Figure 1). Potential problems exist when the NCDC and USHCN slopes for the same climate division drift apart through time (e.g., VT-1). In most cases, both datasets had increasing trends, but with differing rates of change (Figure 2). In some cases the trends in the two datasets ran parallel, but the magnitudes of the precipitation data differed, e.g. NH-2. This suggests that the two datasets vary dramatically in elevation (or continentality), but the overall temporal trends have not been contaminated.

Connecticut Division 2 (CT-2) had opposing linear trends (Figure 3A). Temporal changes in mean elevation, latitude, and longitude of the NCDC stations within the division were investigated to determine their impact on the trend, relative to the USHCN data. The difference in these annual precipitation totals (NCDC – USHCN) showed an increasing trend (Figure 3B), indicating that NCDC average precipitation has increased significantly relative to USHCN data at $\alpha < 0.01$.

A total of 67 stations were included in the NCDC divisional average for CT-2 over the study period, varying from 6 to 17 at any given time. The average latitude, longitude, and elevation also changed significantly over time ($\alpha < .01$). Through time, the mean station location drifted south and east, and to a lower elevation (Figure 3C-E). The USHCN station, Storrs, which is located in the northeastern portion of the climate division, remained stationary at 41.80˚N, 72.25˚W, at an elevation of 198 meters above msl. The shift in mean station location of the NCDC dataset may be the reason for the opposing trends.

Pearson correlations were calculated between the difference (NCDC-USHCN) and latitude, longitude, and elevation (Table 1). All three variables were negatively associated with the difference between datasets. This indicates that decreasing mean latitude, longitude, and elevation were all associated with increasing precipitation in the NCDC data. Elevation ($r^2 = .30$) best explained the difference between the two datasets, followed by latitude ($r^2 = .21$) and then longitude ($r^2 = .16$). A multiple stepwise regression analysis was performed, which takes into account multi-colinearity between variables (Table 2). Results showed that latitude and longitude together explained 48 percent of the variance. Elevation was insignificant because of its high correlation with latitude (Table 1), suggesting that as the mean station location moved south, it also decreased in elevation. It is somewhat counter-intuitive that rainfall would increase at stations lower in elevation. However, in this case, mean station location migrated closer to the Connecticut coastline, which may explain the increase in precipitation. Some of the remaining unexplained variance may be attributed to changing aspects of the NCDC stations through time.

With VT-1, elevation and latitude are negatively associated with the difference (NCDC- USHCN); however, longitude was positively associated (Table 3). This suggests that decreasing mean elevation and latitude, but increasing mean longitude were associated with increasing precipitation in the NCDC dataset. The mean NCDC station location moved south and west and to a lower elevation. Elevation ($r^2 = .40$) was the most significant variable, followed by latitude ($r^2 = .30$) and longitude ($r^2 = .30$) (Table 3). Results of a multiple stepwise regression showed that elevation and longitude together explained 45 percent of the variance. We surmise that as the mean location of stations moved south and to a lower elevation, a more southerly aspect of the stations served to increase the rainfall in the NCDC dataset.

Regional precipitation change was examined by areally weighting each division within the two datasets (Figure 2), minus MA-1 and CT-3. Both time series showed regional increases in precipitation, but the USHCN data did so at a slightly higher rate. USHCN data depict a less cohesive pattern geographically than the NCDC data.

Summary and Conclusions

Annual precipitation trends in the NCDC climate division data were compared with USHCN data, and sometimes trends within the same climate division do not agree. The migration of the mean location of the stations within these climate divisions is shown to have impacts on long-term precipitation trends relative to the USHCN stations. Results indicate that caution should be taken when using NCDC climate division precipitation data in decadal- to century-scale climate change studies.

References

Guttman, N.B., and R.G. Quayle (1996), A historical perspective of U.S. climate divisions, *Bull. Am. Meteorol. Soc.*, *77*(2), 293-303.

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:10.1029/2002GL016295.

Muller, R.A., B.D. Keim, and J.L. Hoff (1990), Application of climatic divisional data to flood interpretations – an example from Louisiana, *Phys. Geog.*, *11*(4), 353-362.

Trombulak, S.C., and R. Wolfson (2004), Twentieth-century climate change in New England and New York, USA, *Geophys. Res. Lett.*, *31*, L19202, doi:10.1029/2004GL020574.

Figure 1. Time series and linear trends of annual precipitation comparing NCDC (solid lines) and USHCN (dashed lines) divisional data for selected New England Climate divisions, 1931-2000. Pearson-R correlation values are presented in the upper left of each panel.

Figure 2. Climate division of the New England States, linear precipitation trends (cm), from 1931-2000 by division (italicized = positive, bold = negative, boxed values = significant at $\alpha < .05$) for A) NCDC divisional data, and B) USHCN divisional data. Number of USHCN stations included by division.

Figure 3. Connecticut Division 2 (CT-2) annual time series of A) NCDC and USHCN divisional precipitation; B) precipitation difference between NCDC and USHCN; C) mean latitude on 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.

Table 1. Pearson Cross Correlation Statistics for Connecticut Climate Division 2 (CT-2)

	Latitude	Elevation	Longitude
Elevation	**.83**		
Longitude	**-.24**	.11	
Difference	**-.46**	**-.55**	**-.40**

Difference in USHCN Divisional Average subtracted from the NCDC Divisional Precipitation. Each bold R value is significant at $\alpha < .05$.

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 Connecticut Climate Division 2 (CT-2).

	Variable	r^2	Probability
Step 1	Longitude	.42	.0005
Step 2	Latitude	.48	.0047

Table 3. Pearson Cross Correlation Statistics for Vermont Climate Division 1 (VT-1)

	Latitude	Elevation	Longitude
Elevation	**.69**		
Longitude	**-.68**	**-.56**	
Difference	**-.55**	**-.63**	**.55**

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

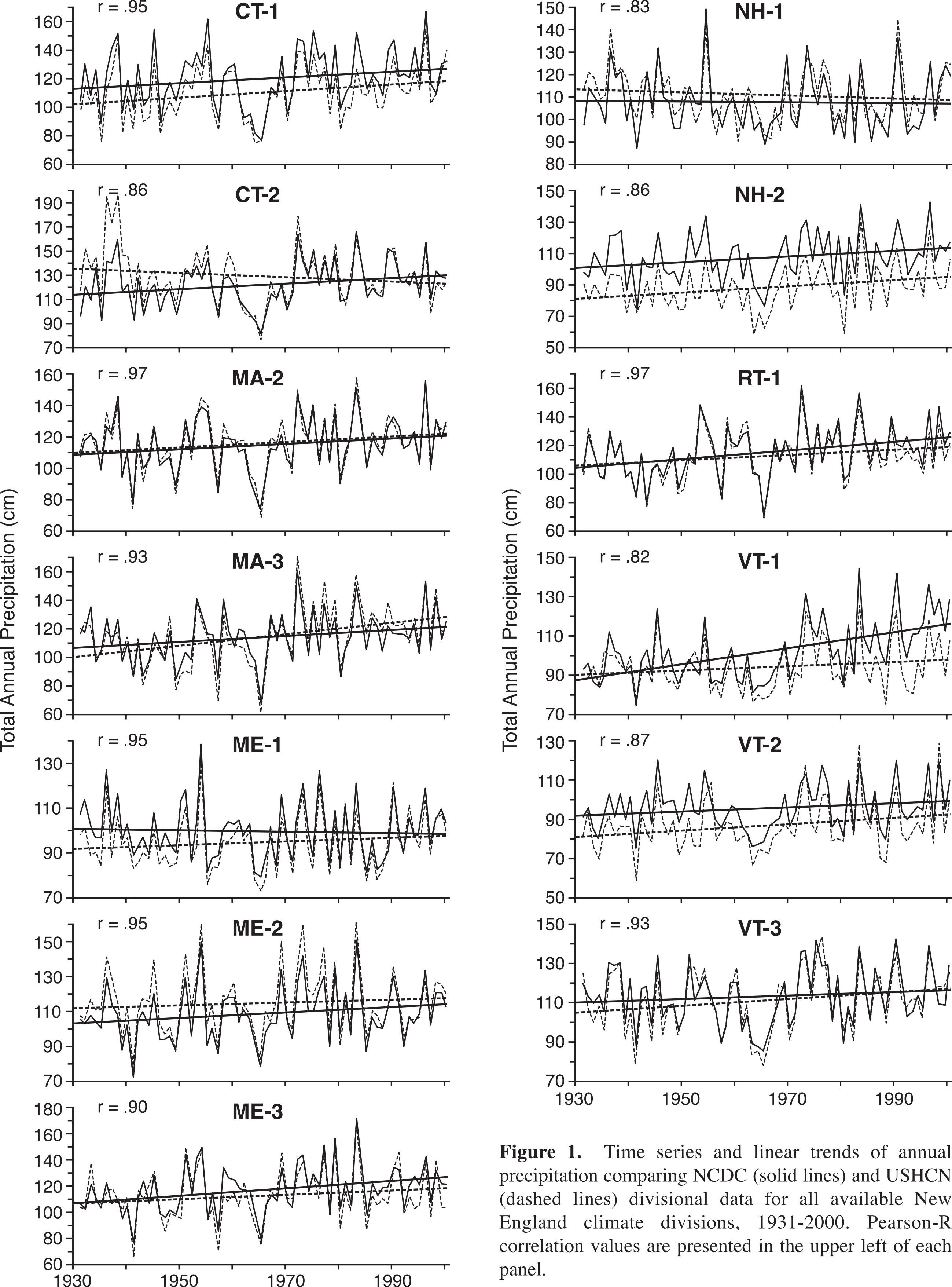

Figure 1. Time series and linear trends of annual precipitation comparing NCDC (solid lines) and USHCN (dashed lines) divisional data for all available New England climate divisions, 1931-2000. Pearson-R correlation values are presented in the upper left of each panel.

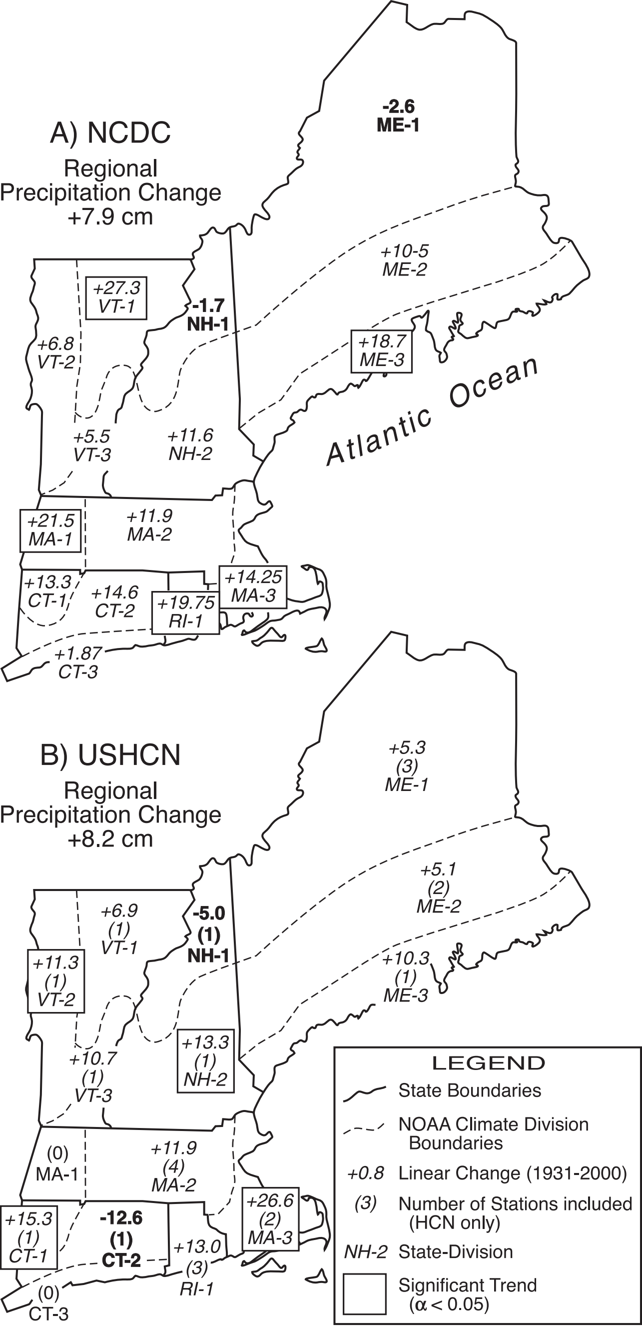

Figure 2. Climate division of the New England States, linear precipitation trends (cm) from 1931-2000 by division (italicized = positive, bold = negative, boxed values = significant at $\alpha < .05$) for A) NCDC divisional data, and B) USHCN divisional data. Number of USHCN stations included by division.

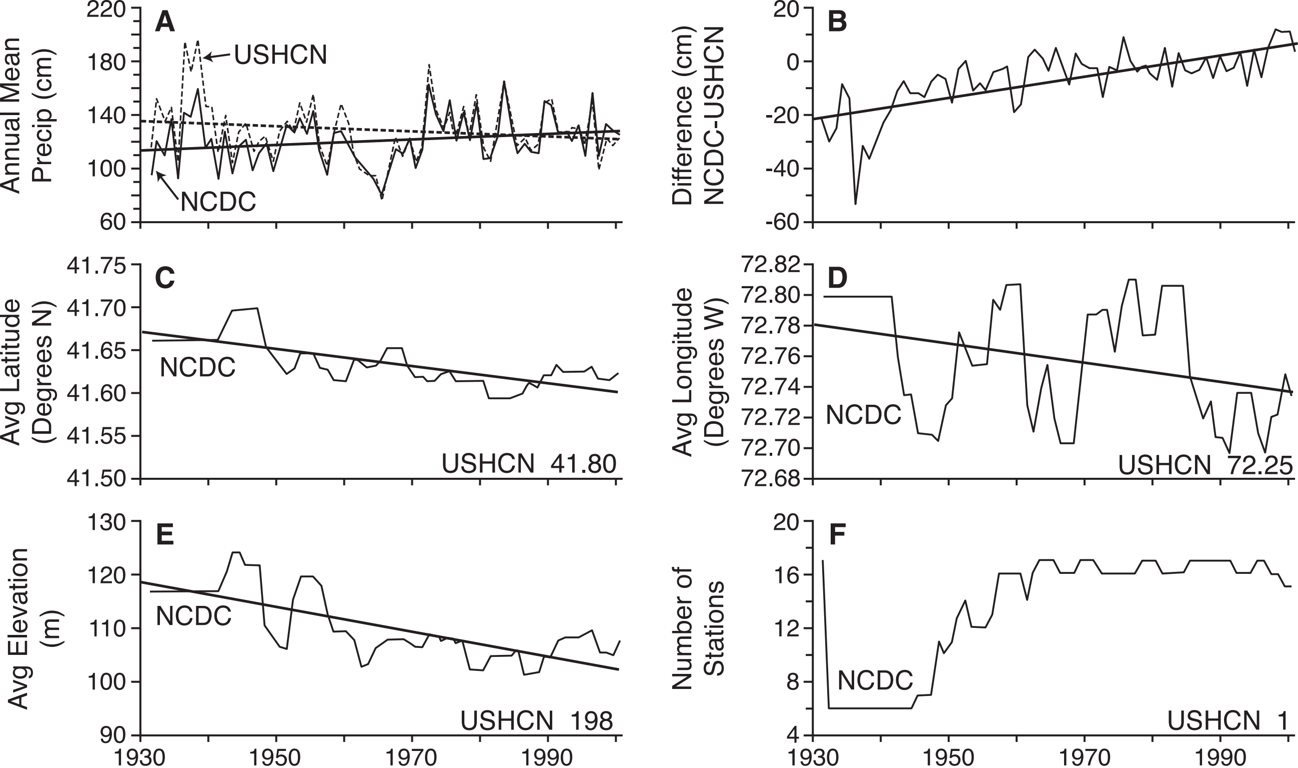

Figure 3. Connecticut Division 2 (CT-2) annual time series of A) NCDC and USHCN divisional precipitation; B) precipitation difference between NCDC and USHCN; C) mean latitude on NCDC stations; D) mean longitude of NCDC stations; E) mean elevation of NCDC stations, and F) number of stations included in the NCDC divisional average.