View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by MURAL - Maynooth University Research Archive Library

GEOPHYSICAL RESEARCH LETTERS, VOL. 39, L10703, doi:10.1029/2012GL051387, 2012

# An intercomparison of temperature trends in the U.S. Historical Climatology Network and recent atmospheric reanalyses

Russell S. Vose,[1] Scott Applequist,[1] Matthew J. Menne,[1] Claude N. Williams Jr.,[1] and Peter Thorne[2]

Received 17 February 2012; revised 13 April 2012; accepted 15 April 2012; published 17 May 2012.

[1] Temperature trends over 1979–2008 in the U.S. Historical Climatology Network (HCN) are compared with those in six recent atmospheric reanalyses. For the conterminous United States, the trend in the adjusted HCN (0.327 °C $dec^{-1}$) is generally comparable to the ensemble mean of the reanalyses (0.342 °C $dec^{-1}$). It is also well within the range of the reanalysis trend estimates (0.280 to 0.437 °C $dec^{-1}$). The bias adjustments play a critical role, as the raw HCN dataset displays substantially less warming than all of the reanalyses. HCN has slightly lower maximum and minimum temperature trends than those reanalyses with hourly temporal resolution, suggesting the HCN adjustments may not fully compensate for recent non-climatic artifacts at some stations. Spatially, both the adjusted HCN and all of the reanalyses indicate widespread warming across the nation during the study period. Overall, the adjusted HCN is in broad agreement with the suite of reanalyses. **Citation:** Vose, R. S., S. Applequist, M. J. Menne, C. N. Williams Jr., and P. Thorne (2012), An intercomparison of temperature trends in the U.S. Historical Climatology Network and recent atmospheric reanalyses, *Geophys. Res. Lett.*, *39*, L10703, doi:10.1029/2012GL051387.

## 1. Introduction

[2] NOAA's National Climatic Data Center (NCDC) releases a report each month assessing the *State of the Climate* of the previous month. For its national-scale assessment, NCDC relies heavily upon the U.S. Historical Climatology Network surface temperature database (HCN version 2 [*Menne et al.*, 2009]). HCN was selected for operational monitoring by NCDC for two primary reasons. First, the database is a subset of the Cooperative Observing Network, stations having been chosen based on their spatial coverage, record length, data completeness, and historical stability. Second, HCN contains multiple adjustments that account for historical changes in observation time, station location, temperature instrumentation, and siting conditions.

[3] While HCN is important operationally, atmospheric reanalyses are also useful in a climate monitoring context [*Dee et al.*, 2011a]. In brief, a reanalysis is a fixed data assimilation and numerical weather prediction system run in hindcast mode over a period sufficient in length for climate applications. A reanalysis synthesizes all of the observations presented to it in a manner consistent with model physics, generating spatially complete fields for the period of record. Although reanalyses assimilate some surface observations (Table S1 in the auxiliary material), none directly ingest near-surface air temperature data in their atmospheric analysis.[1] Consequently, reanalyses can serve as a counterpart to the purely instrumental temperature record in climate monitoring activities.

[4] Given the potential utility of both, this paper intercompares recent trends in HCN and a suite of contemporary atmospheric reanalyses. In particular, the study uses six reanalyses to bracket a range of plausible U.S. trends, noting the relative position of HCN within the larger, reanalysis-based context. The analysis consists of two parts, the first quantifying trends in area-averaged time series, and the second exploring spatial patterns in trends. Mean annual temperature is evaluated for the full suite of reanalyses, and maximum and minimum temperature are evaluated for two reanalyses having hourly temporal resolution. The study is restricted the period 1979–2008, the era common to HCN and all of the reanalyses (though most are in fact available in near-real time for operational climate monitoring).

## 2. HCN Version 2

[5] HCN version 2 consists of 1218 stations distributed in a quasi-uniform fashion across the conterminous United States [*Menne et al.*, 2009]. The database is commonly employed in climate change research because its temperature records contain adjustments for historical changes in observing practice. The efficacy of both the station network and the adjustment approach has been well documented for large-scale applications [e.g., *Vose et al.*, 2003; *Vose and Menne*, 2004; *Menne et al.*, 2009]. Recent efforts to benchmark system performance through recourse to analogs further confirm the efficiency of the approach [*Williams et al.*, 2012]. During the analysis period the adjustments compensate for two net cold biases in maximum temperature and generally balancing cold/warm biases in minimum temperature. The adjustments also minimize the large-scale trend bias resulting from suboptimal siting, which may afflict a significant portion of the network [*Fall et al.*, 2011]. In particular, since 1980 the national-scale time series derived from poor exposure sites is comparable to that from good exposure sites after the application of the adjustments [*Menne et al.*, 2010; *Fall et al.*, 2011].

[1]National Climatic Data Center, Asheville, North Carolina, USA.
[2]Cooperative Institute for Climate and Satellites–North Carolina, Asheville, North Carolina, USA.

Corresponding author: R. S. Vose, National Climatic Data Center, 151 Patton Ave., Asheville, NC 28801, USA. (russell.vose@noaa.gov)

This paper is not subject to U.S. copyright.
Published in 2012 by the American Geophysical Union.

[1]Auxiliary materials are available in the HTML. doi:10.1029/2012GL051387.

[6] From an operational perspective, there are two separate bias corrections that are applied in sequence. The first [*Karl et al.*, 1986] accounts for documented changes in observation time that result in both day-to-day "carry-over" bias and end-of-the-month "drift" bias. Approximately 20% of the network experienced a change in observation time during the study period. This change was almost always from an afternoon to a morning observation time, the latter having a known cool bias [*Hubbard and Lin*, 2006]. The net effect at the national scale was an artificial decrease of ~0.20 °C for both maximum and minimum temperature, a decrease that synthetically reduced the true U.S. temperature trend in recent decades [*Menne et al.*, 2009].

[7] The second bias correction [*Menne and Williams*, 2009] attempts to account for all other changes, such as in station location, temperature instrumentation, and siting conditions. The implementation of the Maximum Minimum Temperature System (MMTS) was particularly important during the period, as it impacted some 70% of the network and often coincided with new siting conditions. The resulting change was readily detectable for the majority of stations by the adjustment approach, which relies upon extensive pairwise interstation comparisons. On average, MMTS recorded slightly lower maxima relative to liquid-in-glass thermometers, and vice-versa for minima [*Quayle et al.*, 1991]. The net effect of MMTS and other such changes was an artificial decrease of ~0.25 °C for maximum temperature and an increase of ~0.20 °C for minimum temperature [*Menne et al.*, 2009].

## 3. Contemporary Reanalyses

[8] This paper employs near-surface air temperature fields from the six major atmospheric reanalyses completed since 2006. These include the 20th Century Reanalysis (20CR) [*Compo et al.*, 2011]; the Climate Forecast System Reanalysis (CFSR) [*Saha et al.*, 2010]; ERA-Interim (ERA-INT) [*Dee et al.*, 2011b]; the Japanese 25-year Reanalysis (JRA-25) [*Onogi et al.*, 2007]; the Modern Era Reanalysis for Research and Applications (MERRA) [*Rienecker et al.*, 2011]; and the North America Regional Reanalysis (NARR) [*Mesinger et al.*, 2006].

[9] There are substantial differences across the reanalysis suite. For example, temporal resolution varies considerably, with JRA-25 having 6-hourly fields; 20CR, ERA-INT, and NARR having 3-hourly fields; and CFSR and MERRA having hourly fields. Each reanalysis has a different spatial resolution, half being coarser than ~125 km (20CR, ERA-INT, JRA-25) and half being finer than ~50 km (CFSR, MERRA, and NARR). (Notably, each has more than 300 grid points over the study area, far in excess of the 135 recommended to detect trends in the mean annual temperature anomaly of the conterminous United States [*Vose and Menne*, 2004]). Each reanalysis employs a distinct data assimilation scheme and different input datasets; for instance, 20CR only assimilates surface pressure reports whereas CFSR is a fully coupled ocean-atmosphere reanalysis that ingests subsurface ocean data in addition to surface and atmospheric data. No reanalysis directly assimilates near-surface air temperature data in its atmospheric analysis, though ERA-INT uses unadjusted temperature and humidity data in its land-surface analysis to nudge the soil moisture stream. (For more information on each reanalysis, see http://reanalyses.org.)

[10] There are two main reasons for the inclusion of multiple reanalyses in this study. First, an inclusive approach makes no a priori assumption about the inferiority of any reanalysis for a specific application. Rather, it only assumes that each will have generally credible trends for the study area, which is reasonable given that several recent studies have documented the plausibility of reanalysis trends over North American in particular (e.g., NARR [*Fall et al.*, 2010] and ERA-INT [*Simmons et al.*, 2010]). Second, an inclusive approach allows for the quantification of the structural uncertainty [*Thorne et al.*, 2005] resulting from differences in reanalysis data assimilation schemes and input datasets. This uncertainty generally manifests itself as an array of possible trend values, the integrity of both HCN and the reanalyses being supported if the latter exhibit reasonably good agreement with the former. In theory, such a comparison could be impacted by systematic biases across the reanalyses [*Bosilovich et al.*, 2009] and by the effect of observing system changes in multiple reanalyses [*Bosilovich et al.*, 2011]. However, such impacts are minimized in this study through the inclusion of several distinct approaches (ranging from surface-only to coupled ocean-atmosphere reanalyses) that jointly span an extensive portion of the structural uncertainty phase space.

## 4. U.S. Trends

[11] The analysis described here examined trends over the conterminous United States as a whole at the annual time scale. The primary goal was to compare HCN with the reanalysis suite. This was accomplished by considering both the ensemble mean of the reanalyses as well as the range of estimates across the six reanalyses (as in *Kumar and Hu* [2011]). The secondary goal was to determine whether the HCN adjustments improved consistency with the reanalysis suite. This was accomplished by examining trends in both the raw dataset and the fully adjusted, serially complete HCN.

[12] Trends were computed via least-squares regression on area-averaged time series. For HCN, trends were calculated using the same U.S. time series as in *Menne et al.* [2009], only updated through 2008 and converted to a 1979–2008 base period (i.e., by computing the average for that 30-year period, then subtracting that average from each value). For the reanalyses, a U.S. time series had to be prepared for each element prior to calculating trends. In brief, each series was simply the area-weighted average of anomalies at grid boxes whose midpoint fell within the conterminous United States (anomalies being used to account for differences in elevation, latitude, and coastal proximity across the grid points). For mean temperature, grid-box anomalies were produced by first computing each box's long-term average using all hours across all days from 1979–2008, then subtracting that average from the annual value computed using all hours across all days in that particular year. The process was essentially the same for maximum (minimum) temperature, but using only the highest (lowest) of the 24 hourly values on each day. Maximum and minimum temperatures were only computed for reanalyses with hourly fields (i.e., CFSR and MERRA).

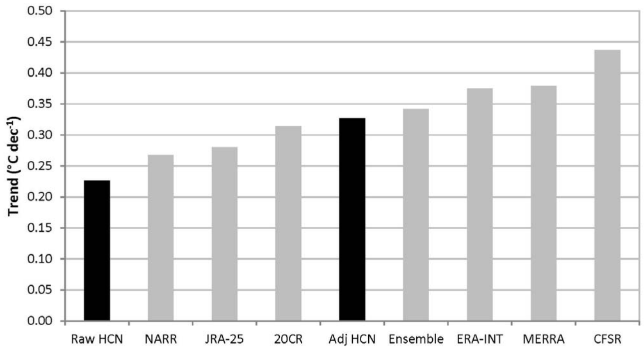


**Figure 1.** Least-squares trends (°C $dec^{-1}$) in mean annual temperature over the conterminous United States during the period 1979–2008. All trends are significant at the 0.05 level. The "Ensemble" value is just the average of the six reanalysis trends.

[13] Figure 1 presents trends in mean annual temperature over the conterminous United States during the period 1979–2008. All of the trends are positive and statistically significant, a clear indication of warming over the past three decades. The trend in the adjusted HCN (0.327 °C $dec^{-1}$) is roughly comparable to the ensemble mean of the reanalyses (0.342 °C $dec^{-1}$). It is also well within the range of the reanalysis trend estimates (0.280 to 0.437 °C $dec^{-1}$). The bias adjustments are salient in this regard, as the raw HCN dataset displays substantially less warming than any reanalysis product. Three of the reanalyses exhibit more warming than the adjusted HCN (i.e., CFSR, ERA-INT, and MERRA), one has about the same rate of warming (i.e., 20CR), and two exhibit less warming (i.e., JRA-25 and NARR).

[14] Figure 2 presents area-averaged trends in mean annual maximum and minimum temperature. As with mean annual temperature, all trends are positive and statistically significant. The adjusted HCN exhibits better agreement with the reanalyses than does its raw counterpart, which has smaller trends for both temperature elements. The impact of the HCN adjustments is considerably larger for maximum temperature because the corrections removed two distinct cold biases for that element (versus the loosely offsetting cold/warm biases for the minimum). The trend in maximum temperature exceeds that of minimum temperature by about 0.05 °C $dec^{-1}$ for both the adjusted HCN and the reanalyses. Again, the bias corrections account for this result in the adjusted HCN, as the raw HCN dataset has the exactly the opposite relationship (i.e., more warming in minimum temperature).

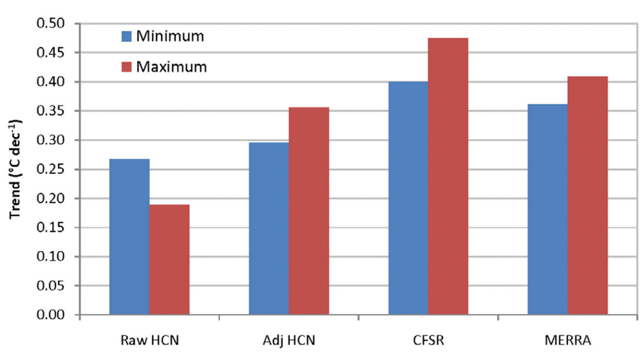


**Figure 2.** Least-squares trends (°C $dec^{-1}$) in mean annual maximum and minimum temperature over the conterminous United States during the period 1979–2008. All trends are significant at the 0.05 level.

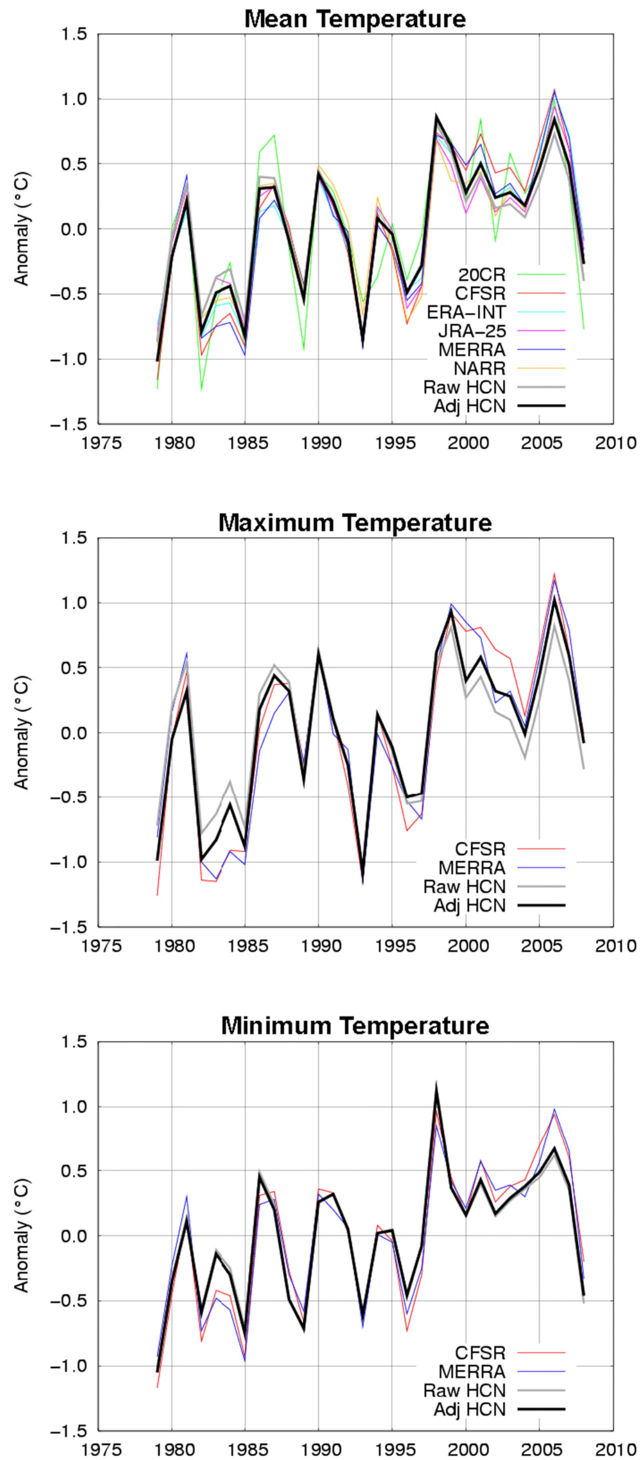


**Figure 3.** Time series of annual mean, maximum, and minimum temperature anomalies over the conterminous United States during the period 1979–2008.

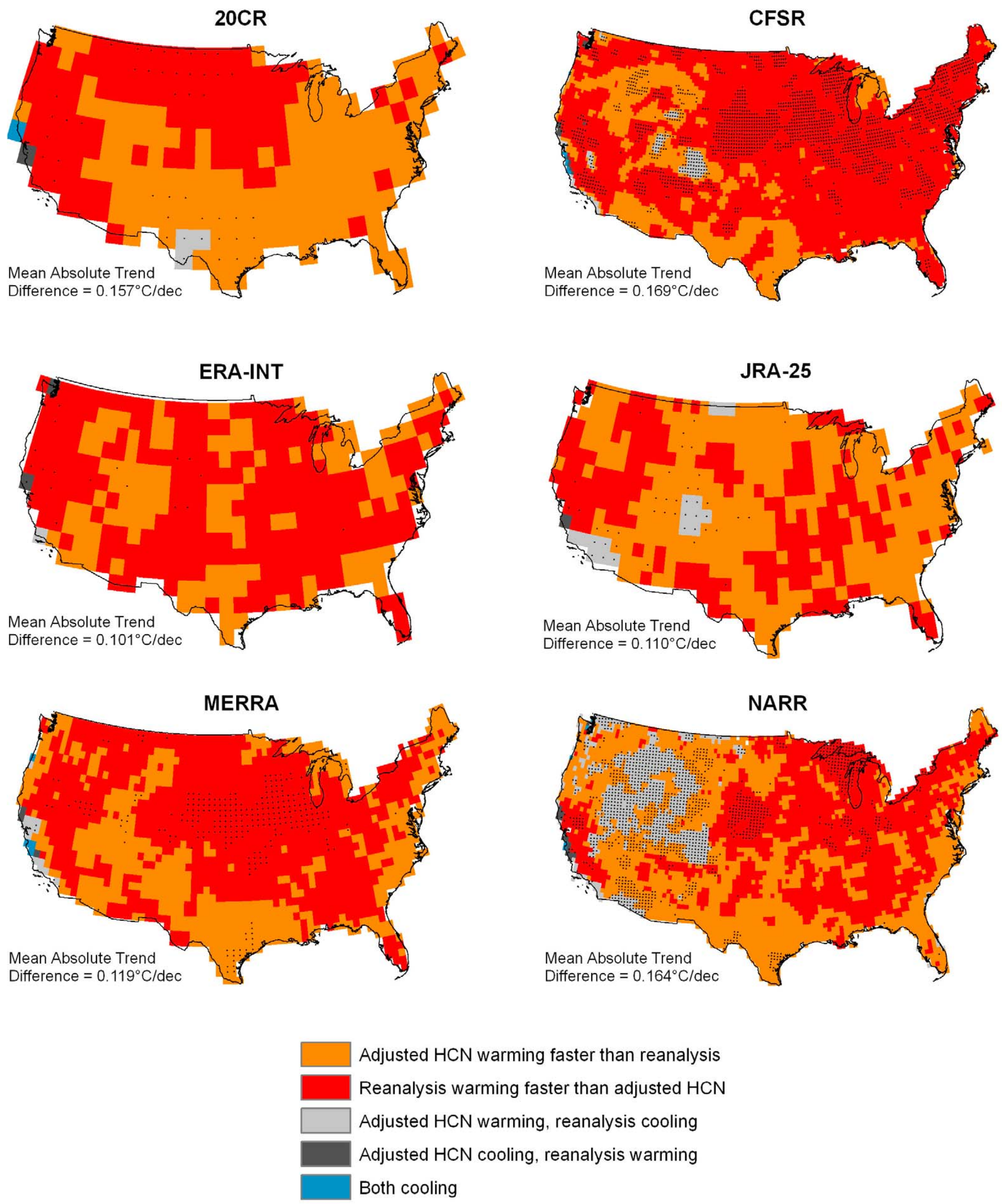


**Figure 4.** Categorical depiction of grid-box trends in mean annual temperature during the period 1979–2008. Stippled areas contain trend differences exceeding 0.25 °C dec$^{-1}$. The mean absolute trend difference is the area-weighted average of absolute trend differences across all grid boxes (i.e., the trend difference was computed for each grid box, then the sign was removed from that difference, then the unsigned grid-box differences were area-averaged into a single value for the conterminous United States).

[15] Figure 3 depicts the area-averaged time series for HCN and the reanalyses. Generally speaking, all of the series exhibit a high degree of similarity, and the adjusted HCN is clearly in agreement with the suite of reanalyses. In any particular year, the adjusted HCN rarely constitutes the most extreme estimate of the mean temperature anomaly. For maximum and minimum temperature, the adjusted HCN is slightly warmer than the reanalyses in the early 1980s and slightly cooler since the late 1990s. Consistent with *Fall et al.* [2011] and *Williams et al.* [2012], this suggests that the adjusted HCN may not entirely compensate for recent changes at some stations (e.g., the implementation of electronic instrumentation, often with concurrent changes in siting). In other words, the adjusted HCN may slightly underestimate recent rates of warming, particularly for maximum temperature.

## 5. Spatial Patterns

[16] The analysis described here examined spatial patterns in trends of mean annual temperature. Again, the primary goal was to compare HCN with the reanalysis suite, as depicted by the range of estimates across the six reanalyses. The first step involved converting the adjusted HCN series at each station to an anomaly series (i.e., by computing the mean from 1979–2008, then subtracting that mean from

each value), thereby accounting for differences in elevation, latitude, and coastal proximity across stations. The annual anomalies were then interpolated to the grid nodes of each reanalysis using the method of *Willmott et al.* [1985]. Station-based anomalies were interpolated to reanalysis grid nodes (rather than the reverse) for two reasons: HCN is of greater spatial density than three of the reanalyses (20CR, JRA-25, and ERA-INT), and the HCN adjustments were primarily designed for computing robust areal averages rather than assuring pristine single-station series [*Easterling et al.*, 1996]. The least-squares trend was then computed for each gridded HCN series (as well as for each reanalysis grid-box series). Finally, for each reanalysis, each grid box was assigned to one of several categories that depicted the sign of the trend in the adjusted HCN, the sign of the trend in the reanalysis, and the relative rate of change in each.

[17] Figure 4 depicts spatial patterns in trends of mean annual temperature. For each reanalysis, most grid boxes fall into one of two "warming" categories (i.e., the adjusted HCN is warming faster than the reanalysis, or vice-versa). Although warming rates differ at the grid-box level, the adjusted HCN and all of the reanalyses agree in their depiction of a widespread temperature increase since 1979. By extension, relatively few grid boxes fall into one of the three "cooling" categories (i.e., the adjusted HCN is warming and the reanalysis is cooling, or vice-versa, or both are cooling). In fact, the only notable areas of cooling are the intermountain West (particularly for NARR) and the coast of California (for HCN and/or the reanalyses). Overall, there is no large area where the adjusted HCN regularly exhibits more (or less) warming than all of the reanalyses; in fact, such regularity is only apparent occasionally and only over relatively small areas (e.g., the adjusted HCN always has more warming over southern Texas).

[18] Other than broad-scale warming, it is difficult to verify the subtleties in the HCN trend pattern given the lack of consistency across the reanalyses. For instance, in the eastern United States 20CR has lower trends than the adjusted HCN whereas CFSR has larger trends. The pattern is smooth in some cases (as in 20CR) and noisy in others (as in NARR). Regions with large trend differences (denoted as stippling on the maps) vary from one reanalysis to another; for example, CFSR and MERRA contain large trend differences over the central United States while JRA-25 and NARR have significant differences over the West. Divergent patterns even exist when area-averaged metrics of similarity are comparable; for instance, JRA-25 and MERRA have about the same mean absolute trend difference with the adjusted HCN, yet their respective maps are completely different. Overall, this lack of consistency strikes a cautionary note regarding the use of a single reanalysis to scrutinize small-scale trends in observed surface temperatures (e.g., as in *Pielke et al.* [2007] and *Fall et al.* [2010]).

## 6. Summary and Conclusions

[19] This paper compared trends in HCN during the period 1979–2008 with those in six recent atmospheric reanalyses. For the conterminous United States as a whole, all trends in mean annual temperature are positive and statistically significant, with the trend in the adjusted HCN being roughly comparable to the ensemble mean of the reanalyses. The raw HCN dataset exhibits much less warming than all of the reanalyses, confirming the necessity for and the utility of the HCN bias adjustments. Reanalyses with hourly temporal resolution have slightly higher maximum and minimum temperature trends than the adjusted HCN, suggesting the HCN bias adjustments may not fully compensate for recent changes in siting and instrumentation at some stations. From a spatial perspective, both the adjusted HCN and all of the reanalyses indicate widespread warming across the nation; however, variability across the reanalyses precludes verification of the finer features within the HCN trend pattern itself. Overall, there is broad agreement between the adjusted HCN and the reanalysis suite, building confidence in both approaches from a climate monitoring perspective.

[20] **Acknowledgments.** The Editor thanks the two anonymous reviewers for assisting with the evaluation of this paper.

## References

Bosilovich, M. G., D. Mocko, J. O. Roads, and A. Ruane (2009), A multimodel analysis for the Coordinated Enhanced Observing Period (CEOP), *J. Hydrometeorol.*, *10*, 912–934, doi:10.1175/2009JHM1090.1.

Bosilovich, M. G., F. R. Robertson, and J. Chen (2011), Global energy and water budgets in MERRA, *J. Clim.*, *24*, 5721–5739, doi:10.1175/2011JCLI4175.1.

Compo, G. P., et al. (2011), The Twentieth Century Reanalysis project, *Q. J. R. Meteorol. Soc.*, *137*, 1–28, doi:10.1002/qj.776.

Dee, D., et al. (2011a), The use of reanalysis data for monitoring the state of the climate, *Bull. Am. Meteorol. Soc.*, *92*, S34–S35.

Dee, D., et al. (2011b), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*, 553–597, doi:10.1002/qj.828.

Easterling, D. R., T. R. Karl, E. H. Mason, P. Y. Hughes, and D. P. Bowman (1996), 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., U.S. Dep. of Energy, Oak Ridge, Tenn.

Fall, S., D. Niyogi, R. Pielke, A. Gluhovsky, E. Kalnay, and G. Rochon (2010), Impacts of land use land cover on temperature trends over the continental United States: Assessment using the North American Regional Reanalysis, *Int. J. Climatol.*, *30*, 1980–1993, doi:10.1002/joc.1996.

Fall, S., A. Watts, J. Nielsen-Gammon, E. Jones, D. Niyogi, J. Christy, and R. A. Pielke Sr. (2011), Analysis of the impacts of station exposure on the U.S. Historical Climatology Network temperatures and temperature trends, *J. Geophys. Res.*, *116*, D14120, doi:10.1029/2010JD015146.

Hubbard, K. G., and X. Lin (2006), Reexamination of instrument change effects in the U.S. Historical Climatology Network, *Geophys. Res. Lett.*, *33*, L15710, doi:10.1029/2006GL027069.

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. Clim. Appl. Meteorol.*, *25*, 145–160, doi:10.1175/1520-0450(1986)025<0145:AMTETT>2.0.CO;2.

Kumar, A., and Z. Hu (2011), Uncertainty in the ocean-atmosphere feedbacks associated with ENSO in reanalysis products, *Clim. Dyn.*, doi:10.1007/s00382-011-1104-3.

Menne, M. J., and C. N. Williams Jr. (2009), Homogenization of temperature series via pairwise comparisons, *J. Clim.*, *22*, 1700–1717, doi:10.1175/2008JCLI2263.1.

Menne, M. J., C. N. Williams, and R. S. Vose (2009), The United States Historical Climatology Network monthly temperature data, version 2, *Bull. Am. Meteorol. Soc.*, *90*, 993–1007, doi:10.1175/2008BAMS2613.1.

Menne, M. J., C. N. Willaims Jr., and M. A. Palecki (2010), On the reliability of the U.S. surface temperature record, *J. Geophys. Res.*, *115*, D11108, doi:10.1029/2009JD013094.

Mesinger, F., et al. (2006), North American Regional Reanalysis, *Bull. Am. Meteorol. Soc.*, *87*, 343–360, doi:10.1175/BAMS-87-3-343.

Onogi, K., et al. (2007), The JRA-25 Reanalysis, *J. Meteorol. Soc. Jpn.*, *85*, 369–432, doi:10.2151/jmsj.85.369.

Pielke, R., Sr., et al. (2007), Documentation of uncertainties and biases associated with surface temperature measurement sites for climate change assessment, *Bull. Am. Meteorol. Soc.*, *88*, 913–928, doi:10.1175/BAMS-88-6-913.

Quayle, R. G., D. R. Easterling, T. R. Karl, and P. Y. Hughes (1991), Effects of recent thermometer changes in the cooperative station network, *Bull. Am. Meteorol. Soc.*, *72*, 1718–1723, doi:10.1175/1520-0477(1991)072<1718:EORTCI>2.0.CO;2.

Rienecker, M. M., et al. (2011), MERRA—NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.*, *24*, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Saha, S., et al. (2010), The NCEP Climate Forecast System Reanalysis, *Bull. Am. Meteorol. Soc.*, *91*, 1015–1057, doi:10.1175/2010BAMS3001.1.

Simmons, A. J., K. M. Willett, P. D. Jones, P. W. Thorne, and D. P. Dee (2010), Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets, *J. Geophys. Res.*, *115*, D01110, doi:10.1029/2009JD012442.

Thorne, P. W., D. E. Parker, J. R. Christy, and C. A. Mears (2005), Uncertainties in climate trends: Lessons from upper-air temperature records, *Bull. Am. Meteorol. Soc.*, *86*, 1437–1442, doi:10.1175/BAMS-86-10-1437.

Vose, R. S., and M. J. Menne (2004), A method to determine station density requirements for climate observing networks, *J. Clim.*, *17*, 2961–2971, doi:10.1175/1520-0442(2004)017<2961:AMTDSD>2.0.CO;2.

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.

Williams, C. N., M. J. Menne, and P. W. Thorne (2012), Benchmarking the performance of pairwise homogenization of surface temperatures in the United States, *J. Geophys. Res.*, *117*, D05116, doi:10.1029/2011JD016761.

Willmott, C. J., C. M. Rowe, and W. D. Philpot (1985), Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring, *Cartogr. Geogr. Inf. Sci.*, *12*, 5–16, doi:10.1559/152304085783914686.