6B.5

Transitioning from the traditional divisional dataset to the Global Historical Climatology Network-Daily gridded divisional dataset

Chris Fenimore, Derek Arndt, Karin Gleason, and Richard R. Heim Jr.
NOAA/NESDIS/NCDC
Asheville, NC

1. INTRODUCTION

The National Climatic Data Center's (NCDC) Climate Monitoring Branch (CMB) uses climate division data to identify trends, generate ranks and summarize significant climate events in its monthly suite of temperature, precipitation and drought products. These products include a variety of maps and graphics, some of which are contained within the State of the Climate (SotC) report. These maps and graphics are produced from a well established network of surface stations that are comprised of National Weather Service First Order and Cooperative Observer Network (COOP) stations. These surface stations are used to generate average temperature and precipitation for climate divisions across the country and will be referenced in this paper as the traditional climate division dataset (and abbreviated, for the purposes of this manuscript, TCDD).

By 2013, NCDC plans to transition to a newly-developed dataset to compute divisional values. This new dataset is based on the Global Historical Climatological Network-Daily (GHCN-D). NCDC recognizes that many user and applied-science communities have built products and operations upon the TCDD. This manuscript and its related presentation are part of our strategy to make the transition to the new divisional dataset as transparent and smooth as possible.

The construction and methodology of the new gridded divisional dataset (for the purposes of this manuscript, GrDD) will be described in a separate (peer-reviewed) manuscript (in preparation) before it is adopted as CMB's official divisional dataset. In this extended abstract, we describe the temporal and spatial effects this transition will have on the temperature, precipitation and drought indices used in the CMB monitoring products.

Corresponding author address: Chris Fenimore, NOAA / NESDIS / NCDC, 151 Patton Avenue, Asheville NC, 28801.
E-mail: chris.fenimore@noaa.gov

2. BACKGROUND

For many decades, the TCDD was the only long-term temporally and spatially complete database from which to generate historical climate analyses (1895 to the present) for the contiguous United States (CONUS). It was originally developed for climate-division, statewide, regional, national, and population-weighted monitoring of drought, temperature, precipitation, and heating/cooling degree day values. Since the database was at the divisional spatial scale, it naturally lent itself to agricultural and hydrological applications.

In the TCDD, monthly station temperature and precipitation values are computed from the daily observations. There are 344 climate divisions in the contiguous U.S. The monthly values for all of the COOP stations in each division are averaged to compute divisional monthly temperature and precipitation averages/totals. The divisional values are weighted by area to compute statewide values and the statewide values are weighted by area to compute regional values. This is valid for values computed from 1931 to the present. For the 1895-1930 period, statewide values were computed directly from stations within each state. Divisional values for this early period were computed using a regression technique against the statewide values (Guttman and Quayle, 1996).

The GHCN-D 5km gridded divisional dataset is based on a similar station inventory as the TCDD; however, new methodologies are used to compute temperature, precipitation, and drought for United States climate divisions. These new methodologies include the transition from a station-based to a grid-based calculation, the inclusion of many more stations from the pre-1930s, and the use of NCDC's modern array of quality control algorithms. These are expected to improve the data coverage and the quality of the dataset, while maintaining the current product stream.

The GrDD is designed to address the following general issues inherent in the TCDD:

1. For the TCDD, each divisional value from 1931-present is simply the arithmetic average of the station data within it, a computational practice that results in a bias when a division is spatially undersampled in a month (e.g., because some stations did not report) or is climatologically inhomogeneous in general (e.g., due to large variations in topography).
2. For the TCDD, all divisional values before 1931 stem from state averages published by the U.S. Department of Agriculture (USDA) rather than from actual station observations, producing an artificial discontinuity in both the mean and variance for 1895-1930 (Guttman and Quayle, 1996).
3. In the TCDD, many divisions experienced a systematic change in average station location and elevation during the 20th Century, resulting in spurious historical trends in some regions (Keim et al., 2003; Keim et al., 2005; Allard et al., 2009).
4. Finally, none of the TCDD's station-based temperature records contain adjustments for historical changes in observation time, station location, or temperature instrumentation, inhomogeneities which further bias temporal trends (Peterson et al., 1998).

The GrDD's initial (and more straightforward) improvement is to the underlying network, which now includes additional station records and contemporary bias adjustments (i.e., those used in the U.S. Historical Climatology Network version 2; Menne et al., 2009).

The second (and far more extensive) improvement is to the computational methodology, which now addresses topographic and network variability via climatologically aided interpolation (Willmott and Robeson, 1995).

The outcome of these improvements is a new divisional dataset that maintains the strengths of its predecessor while providing more robust estimates of areal averages and long-term trends.

3. METHODS

Simultaneous "snapshots" were taken of the divisional temperature and precipitation data for each of the two (TCDD and GrDD) datasets. For larger-scale areas (i.e., states and climate regions), divisional values were composited using the same area-weighted methods as are used operationally for the TCDD. This area-weighting method will continue to be used operationally when the GrDD is commissioned. The period of comparison was January 1895 through December 2009.

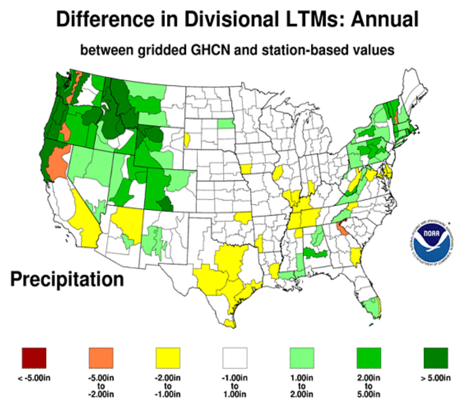
Basic statistical comparisons were used to calculate the difference in the long term mean (GrDD minus TCDD) and the difference in long-term trend (GrDD slope minus TCDD slope). These were aggregated across all climate divisions, states, and climate regions. Geographic Information System (GIS) software was also used to help determine elevation characteristics between the two datasets. Combining spatial analysis tools with a digital elevation model (DEM), we determined spatial patterns and differences in the two datasets.

A visualization toolkit was also created to help users examine snapshots of both datasets for the comparison period (i.e., through December 2009). The tool allows the user to select criteria which are of interest to her/him, and investigate the comparisons themselves. Parameters included in the toolkit are temperature, precipitation, and a variety of drought indices. Changes in monthly, seasonal and annual variability can be examined through the use of the interactive time series plots. In addition, slope (trend) values by decade and 30-year period may also be added to the output plots. This allows the user to take a closer look at the behavior of the data at a variety of smaller time scales throughout the record.

4. IMPACTS (PRELIMINARY FINDINGS)

Using the methods mentioned above, we identified the climate divisions, states and regions for which the data are impacted the most by this transition. In the aggregate, the GrDD dataset is slightly cooler and wetter when compared to the TCDD, but significant regional differences are apparent (Fig. 1). Typically, the greatest deviations occur in areas with large elevation differences, divisions which border Canada and Mexico, and divisions lacking a large network of stations early in the record.

a)



b)

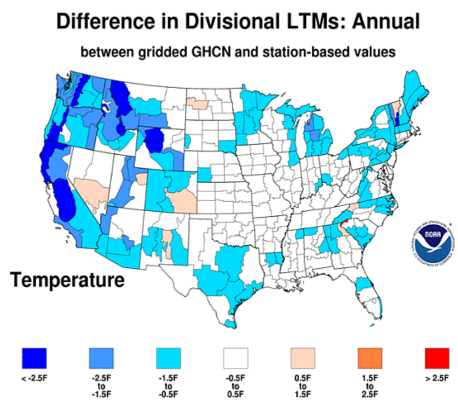
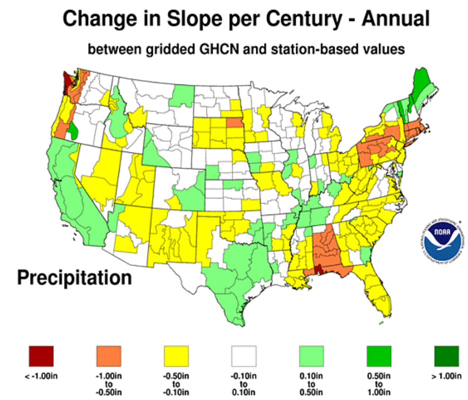


Figure 1. Difference in long-term means (LTMs) between the Gridded GHCN-D (GrDD) and Traditional methods of determining climate division (a) precipitation (inches), and (b) temperature (degrees Fahrenheit). Mountainous climate divisions in the West tend to show cooler and wetter LTMs in the GrDD method. GHCN-D closely follows the traditional divisional dataset east of the Rockies.

Divisions with more variable terrain reflect greater differences in values. The GrDD is more equipped to represent higher-elevation information, both through the inclusion of high-elevation (e.g., SNOTEL) stations, and through algorithmic treatment of topography. As a result high-terrain states like Idaho, Montana, Oregon, Washington, and Wyoming reflect a wetter and cooler annual period. The difference in the long-term trend for precipitation (GrDD trend minus TCCD trend, measured in inches per century) shows more geographical (as in spatial) patterns, as opposed to the more terrain-oriented differences in the precipitation LTMs.

a)



b)

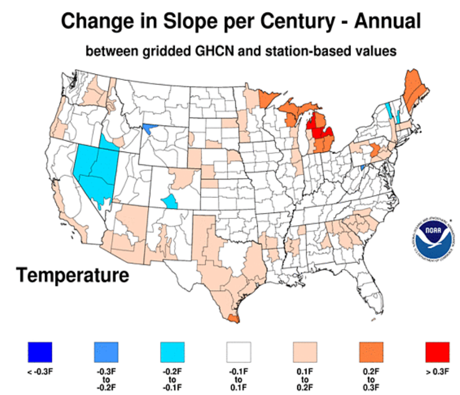


Figure 2. Difference in slope (trend) between the Gridden GHCN-D (GrDD) and Traditional (TCDD) methods of calculating (a) precipitation (inches per century) and (b) temperature (degrees Fahrenheit per century). Compared to the TCDD, the GrDD has a more positive precipitation trend in Maine, drier in Alabama, warmer in Michigan, and cooler in Nevada. These reflect peculiarities related to the regression methods used to obtain the averages prior to 1930 (see text for details).

Much of the eastern U.S. has a smaller (less positive) precipitation trend in the GrDD, compared to the TCDD; this is enhanced quite distinctly for certain states (Alabama, Pennsylvania, Massachusetts, Rhode Island and Connecticut). (Fig. 2a).

The largest differences in the annual temperature LTMs are found in mountainous areas of the West, where the GrDD are more than 2.5°F cooler than the TCDD (Fig. 1b). In the majority of the southern, central and eastern U.S., the GrDD is warmer during the most recent

a) b)

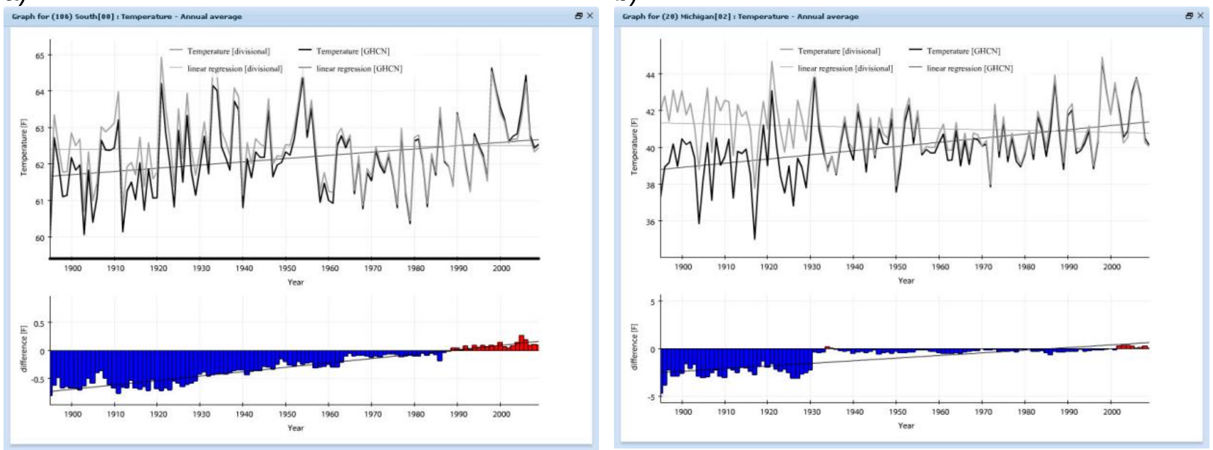


Figure 3. Average annual temperature comparison for (a) the South climate region (a composite of Arkansas, Kansas, Louisiana, Mississippi, Oklahoma, and Texas) and (b) Michigan Climate Division 2. The bottom panel in each shows the difference (GrDD minus TCDD) between the two data sets. Please note that the scales for each panel are very different. During the past 20-years, the GrDD is warmer than the TCDD (red bars). For Michigan CD2, the impact of eliminating the pre-1930s regression methods can be clearly seen. The tool used to construct these comparisons is publicly available at: http://nidis1.ncdc.noaa.gov/GHCNViewer/

20 years, as reflected in the trend comparison (Fig. 3a). For much of the rest of the country, the differences between the two datasets are minimal. As with precipitation, but not as dramatically, several states appear to have distinct outlier characteristics. Michigan and Maine have the strongest increases in the warming trend (Fig. 2b); this reflects the large difference resulting from the replacement of the pre-1930s regression methods with station-based information. Nevada, on the other hand, shows a distinct decrease in the temperature trend between datasets.

For long-term means (LTMs) the largest differences in both temperature and precipitation are typically found in climate divisions with significant topography. The following general patterns are noted:

- For divisions with average station elevation lower than the average topographical elevation ("stations in the valleys"), the GrDD typically display cooler and wetter LTMs.
- For divisions with average station elevation higher than the average topographical elevation ("stations on the ridges"), the GrDD presents warmer and drier LTMs.

It should be noted that the former scenario ("stations in the valleys") is much more common than the latter ("stations on the ridges").

There appears to be some border effect on temperature LTMs climate divisions that border Canada east of North Dakota. The LTMs for the GrDD set are generally cooler. This may reflect a cross-border "reach" effect in that GrDD utilizes Canadian stations while TCDD does not. An average of 20 neighbors is used in the anomaly interpolation to create the GrDD, so the radius will vary with the density of the station network in each year and month. However, no definitive statement about the cause (or coincidence) is available at the time of this publication.

For some divisions, the replacement of regression techniques with more station-dense values in the GrDD dataset had the effect of "cooling" the pre-1930's data. This will have an effect on the long-term trend of the slope for these divisions. For example, in Michigan Division 02, GrDD runs much cooler (2-4°F or 1-2°C) in the pre-1930's than the TCDD, resulting in pre-1930's temperatures that are cooler than the 1950's temperatures (Fig. 3b). This period is actually warmer than the 1950's in the TCDD dataset. This change affects the linear regression slope of the annual temperature, changing it from a negative slope for the TCDD to a positive slope for the GrDD (Fig. 3b).

5. CONCLUSION

The GrDD is a modern, quality-assured database that improves upon the historical monthly temperature and precipitation data that are currently available with the TCDD. Use of these data will improve our understanding of observed changes in climate across the contiguous U.S. (CONUS). Regression techniques used to derive pre-1930 divisional data in the TCDD have been replaced by real station data in the GrDD, improving comparisons made to early 20th Century data. Because of the different algorithms used, slight variances in temperature and precipitation averages may be seen throughout the data record. The average change in trend was about 0.06°F per century. The annual temperature trend in each division is between -0.3°F and +0.3°F per century and only three climate divisions had differences in their mean larger than 0.3°F per century. In terms of precipitation, the mean change in slope is slightly negative for the annual period.

6. REFERENCES

Allard, J**.,** B.D. Keim**,** J.E. Chassereau, D. Sathiaraj. 2009. Spuriously induced precipitation trends in the southeast United States. *Theoretical and Applied Climatology*. DOI: 10.1007/s00704-008-0021-9.

Guttman, N. V. and R. G. Quayle, 1996: A historical perspective of U.S. climate divisions. *Bull. Amer. Meteor. Soc.,* 77, 293-303.

Karl, T.R., C.N. Williams, Jr., P.J. Young, and W.M. Wendland, 1986: A model to estimate the time of observation bias associated with monthly mean maximum, minimum, and mean temperature for the United States, *J. Climate Appl. Meteor*., 25, 145-160.

Karl , T.R., and Koss, W.J., 1984: Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895-1983. *Historical Climatology Series* 4-3, National Climatic Data Center, Asheville, NC, 38 pp.

Keim, B. D., A. Wilson, C. 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

Keim, B.D., M.R. Fischer, and A.M. Wilson, 2005: Are there spurious precipitation trends in the United States Climate Division database? *Geophys. Res. Lett*., **32**, L04702, doi: 10.1029/2004GL021985.

Menne, M.J., C.N. Williams, and R.S. Vose, 2009: The United States Historical Climatology Network Monthly Temperature Data - Version 2. *Bulletin of the American Meteorological Society*, **90**, 993-1107.

Peterson, T.C., T.R. Karl, P.F. Jamason, R. Knight, and D.R. Easterling, 1998: The first difference method: maximizing station density for the calculation of long-term global temperature change. *J. Geophys. Res., Atmosphere*s, **103 (D20)**, 25967-25974.

Willmott, C.J. and S.M. Robeson, 1995. Climatologically aided interpolation (CAI) of terrestrial air temperature. *International Journal of Climatology*, **15**(2), 221-229.