

## An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network

Russell S. Vose, Claude N. Williams Jr., Thomas C. Peterson, Thomas R. Karl, and David R. Easterling

National Climatic Data Center, Asheville, North Carolina, USA

Received 7 July 2003; revised 8 September 2003; accepted 10 September 2003; published 23 October 2003.

[1] The U.S. Historical Climatology Network (HCN) database contains statistical adjustments that address historical changes in observation time at each observing station in the network. A paper in 2002 suggested that these adjustments cause HCN temperature trends to be “spuriously” warm relative to other datasets for the United States. To test this hypothesis, this paper evaluates the reliability of these “time of observation bias” adjustments in HCN. The results indicate that HCN station history information is reasonably complete and that the bias adjustment models have low residual errors. In short, the time of observation bias adjustments in HCN appear to be robust. *INDEX TERMS*: 3309 Meteorology and Atmospheric Dynamics: Climatology (1620); 1610 Global Change: Atmosphere (0315, 0325); 3399 Meteorology and Atmospheric Dynamics: General or miscellaneous. *Citation*: Vose, R. S., C. N. Williams Jr., T. C. Peterson, T. R. Karl, and D. R. Easterling, 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, 2003.

### 1. Introduction

[2] The U.S. Historical Climatology Network (HCN) database [Easterling *et al.*, 1996] is commonly used to quantify climate change because it contains adjustments that account for historical variations in observation time [Karl *et al.*, 1986], station location/instrumentation [Karl and Williams, 1987], and population growth [Karl *et al.*, 1988]. In a comparative analysis of the conterminous United States, Balling and Idso [2002] determined that adjusted HCN temperature trends for the past 30 years were slightly more positive than those derived from other datasets [e.g., Jones, 1994; Christy *et al.*, 2000]. This led to the hypothesis that the HCN contained a “spurious” warming that likely resulted from the adjustments for historical variations in observation time. Given this supposition, the purpose of this paper is to demonstrate the reliability of the adjustments for this “time of observation” bias in HCN. The analysis consists of three parts. First, the accuracy of HCN station history information is verified using an independent source of metadata. Second, the predictive skill of the adjustment approach is evaluated using data from 500 hourly stations over the period 1965–2001 (the approach was originally developed using data from 79 stations over

the period 1957–64). Finally, HCN temperature trends are shown to be consistent with those in Jones [1994].

### 2. Background

[3] The majority of weather stations in the U.S. Cooperative Observing Network (and therefore in HCN) are staffed by volunteers. Consequently, the network has no mandatory time at which daily measurements must be taken. Most individuals prefer observing times other than midnight, resulting in an observation day that differs from the standard calendar day. For example, at a station where the volunteer reads the thermometers at 0800 LST, the observation day extends from 0800 LST the previous day to 0800 LST on the current day. At a station where the volunteer reads the thermometers at 1700 LST, the observation day starts and ends 9 hours later. Nevertheless, the observations at both stations are recorded for the same calendar day.

[4] When the observation day differs from the calendar day, a “carry over” bias of up to 2.0°C is introduced into monthly mean temperatures. This bias occurs when atmospheric conditions cause a temperature from one day to be ascribed to the following day. For instance, suppose an observer reads the maximum and minimum thermometers at 1700 LST on April 1, then a cold front passes through the area overnight. If the temperature on April 2 never exceeds the value at 1700 LST on April 1 (when the thermometers were last reset), then the recorded maximum will actually be the temperature at 1700 LST on April 1. This temperature will be higher than if the 24-hour measurement ended at midnight, and because the monthly mean is computed by averaging the daily maximums and minimums, the mean for April will likewise be artificially high. In general, this carry-over phenomenon results in a warm bias for observation days ending in the afternoon and a cool bias for those ending in the morning.

[5] Non-calendar day observations also result in a “drift” bias of up to 1.5°C in monthly means. This bias results from the inclusion of data from the end of the previous month, as well as the exclusion of data from the end of the current month, in the computation of a monthly mean. For example, at a station where the observation day ends at 0700 LST, the average for April will be based upon data from the last 17 hours of March 31 through the first 7 hours of April 30. Because the end of March is usually cooler than the end of April, the average for April will typically be lower than if the station’s observation day ended at midnight. In general, drift bias is most pronounced during transition seasons and for morning observation times.

[6] There has been a systematic change in the preferred observation time in the U.S. Cooperative Observing Net-

work over the past century. Prior to the 1940s most observers recorded near sunset in accordance with U.S. Weather Bureau instructions, and thus the U.S. climate record as a whole contains a slight warm bias during the first half of the century. A switch to morning observation times has steadily occurred during the latter half of the century to support operational hydrological requirements, resulting in a broad-scale nonclimatic cooling effect. In other words, the systematic change in the time of observation in the United States in the past 50 years has artificially reduced the temperature trend in the U.S. climate record [Hansen *et al.*, 2001].

### 3. HCN Metadata

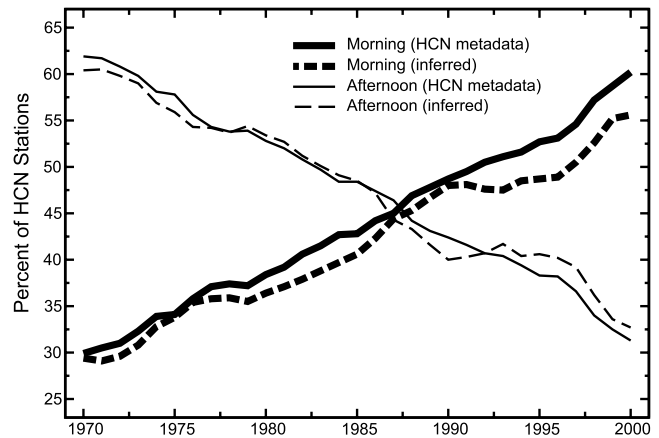
[7] This time of observation bias is addressed in the adjusted HCN using the method described in Karl *et al.* [1986]. This adjustment approach requires as input an a priori knowledge of the actual observation time in each year at each HCN station. If this information is inaccurate for a significant number of stations, then the resulting adjustments would be erroneous for a large portion of the network, thus impacting national-scale trend analyses. Given this possibility, the accuracy of HCN station history information is verified here using an independent source of metadata.

[8] The verification consists of a simple comparison of HCN metadata with time of observation metadata that were “inferred” using the method described in DeGaetano [1999]. This technique infers the observation time at annual resolution on a station-by-station basis using daily maximum and minimum temperature data. The estimated observation time is morning (0600–0800 LST), afternoon (1600–1900 LST), or midnight (0000 LST); for comparative purposes HCN metadata were thus aggregated in a similar fashion. According to DeGaetano [1999], the procedure correctly classified the observation schedule in 90% of the station-years tested in a network of over 1000 U.S. stations during a 40-year period, suggesting that the inferred metadata are reasonably accurate. The daily data for the analysis were extracted from the U.S. HCN daily database [Easterling *et al.*, 1996], and the comparison focused on the last three decades because that was the period in which Balling and Idso [2002] noted a difference in trends between the adjusted HCN and other datasets.

[9] Figure 1 depicts the percentage of HCN stations with morning and afternoon observing times according to inferred metadata and actual metadata (midnight is not depicted because it accounts for less than 10% of the total). In general, the inferred metadata closely agree with the actual metadata on the national scale. Both sources indicate that the percentage of afternoon readers decreased over the past half century while the percentage of morning observers increased, a change which has been widely documented for some time. As described earlier, this systematic change would impart a distinct cold bias to the HCN database if not properly treated.

### 4. Bias Adjustments

[10] The analysis described here examines the ability of the HCN adjustments to account for the time of observation



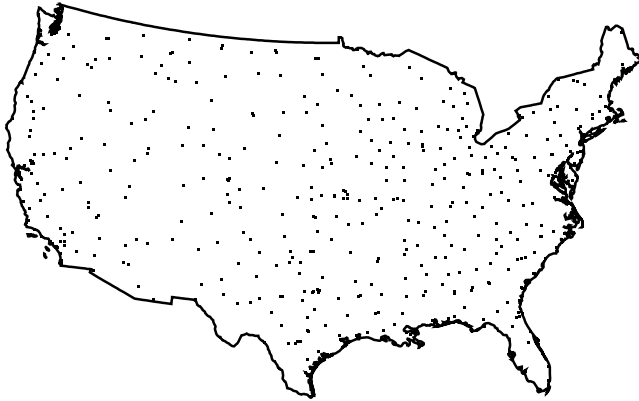
**Figure 1.** Percent of HCN stations with morning and afternoon observation times according to HCN metadata and metadata inferred using the method of DeGaetano [1999].

bias at the national scale. The first step entailed computing a suite of unbiased, biased, and bias-adjusted temperature time series for all hourly observing stations in the country. These series were then used to estimate (a) the time of observation bias at each hour at each station and (b) the residual bias remaining at each hour after the application of the adjustment. Finally, the station-based biases and residuals were areally averaged into hourly means for the conterminous United States, and the importance of the residual biases was assessed by computing the number of HCN stations using each observation time.

[11] Data for the analysis were extracted from the Surface Airways Hourly database [Steurer and Bodosky, 2000] archived at the National Climatic Data Center. The analysis employed data from 1965–2001 because the adjustment approach itself was developed using data from 1957–64. The Surface Airways Hourly database contains data for 500 stations during the study period; the locations of these stations are depicted in Figure 2. The period of record varies from station to station, and no minimum record length was required for inclusion in this analysis.

[12] The first step entailed calculating a time series of mean monthly temperatures for each station. Each monthly mean in each series was the average of the daily mean temperatures in the month, and each daily mean was the average of the maximum and minimum temperatures during the calendar day (i.e., the “observation time” was midnight). The daily maximum was the highest of the 24 hourly temperatures, and the minimum was the lowest of the hourly values. The use of hourly observations rather than true daily extremes does slightly impact the monthly mean [Baker, 1975], but the corresponding effect on the estimate of the time of observation bias in any month is small [i.e., usually less than  $.03^{\circ}\text{C}$ ; Karl *et al.*, 1986].

[13] A time series of biased mean monthly temperatures was then developed for each station for each hour of the day. Bias was introduced by simply changing the “observation time” from midnight to another hour before computing the monthly means. For example, the first biased series for each station had 0100 LST as its observation time, the second had 0200 LST as its observation time, and so on



**Figure 2.** Locations of the 500 stations used to evaluate the *Karl et al.* [1986] adjustment.

through 2300 LST. Each monthly mean in each biased series was the average of the daily mean temperatures in the month, and each daily mean was the average of the maximum and minimum temperatures during the 24-hr period ending at the specified observation time.

[14] Finally, a time series of bias-adjusted mean monthly temperatures was derived from each biased series using the *Karl et al.* [1986] technique. In this approach, drift bias for each month in each series was quantified using a weighted average of the current month's temperature, the previous month's temperature, and the observation hour. Carry-over bias in each month was predicted using multilinear regression equations that require as input the station's coordinates, observation hour, average daily diurnal temperature range ( $\rho$ ), and average daily day-to-day temperature difference ( $\delta$ ). Rather than using daily data, which are not available in HCN prior to 1948, the model interpolates  $\rho$  and  $\delta$  from empirical estimates derived from hourly data for the period 1957–64 at four neighboring first-order stations.

[15] The next step involved estimating (a) the time of observation bias at each hour at each station and (b) the residual bias remaining at each hour after the application of the *Karl et al.* [1986] adjustment model. The time of observation bias (*TOB*) at a specific hour was defined as the average difference between the unbiased series for that hour (i.e., midnight) and the biased series:

$$\frac{1}{(n \times 12)} \sum_{i=1}^n \sum_{j=1}^{12} (B_{ij} - U_{ij}),$$

where  $U_{ij}$  is the temperature in the unbiased series in month  $j$  in year  $i$ ,  $B_{ij}$  is the corresponding temperature in the biased series for that hour, and  $n$  is the number of years with data for that station. Similarly, the residual bias (*RTOB*) was defined as the average difference between the biased series and the bias-adjusted series for that hour:

$$\frac{1}{(n \times 12)} \sum_{i=1}^n \sum_{j=1}^{12} (B_{ij} - A_{ij}),$$

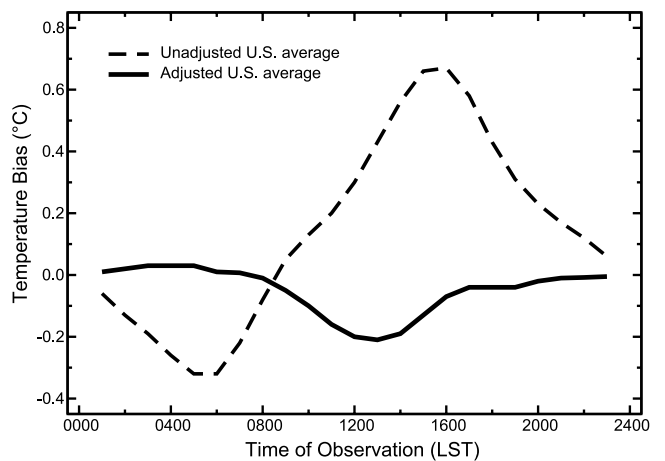
where  $A_{jk}$  is the temperature in the bias-adjusted series in month  $j$  in year  $i$  for that hour.

[16] The *TOB* and *RTOB* values at all stations were averaged into a spatial mean for the conterminous United

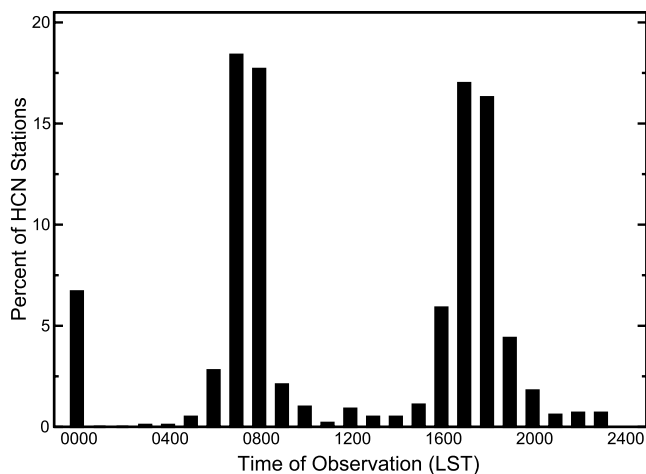
States for each hour of the day. Spatial averaging was accomplished by interpolating station-based values to the nodes of a  $0.25^\circ \times 0.25^\circ$  latitude-longitude grid, then computing the cosine-weighted average of the gridded values. Interpolation was performed using the inverse distance weighting model of *Willmott et al.* [1985], which has been widely used to grid temperature data in the United States [e.g., *Robeson*, 1997]. The model explicitly accounts for the sphericity of the earth when computing the distance between each station and grid point, and it permits extrapolation beyond the range of data values in the neighboring stations.

[17] Figure 3 depicts the spatial mean *TOB* and *RTOB* for each hour of the day over the conterminous United States during the period 1965–2001. In general, morning observation schedules result in a cool *TOB* in mean annual temperature whereas afternoon observations result in a warm *TOB*. The largest positive bias occurs at roughly the same time as the daily maximum temperature, and the greatest negative bias occurs near the time of the daily minimum temperature. The sinusoidal pattern of *TOB*, the magnitude of the biases, and the timing of the extremes are consistent with previous results [e.g., *Winkler et al.*, 1981; *Karl et al.*, 1986]. The magnitude of *RTOB* is less than *TOB*, suggesting that the *Karl et al.* [1986] adjustment results in an improved estimate of calendar-day mean monthly temperatures. The magnitude of *RTOB* exceeds  $0.05^\circ\text{C}$  between 1000 and 1500 LST, with the worst residuals ( $-0.20^\circ\text{C}$ ) occurring around noon LST.

[18] Time of observation metadata suggest that the potential impact of *RTOB* is minimal in HCN (Figure 4). For instance, over the period 1965–2001, only 4% of all HCN stations had an observing hour between 1000 and 1500 LST, indicating that the somewhat larger *RTOB* during that period should be of little consequence in national-scale analyses. More than a third of the stations observed at either 0700 or 0800 LST, when the average *RTOB* was only  $-0.003^\circ\text{C}$  (versus an average *TOB* of  $-0.15^\circ\text{C}$ ). Another third observed at either 1700 or 1800 LST, when *RTOB* was  $-0.04^\circ\text{C}$  (versus an average *TOB* of  $0.51^\circ\text{C}$ ). Aside from midnight, no other hour accounted for more than about 5%



**Figure 3.** Time of observation temperature bias in the conterminous United States before and after the application of the *Karl et al.* [1986] adjustment.



**Figure 4.** Percent of HCN stations using various observation times from 1965–2001 as documented by HCN metadata.

of the total record, and none had an *RTOB* with a magnitude in excess of  $0.05^{\circ}\text{C}$ .

## 5. Temperature Trends

[19] *Balling and Idso* [2002] found that adjusted HCN temperature trends for the past 30 years were slightly more positive than those calculated using an updated version of the *Jones* [1994] dataset. This is surprising because the latter contains a nearly complete version of HCN that includes adjustments for the time of observation bias. To resolve this discrepancy, temperature trends derived from the fully adjusted HCN database are compared here with those derived from two subsets of *Jones*. The first subset consists of all U.S. stations in *Jones* (1578 in total). The second consists of 248 stations that are not in HCN and that require no adjustments for variations in observation time because they always have an observation hour of midnight.

[20] The same technique was used to compute temperature trends for HCN and the two subsets of *Jones*. The first step involved removing stations that had less than 20 years of data during the 1961–90 base period. Next, each annual value at each station was converted to an anomaly from its mean during the base period. The annual anomalies were then interpolated to the nodes of a  $0.25^{\circ} \times 0.25^{\circ}$  latitude-longitude grid using the method of *Willmott et al.* [1985], and the grid point values were area weighted into a mean anomaly for the conterminous United States for each year. Finally, least squares regression was applied to compute trends over the period 1970–2000. Consistent with *Balling and Idso* [2002], HCN has a larger trend during that period ( $0.29^{\circ}\text{C dec}^{-1}$ ) than either subset of *Jones* (each with a trend of  $0.23^{\circ}\text{C dec}^{-1}$ ). However, the difference in trend results from a drastic change in the size of the *Jones* dataset in 1996; prior to that point the network contains at least 1000 U.S. stations per year (the majority being HCN stations), whereas thereafter it contains no more than 150. When trends are computed for the period 1970–95, HCN and both subsets of *Jones* exhibit the same rate of warming ( $0.25^{\circ}\text{C dec}^{-1}$ ). The fact that the non-HCN subset of *Jones* has the same trend as HCN suggests that the time of observation bias has been properly treated in the latter.

## 6. Summary and Conclusions

[21] Given the documented change from afternoon to morning observation schedules over the past half century and the known cool (warm) bias for morning (afternoon) observers, it is clear that the U.S. climate record contains a nonclimatic cooling effect that must be addressed when estimating the magnitude of a temperature trend. Consequently, this paper reviewed the efficacy the approach used to address this bias in HCN. First, the accuracy of HCN station history information was verified with an independent source of metadata inferred using the method of *DeGaetano* [1999]. The predictive skill of the adjustment approach was then examined using data from a network of 500 stations over the period 1965–2001. While nontrivial bias-adjustment error was apparent from 1000–1500 LST, those observation hours were found to be used by very few stations. Finally, adjusted HCN temperature trends were shown to be consistent with those derived from *Jones* [1994]. In short, the adjustment for the time of observation bias in HCN appears to be robust.

[22] **Acknowledgments.** Partial support for this work was provided by the Office of Biological and Environmental Research, U.S. Department of Energy, and the NOAA Office of Global Programs.

## References

- Baker, D. G., Effect of observation time on mean temperature estimation, *J. Appl. Meteor.*, *14*, 471–476, 1975.
- 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*, 2002.
- Christy, J. R., R. W. Spencer, and W. D. Braswell, MSU tropospheric temperatures: Data set construction and radiosonde comparisons, *J. Atmos. Oceanic Tech.*, *17*, 1153–1170, 2000.
- DeGaetano, A. T., A serially complete simulated observation time metadata file for U.S. Daily Historical Climatology Network stations, *Bull. Amer. Meteorol. Soc.*, *81*, 49–67, 1999.
- Easterling, D. R., T. R. Karl, E. H. Mason, P. Y. Hughes, D. P. Bowman, R. C. Daniels, and T. A. Boden, United States Historical Climatology Network (U.S. HCN) Monthly Temperature and Precipitation Data, Carbon Dioxide Information and Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, Publication No. 4500, 83 pp., 1996.
- Hansen, J., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. R. Easterling, T. C. Peterson, and T. R. Karl, A closer look at United States and global surface temperature change, *J. Geophys. Res.*, *106*, 23,947–23,964, 2001.
- Jones, P. D., Hemispheric surface air temperature variations: A reanalysis and an update to 1993, *J. Clim.*, *7*, 1794–1802, 1994.
- Karl, T. R., and C. N. Williams Jr., An approach to adjusting climatological time series for discontinuous inhomogeneities, *J. Climate Appl. Meteorol.*, *26*, 1744–1763, 1987.
- Karl, T. R., C. N. Williams Jr., P. J. Young, and W. M. Wendland, A model to estimate the time of observation bias associated 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.
- Robeson, S. R., Spherical methods for interpolation, *Cartogr. Geogr. Inf. Sys.*, *24*, 3–20, 1997.
- Steurer, P., and M. Bodosky, Data documentation for data set 3280 Surface Airways Hourly, National Climatic Data Center, 2000.
- Willmott, C. J., C. M. Rowe, and W. D. Philpot, Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring, *Am. Cartographer*, *12*, 5–16, 1985.
- Winkler, J. A., R. H. Skaggs, and D. G. Baker, Effect of temperature adjustments on the Minneapolis-St. Paul Urban Heat Island, *J. Appl. Meteorol.*, *20*, 1295–1300, 1981.

R. S. Vose, C. N. Williams Jr., T. C. Peterson, T. R. Karl, and D. R. Easterling, National Climatic Data Center, 151 Patton Avenue, Asheville, NC 28801, USA. (russell.vose@noaa.gov)