

Documentation of Uncertainties and Biases Associated with Surface Temperature Measurement Sites for Climate Change Assessment

BY ROGER PIELKE SR., JOHN NIELSEN-GAMMON, CHRISTOPHER DAVEY, JIM ANGEL, ODIE BLISS, NOLAN DOESKEN, MING CAI, SOULEYMANE FALL, DEV NIYOGI, KEVIN GALLO, ROBERT HALE, KENNETH G. HUBBARD, XIAOMAO LIN, HONG LI, AND SETHU RAMAN

The use of temperature data from poorly sited stations can lead to a false sense of confidence in the robustness of multidecadal surface air temperature trend assessments.

Davey and Pielke (2005) presented photographic documentation of poor observation sites within the U.S. Historical Climate Reference Network (USHCN) with respect to monitoring long-term surface air temperature trends. [These photographs were first shown to the community at the 2002 Asheville, North Carolina, meeting of the American Association of State Climatologists (see information online at www.stateclimate.org/meetings/minutes/2002minutes).] Peterson (2006) compared the adjusted climate records of many of these stations and concluded that

. . . the similarity between the homogeneity-adjusted time series from the good and poorly sited stations supports the view that even stations that do not, upon visual inspection, appear to be spatially representative can, with proper homogeneity adjustments, produce time series that are indeed representative of the climate variability and change in the region.

One of the objectives of the USHCN, however, as stated in Easterling et al. (1996),

AFFILIATIONS: PIELKE—CIRES and ATOC, University of Colorado, Boulder, Colorado; NIELSEN-GAMMON—Department of Atmospheric Sciences, Texas A&M University, College Station, Texas; DAVEY—Desert Research Institute, Reno, Nevada; ANGEL—Illinois State Water Survey, Department of Natural Resources, Champaign, Illinois; BLISS AND DOESKEN—Atmospheric Science, Colorado State University, Fort Collins, Colorado; CAI—Department of Meteorology, The Florida State University, Tallahassee, Florida; FALL AND NIYOGI—Departments of Agronomy and Earth & Atmospheric Sciences, Purdue University, West Lafayette, Indiana; GALLO—Center for Satellite Applications and Research, NOAA/NESDIS, Camp Springs, Maryland; HALE—CIRES, Colorado State University, Fort Collins, Colorado; HUBBARD AND LIN—School of Natural Resources, University of Nebraska—Lincoln, Lincoln, Nebraska; LI—Department of Atmospheric and Oceanic Science, University of Maryland,

College Park, College Park, Maryland; RAMAN—Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina

CORRESPONDING AUTHOR: Roger Pielke Sr., CIRES and ATOC, Folsom Stadium 255-16, University of Colorado, Boulder, CO 80309
E-mail: pielkesr@cires.colorado.edu

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/BAMS-88-6-913

In final form 8 February 2007
©2007 American Meteorological Society

... was to detect temporal changes in regional rather than local climate. Therefore, only stations not influenced to any substantial degree by artificial changes in their local environments were included in the network.

Peterson's claim relaxes this requirement with the implication that data from stations with siting inhomogeneities, after adjustment, may be used to represent regional changes. There remain significant issues, however, with the methodology applied and the conclusions reached in the Peterson article.

UNDOCUMENTED STATION CHANGES.

In the United States, the primary source of surface observations used to construct the long-term global surface temperature analyses has been National Oceanic and Atmospheric Administration (NOAA) stations, which include first-order stations and a subset of NOAA Cooperative Observer Program (COOP) sites that compose the USHCN (information online at <http://cdiac.ornl.gov/epubs/ndp/ushcn/newushcn.html>). The COOP network has long served as the main climate observation network in the United States, with once-daily measurements of temperature, precipitation, and sometimes snowfall/snow depth made by volunteers using equipment supplied, installed, and maintained by NOAA. The metadata for these sites, including information on site exposure, has been provided in B-44 forms, and their equivalents, for the past several decades up to a century.

In the early and middle part of the twentieth century, these forms usually included a schematic drawing of the exposure characteristics of these sites. During the 1980s, the format of these B-44 forms changed as computer entry replaced hand-typed forms (including hand-drawn site exposure graphics). Site drawings were replaced by cryptic "nomenclature" of the site exposure using azimuth, range, and elevation to the nearest obstructions. The "distance and direction from previous locations" field was omitted on the more recent forms.

Photographic documentation has been virtually nonexistent throughout the history of the majority of these sites, and so for the period from the mid-1980s until the present, the only information on site exposure has been from abbreviated "azimuth/range/distance" descriptions. Recently, there have been efforts to photograph present USHCN sites, and other candidate locations, to determine whether these sites should be further considered for inclusion in NOAA's Environmental Real-Time Observation Network (NERON; information online at www.isos.noaa.gov/overview/).

This effort, however, has not been expanded to all NOAA sites (either first order or COOP).

Efforts are under way to continue to improve the statistical assessment of data inhomogeneities (e.g., Mitchell and Jones 2005). However, significant homogeneity issues are still missed. The serious undocumented problem at Holly, Colorado, was first identified by Davey and Pielke (2005), and was not flagged by statistical techniques until the recently developed Menne and Williams (2005) test was applied by Peterson (2006). Photographic documentation and other metadata, if maintained and compared over time, is therefore valuable, both for confirming station inhomogeneities flagged by statistical techniques and for identifying station inhomogeneities that are too subtle to be unambiguously identified by statistical techniques. In a separate study, Mahmood et al. (2006) used improved metadata involving 12 COOP and USHCN stations in Kentucky, and found that undesirable instrument exposure associated with both anthropogenic and natural influences resulted in large variations in the measurements of temperature.

Moreover, there is an undocumented move with one of the sites used in the Peterson analysis (Las Animas, Colorado). The candidate dates for homogeneity adjustments at Las Animas listed by Peterson (2006; his Table 2) are at and after 1986. The B-44 immediately before 1986, chronologically, was the last B-44 that had a schematic of the Las Animas site exposure. This particular B-44, however, showed the Las Animas site as being located just over 50 m northwest of its current site. The current site has been photographically documented (Davey and Pielke 2005). Neither the 1986 B-44, which was issued to indicate a change in instrumentation/sensor suite, mentions any change in location, nor do any subsequent B-44s. It is therefore likely that the Las Animas site has had an undocumented change in location.

To look at possible undocumented changes at both the Holly and Las Animas stations, the time-of-observation-adjusted annual data were used for these two stations. The annual mean time series of both maximum and minimum temperature at both stations were statistically tested using the following two temperature homogeneity test methods described in Menne and Williams (2005): the standard normal homogeneity test (SNHT) (Alexandersson 1986) and the two-phase regression method with a constant trend model (TPR). Two hundred nearest-neighbor stations were preselected separately for the Holly and Las Animas stations, and then pretested for their statistical homogeneities without using any reference series. Only the annual series of neighboring stations that were identified as homogeneous were selected to create

a reference series to test the Holly and Las Animas stations.

A method based on SNHT (Alexandersson 1986; Ducre-Robitaille et al. 2003) was used for creating reference series from the five most highly correlated and qualified homogeneous neighbors (correlation at least greater than 0.7). Results indicated that the maximum temperature series for Holly and Las Animas identified by the SNHT and TPR methods were homogeneous, but that their minimum temperature series were not (Fig. 1), when the correlations applied were obtained from annual mean temperatures. However, if using correlations calculated from annual maximum and minimum temperatures, the maximum temperature series at both stations were inhomogeneous, as were the minimum temperature series at Las Animas (Table 1 and Fig. 1). While the maximum temperature inhomogeneity was around the time of a documented instrument change, the minimum temperature inhomogeneity was not. It was not possible to create a reference series for minimum temperature at the Holly station because there were no qualified homogeneous neighboring stations with a correlation greater than 0.7 in all of the 200 nearest-neighbor series of minimum temperature. Therefore, undocumented discontinuities likely existed, and their magnitudes (if a step change) were also different from the magnitudes adjusted in the USHCN for annual maximum and minimum series at the two poorly sited Holly and Las Animas sites (Davey and Pielke 2005).

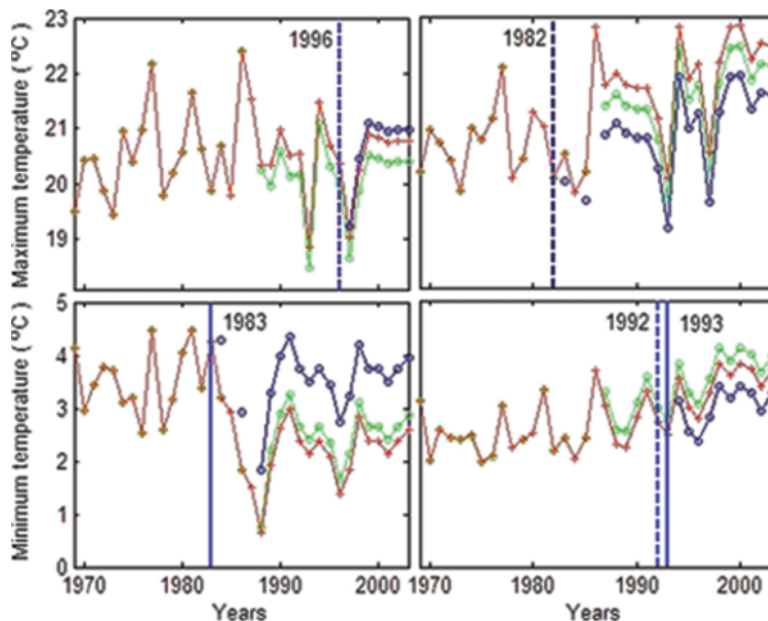


FIG. 1. Annual mean time series of maximum and minimum temperatures at the (left) Holly and (right) Las Animas stations. The red lines are the TOB-adjusted series from USHCN, the green lines are full adjusted series from USHCN (except for urbanization adjustments), and blue lines are adjusted by statistical homogeneity tests starting from the discontinuity + 1 yr. The years along with solid blue vertical lines indicate the positions of statistical discontinuities by using correlations calculated from annual mean homogeneous neighbors and years along with dashed blue vertical lines refer to the positions of statistical discontinuities by using correlations from annual maximum or minimum homogeneous neighbors.

The analysis described in Peterson (2006) excluded Holly, because of an undocumented station change. It is therefore reasonable in hindsight, based on the B-44 form evidence and our statistical analysis, that Las Animas should have been excluded as well.

Other studies have also reported undocumented station changes. Christy (2002), Christy et al. (2006), and Holder et al. (2007, manuscript submitted to *Climate Res.*), for example, discovered several instances in which significant but undocumented

Neighbor stations selected from		Annual mean homogeneous neighbors			Annual Tx or Tn homogeneous neighbors		
Stations	Element	Homogeneity	Position	Magnitude	Homogeneity	Position	Magnitude
Holly, CO	Tx	Homogeneous			Inhomogeneous	1996	-0.57
	Tn	Inhomogeneous	1983	-1.09	No reference*		
Las Animas, CO	Tx	Homogeneous			Inhomogeneous	1982	0.52
	Tn	Inhomogeneous	1993	0.71	Inhomogeneous	1992	0.59

*The reference series was unable to be created from its neighbors.

break points occurred in the individual instrumental records. In one example, for which no documentation was ever found, Athens, Alabama, experienced a spurious 1.5°C warm shift relative to three nearby stations (Christy 2002). Such undocumented inhomogeneities at comparison stations will add further uncertainties to other types of trend adjustments.

UNCERTAINTIES IN ADJUSTMENTS. *Brief background.* In the USHCN, the monthly mean temperatures have been adjusted for the following four factors: 1) an adjustment for the time-of-observation (TOB) bias (Karl et al. 1986), which came about because, at many sites, the observing time has changed during the station's history; 2) a statistical adjustment (instrumentation bias; Quayle et al. 1991, hereafter QUA) to account for the replacement of the Cotton Region Shelter (CRS) by the maximum–minimum temperature system (MMTS); 3) an adjustment based on station moves or relocations (relocation bias; Karl and Williams 1987); and 4) an adjustment for the bias caused by station urbanization (urban bias; Karl et al. 1988). All four adjustments rely heavily on the metadata to identify changepoints. Quality metadata are required for the homogeneity adjustment methods to ensure the robustness of bias modeling, but such historical metadata are not complete. Also, the adjustment can include stations that are not part of the USHCN (online at www.ncdc.noaa.gov/oa/climate/normals/usnormalshist.html). We examine the nature of the uncertainties associated with bias adjustments to the USHCN and the adjustments associated with a subset of five stations (Davey and Pielke 2005; Peterson 2006).

The TOB bias adjustment is the most systematic adjustment with respect to all stations and all time series in the USHCN. From the mean of all stations, both the monthly maximum and minimum temperatures were adjusted upward with time until the mid-1980s. Karl et al. (1986) mentioned that the uncertainties in TOB adjustment are from one-fourth to one-third the magnitude of the TOB bias, which in turn depends on the season and time of observation. However, the evaluation of these TOB biases has indicated that the time-of-observation bias adjustments in USHCN appear to be robust (Vose et al. 2003).

Instrumentation adjustment. The instrumentation adjustment in the USHCN is accomplished with two specific constants that were universally applied at all MMTS stations—one for monthly maximum and one for monthly minimum temperatures. Some concern regarding instrumentation bias for individual sites was raised by Peterson et al. (1998a); the adjustment

“is just a regional average; the exact effect at individual stations may vary somewhat depending on local environmental or climate factors such as the amount of direct sunlight on the shelter,” and this adjustment for instrumentation transition should be reevaluated (Peterson 2003). Pielke et al. (2002) pointed out that the instrumentation bias adjustment in the USHCN is not appropriate for an individual station and that it might increase the heterogeneity of data at individual stations.

To respond to concerns about instrumentation bias adjustments, a subset of data was taken from the TOB-adjusted information in monthly USHCN, and two groups of USHCN stations were selected for this study: MMTS and CRS stations. Station selection was based on 1) no instrument changes being reported for the CRS stations, with only a single CRS-to-MMTS transition in the MMTS stations; 2) no vertical or horizontal station moves being reported; and 3) instrument height for temperature being constantly maintained at 2 m during the selected periods according to metadata files.

The selection procedure for MMTS stations sought not only relatively long MMTS observations, but also equally long observations from the pre-MMTS period. The 116 MMTS and 163 CRS stations were selected, requiring an observation length of 342 months, and the MMTS station length included 171 months for each side of the MMTS-to-CRS transition month (Fig. 2). Both the SNHT and the multiple linear regression (MLR) method (Alexandersson 1986; Peterson et al. 1998b; Ducre-Robitaille et al. 2003) were used for testing the single-most-probable discontinuities in each MMTS series for maximum and minimum temperatures separately. The magnitudes of the metadata-based discontinuity were also estimated using the QUA method. Note that the time series was classified as homogeneous only if the null hypothesis of homogeneity was not rejected at the 95% level using either the SNHT or MLR methods.

At only some of the selected stations did the homogeneity testing indicate a statistically significant inhomogeneity coinciding with the instrument change. Figure 2 shows the average magnitudes of step changes at the discontinuities for the 34 MMTS series of maximum temperature and 24 MMTS series of minimum temperature that were identified as inhomogeneous by the SNHT and MLR tests with identical discontinuity positions (instrument transition dates). For these inhomogeneous series, the result indicates that magnitudes of step changes estimated from the QUA method (not shown) were nearly the same as those estimated by the SNHT or MLR methods, because the reference series

used in the SNHT and MLR were derived in nearly the same way as by the QUA method.

The step changes resulting from the instrument changeover in Fig. 2 for the inhomogeneous series are different from the two constants of -0.38° and $+0.28^{\circ}\text{C}$ applied in the USHCN datasets based on Quayle et al. (1991), and our results indicate that these adjustments vary considerably from station to station, with larger magnitudes for the inhomogeneous series (Figs. 2a and 2b) and relatively smaller magnitudes for the homogeneous series (Figs. 2c and 2d).

The series other than the inhomogeneous and homogeneous series shown in Fig. 2 are either an inho-

mogeneous series, whose most-probable discontinuity according to SNHT and MLR did not match with the metadata (i.e., MMTS installation dates), a series that was not tested because of the over 50% missing data at the candidate sites, or a series with no available neighbor stations (correlations must be larger than 0.7).

Our intent in this section was not to show a network mean of instrumentation bias (because there is a limited MMTS station sampling), but to show, for the identified inhomogeneous series, the discrepancies in step-change magnitudes compared to each adjusted bias in the USHCN MMTS stations where the time period for MMTS observations is of equal length to its predecessor. Note that our results are from only those stations either with a step change large enough to be detected by the homogeneity tests, or where there were no other documented changes during the continuous period. The large step changes shown in the identified inhomogeneous series are not likely the result of changing the sensor and shield alone, but more likely are due to additional, synchronous site microclimate changes (e.g., changes associated with proximity to buildings, site obstacles, and roadways).

Station relocation adjustment. On average, the magnitude of the relocation adjustment is generally as large or larger than other adjustments applied to the USHCN data. Using the studies by Ducre-Robitaille et al. (2003) and DeGaetano (2006) as a basis, an explicit and typical correlation structure for simulation was set up to account for five different neighbor stations and typical interneighbor station correlations. One candidate series and five neighboring series were generated with the correlation matrix R as follows:

$$R = \begin{pmatrix} 1 & 0.95 & 0.90 & 0.90 & 0.85 & 0.80 \\ 0.95 & 1 & 0.85 & 0.85 & 0.85 & 0.85 \\ 0.90 & 0.85 & 1 & 0.85 & 0.85 & 0.85 \\ 0.90 & 0.85 & 0.85 & 1 & 0.85 & 0.85 \\ 0.85 & 0.85 & 0.85 & 0.85 & 1 & 0.85 \\ 0.80 & 0.85 & 0.85 & 0.85 & 0.85 & 1 \end{pmatrix}$$

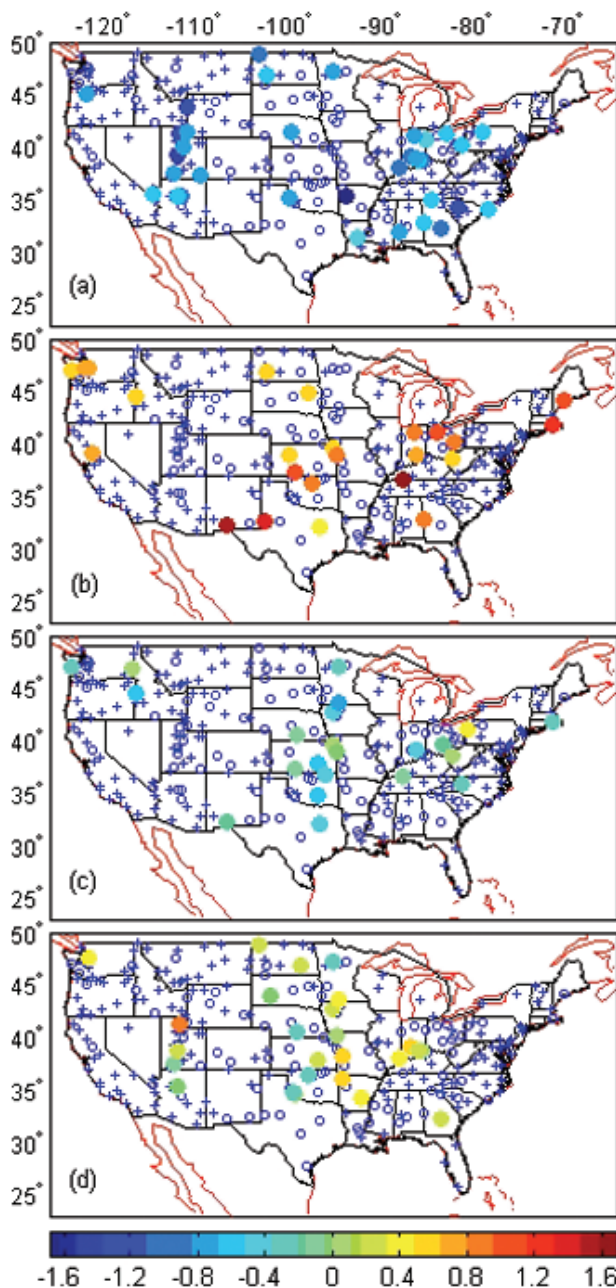


FIG. 2. Average magnitudes of step changes at discontinuities for (a) 34 inhomogeneous MMTS series of maximum temperature, (b) 24 inhomogeneous MMTS series of minimum temperature, (c) the QUA method magnitudes of 27 homogeneous maximum, and (d) the QUA method magnitudes of 24 homogeneous minimum temperature series. The blue open circles are selected 116 MMTS stations and the blue plus symbols are selected 163 CRS stations.

Annual temperature anomaly series were generated to produce 1000 annual time series of 30-yr values each for the simulated candidate station and five simulated neighbor stations. The candidate–neighbor station correlations were preset from 0.95 to 0.8 and the interneighbor station correlations were fixed at 0.85 in order to avoid lower correlations occurring between interneighbor stations during the simulation.

The simulated series were generated for the stations by introducing fields of random temperatures that were normally distributed with a mean of zero and a variance of one. A step change of +0.8°C was imposed at year 15 in each time series. The procedures described for the relocation adjustment (Karl and Williams 1987) then were implemented to estimate the adjustment needed to produce a “homogeneous” time series at the candidate station in each of the 1000 time series. Considering that temperature trends in candidate series may slightly differ from neighboring series in some stations because of the urbanization and land use changes around the candidate stations, the analysis was repeated for each of three imposed trends on the candidate station during the simulation to examine the performance of relocation bias modeling when the candidate series is mean nonstationary. The values of these three imposed trends were selected based on Kalnay and Cai (2003) and Vose et al. (2004).

The result indicates that, on average, the magnitudes of positive relocation bias are overestimated when there is an increasing trend in the simulated time series relative to the neighboring stations (Fig. 3). Larger trends produce larger uncertainties in the relocation bias adjustments when there is a single positive step change. A zero trend at the candidate station is associated on average with the originally introduced step change. However, there is uncertainty, as noted by the spread of the box plot, which is because the averages before and after the step change are not necessarily equal even though random numbers are used in the simulation.

For a negative step change (not shown), the magnitude of the uncertainty in relocation was similar to the positive step-change case, but the magnitude of the step change was underestimated and the degree of underestimation increased with increasing trends. Because a portion of the trend is aliased into the relocation adjustment (DeGaetano 2006), it follows that the trend of the adjusted series is not the same as the original trend imposed on the series.

IMPACT OF ADJUSTMENTS ON ESTIMATED TEMPERATURE TRENDS. In this section, we quantify the impact of the relocation adjustment on *trends* in the climate record.

The relocation adjustment algorithm (Karl and Williams 1987) proceeds from known potential discontinuities in the station histories by computing differences on a seasonal basis between a candidate station and neighboring stations and applying these differences on a monthly basis within the respective season. With each candidate–neighbor pair, mean differences are computed for the intervals before and after a discontinuity, with the intervals extending as far as possible, without spanning another discontinuity in either station. Those differences having the narrowest confidence intervals are used to construct a weighted average of differences, which is then applied to the data of the candidate station prior to the discontinuity. Equation (4) of Karl and Williams (1987) erroneously states that the weights are proportional to the confidence intervals. The National Climatic Data Center (NCDC; M. J. Menne 2006, personal communication) has confirmed that the actual weighting used in the USHCN is with the inverse of the confidence intervals, so that stations with the smallest intervals are weighted most heavily.

To illustrate the effect of this adjustment on estimated climate trends, consider a candidate station whose true secular temperature trend over the interval from 1 to $2N$ years is a_j , but whose data record also includes an artificial jump between years N and $N + 1$ (the middle of the interval) of magnitude j resulting

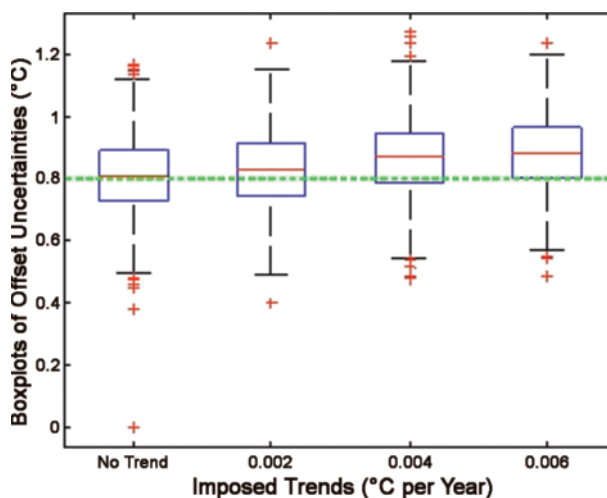


FIG. 3. Box plots of the step-change offsets provided by the relocation model for stationary (leftmost box) and nonstationary series (three trends) with a 0.8°C discontinuity imposed at the 15 yr in 1000 simulated series for each. The box indicates the lower quartile, median, and upper quartile values. The whiskers extend to 1.5 times the interquartile range and outliers are beyond the ends of the whiskers.

from a station move (Fig. 4). In practice, the true trend is not known and is not easily recoverable. Suppose, for convenience, that the temperatures at the neighboring stations to be used in the relocation adjustment match the temperature of the candidate station exactly at the beginning of year 1, but possess a secular temperature trend of a_n . The average temperature difference $d_{\text{bCAN-NEIGH}}$ between the station and its neighbors in the interval after the jump is $1.5N(a_c - a_n) + j$, while the average temperature difference $d_{\text{aCAN-NEIGH}}$ before the jump is $0.5N(a_c - a_n)$. The relocation correction applied to the data record prior to the jump is $d_{\text{bCAN-NEIGH}} - d_{\text{aCAN-NEIGH}} = N(a_c - a_n) + j$, which exceeds the proper correction j by a factor proportional to the difference in trends between the candidate station and its neighbors. This overcorrection causes the homogenized data record at the candidate station to underestimate the true climate trend at that location. In fact, it can be shown that the least squares trend of the homogenized data is a_n ; the true temperature trend at the candidate station is replaced by the temperature trend from the neighboring stations. Thus, the adjustment has the effect of removing any trend information (a_c) that might be present in the original candidate station data during the 10-yr adjustment window and replacing it with a_n .

Peterson (2006) tested the validity of the data from poorly sited stations by comparing the trends from homogeneity-corrected poorly sited stations with the trends from nearby well-sited homogeneous stations. However, the adjustments “produce time series that are indeed representative of the climate variability and change in the region,” in part because the relocation adjustment replaces local information regarding climate change, which is itself contaminated by a station move, with the climate change information from other stations in the region. Because the homogeneity adjustment artificially forces climate trends at adjusted stations to be regionally representative (the same trend as the reference series), the fact that the adjusted trends are consistent with some of the reference series trends should not then be used to demonstrate the validity of the homogeneity adjustment.

How serious is this problem for estimates of regional climate change? Because the relocation adjustment replaces segments of the climate change record at some stations with adjustments from surrounding stations, it is unlikely to introduce a bias into climate change estimates unless, as a whole, the candidate stations exhibit a different partitioning of energy than the surrounding stations. However, not only does the relocation adjustment

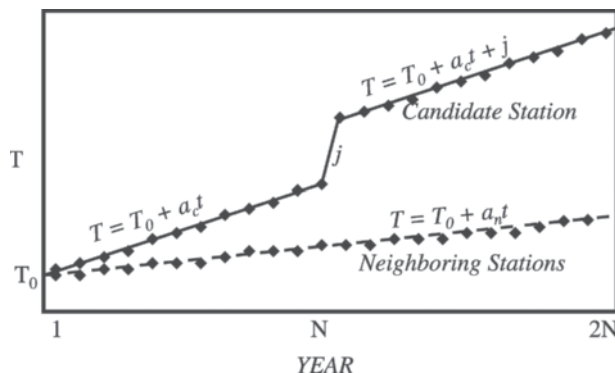


FIG. 4. Sample time series of temperature at a candidate station with a known discontinuity, together with the average time series of temperature at neighboring stations.

reduce the number of independent observations, but it also creates dependent observations from an average of surrounding stations and treats them as independent. Both effects lead to a false sense of confidence in the accuracy of estimates of regional climate change.

COMPARISON WITH REANALYSIS. As shown in a series of publications (Kalnay and Cai 2003, KC hereafter; Cai and Kalnay 2004, 2005; Zhou et al. 2004; Frauenfeld et al. 2005; Lim et al. 2005, 2006; Kalnay et al. 2006), reanalyses [particularly the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (NNR)] can be regarded as an independent estimate of the surface temperature variability, including trends associated with both the large-scale circulation variability and anthropogenic radiative forcings. Because the reanalysis does not reflect the temperature variability and trend due to the local surface properties and the potential biases resulting from “poor siting,” the reanalysis provides an alternative tool to detect and possibly correct the nonclimatic biases resulting from changes in observation practice and poor siting, as suggested in Kalnay et al. (2006).

There have been two published criticisms regarding KC. Vose et al. (2004) argued that KC used the raw surface observations, which have not taken the nonclimatic observation problems into consideration, such as a change of instruments, observation sites, and observation time. However, this is not a criticism of the usage of the reanalysis data itself, but rather the usage of the unadjusted observations.

As pointed out in the rebuttal by Cai and Kalnay (2004), because those adjustments increase the warming trends in the ground observations, the

inclusion of the adjustments would *not* have altered the overall positive sign of the estimate of land use change climate impact (Cai and Kalnay 2004).

Another criticism is that the NNR does not include the observed temporal increases in atmospheric CO₂ in the model, and that water vapor and cloud feedbacks associated with the anthropogenic radiative forcing are not accurately represented in the model. As a result, the NNR would underestimate the surface warming trend (Trenberth 2004).

Cai and Kalnay (2005) showed with a simple analytical study, however, that the reanalysis can capture essentially the full strength of temperature trends caused by the increase of greenhouse gases even if this forcing is absent from the model used in the data assimilation. The work by Andersen et al. (2001) clearly confirms that through data assimilation, the 15-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-15) can capture the Mt. Pinatubo eruption within a few days thereafter even though the model used in ERA-15 has constant aerosols. It follows that the particular issue raised by Trenberth (2004) has little implication regarding the fidelity of the long-term trend in the reanalysis. Furthermore, the publication of Lim et al. (2005) confirmed the main finding of KC by using two different adjusted station observations [the Climate Research Unit (CRU) and Global Historical Climate Network (GHCN) datasets] and two independent reanalysis products [NNR and 40-yr ECMWF Re-Analysis (ER-40)], one of which (ERA-40) includes the observed temporal increases in atmospheric CO₂ in the model.

In this section, we explore whether the reanalysis could be also used to assess to what extent the homogeneity adjustments made to the original observations can “correct” the nonclimatic biases in these poorly sited stations. The main questions to be examined in this section are as follows: i) Are the adjustments made to poor siting observations reported in Peterson (2006) consistent with the reanalysis? ii) Do these adjustments yield additional information about the temperature variability and long-term trend of station data?

The original (and nonurban) adjusted station observations were downloaded from the NCDC Web site (online at www.ncdc.noaa.gov/ol/climate/research/ushcn/ushcn.html). The following two sets of reanalysis datasets are used: daily data of maximum and minimum temperatures derived from the (global) NNR I (Kalnay et al. 1996; Kistler et al. 2001) and the

monthly average data of the daily mean temperature derived from the North American Regional Reanalysis (NARR; Mesinger et al. 2006). (The NARR data are downloaded from <http://nomads.ncdc.noaa.gov>, which only archive the daily mean temperature, but not maximum and minimum temperatures.)

Following Kalnay and Cai (2003), the NNR and NARR temperatures are interpolated from reanalysis grids to individual observation sites. Because most of the nonclimatic changes in the five stations took place after 1978, and there is a potential impact on the NNR climate trend because of the change of the observation system in 1979, we focus on the comparison between the NNR and surface observations for the period from 1979 to 1999.

The monthly NNR temperature anomalies are obtained by removing the annual cycle defined from the period from 1979 to 1999. As a result, any potential NNR trend bias prior to 1979 would have little impact on the monthly temperature anomaly fields to be used in this study. The monthly NARR temperature anomalies are obtained by removing the annual cycle defined in the period from January 1979 to December 2003. The monthly observation anomalies are the departures from the annual cycles of the station observations defined in the period from 1971 to 2000, as in Peterson (2006).¹

Table 2 lists the numerical values of the correlation and root-mean-square (RMS) difference between the monthly anomalies of the maximum and minimum temperature fields derived from the NNR, along with the unadjusted and adjusted station observations at the five locations. Due to their high elevations (ranging from 1033 m at Las Animas to 1839 m at Trinidad), the correlation between the NNR and station observations is smaller than that reported in Kalnay and Cai (2003).

It is found that the correlation is higher for the two well-sited stations (Trinidad and Cheyenne Wells; see Davey and Pielke 2005), despite the fact they are the two highest in elevation among the five stations. It is also apparent that the RMS differences between the NNR and original (unadjusted) observations at these two well-sited stations are the smallest. At Cheyenne Wells, the RMS differences of the monthly T_{\max} and T_{\min} anomalies are 1.5° and 0.86°C, respectively, whereas at Trinidad, whose elevation is nearly twice as high as the other four stations, the RMS differences of the monthly T_{\max} and T_{\min} anomalies are 1.5° and 1.08°C, respectively. Among the three poorly sited

¹ The adjusted station data have their own annual cycles, which are different from the original station data. Therefore, we have also removed the annual cycle resulting from the adjustments from each adjusted time series to obtain the monthly anomalies of the adjusted data.

stations (Eads, Lamar, and Las Animas; see Davey and Pielke 2005), the smallest RMS difference of the monthly T_{\max} and T_{\min} anomalies are 1.81° (at Lamar) and 0.91°C (at Las Animas) and the largest ones are 2.01° (T_{\max} at Eads) and 1.49°C (T_{\min} at Lamar). This comparison indicates that the NNR does yield valuable information about the poorly sited inhomogeneous effect on individual station observations.

Now we compare the NNR and adjusted temperatures. Because the only adjustment made at Trinidad is the climatological seasonal cycle, the difference between unadjusted and adjusted temperature anomalies at Trinidad is essentially zero. It can be seen that the adjustment made at Cheyenne Wells slightly increases the RMS difference of T_{\max} anomalies and has little impact on T_{\min} anomalies. Note that unlike Trinidad, the geographic location of Cheyenne Wells is very close to the three poorly sited stations (less than a half degree in latitude/longitude and less than 62 m in elevation). Therefore, we use the comparison between the NNR and unadjusted temperature anomalies at Cheyenne

Wells as a reference to measure the success of the adjustments made to the original observations at the three poorly sited stations.

As indicated in Table 2, the adjustments at the three poorly sited stations do significantly reduce the difference between the monthly NNR and station temperature anomalies. The largest improvement resulting from the adjustments takes place at Eads, where the adjustments reduce the RMS difference of the monthly T_{\max} anomalies by 0.34°C (from 2.01° to 1.67°C) and T_{\min} anomalies by 0.51°C (from 1.40° to 0.89°C).

The anomaly correlation between NARR and station observations reaches the 90% level, which is about 10% higher than that evaluated using the NNR (Table 3). The RMS difference between the NARR and the station observations also is much smaller than that between the NNR and station observations. As indicated in Table 4, the amplitude of the reanalysis anomalies is very close to that of the observation (the amplitude of NNR anomalies is slightly smaller, about 0.2°C, than the observations).

TABLE 2. Anomaly correlation and RMS difference between the monthly anomalies of the NNR and station-unadjusted observations (adjusted observations in bold).

Station	AC (T_{\max})	RMS (T_{\max})	AC (T_{\min})	RMS (T_{\min})
Trinidad	72% (72%)	1.50 (1.50) °C	71% (71%)	1.08 (1.08) °C
Cheyenne Wells	79% (78%)	1.51 (1.57) °C	85% (85%)	0.86 (0.86) °C
Las Animas	69% (73%)	1.88 (1.80) °C	80% (81%)	0.91 (0.89) °C
Eads	69% (75%)	2.01 (1.67) °C	69% (82%)	1.40 (0.89) °C
Lamar	70% (73%)	1.81 (1.71) °C	64% (79%)	1.49 (0.99) °C

TABLE 3. Anomaly correlation and RMS difference between the monthly anomalies of the NARR and station-unadjusted observations (adjusted observations in bold).

Station	AC	RMS
Trinidad	90% (90%)	0.72 (0.72) °C
Cheyenne Wells	94% (94%)	0.62 (0.61) °C
Las Animas	92% (92%)	0.80 (0.73) °C
Eads	93% (95%)	0.92 (0.57) °C
Lamar	91% (94%)	0.79 (0.66) °C

TABLE 4. Std dev of the monthly mean temperature anomalies for the period of 1979–99.

Station	Raw obs	Adjusted obs	NNR	NARR
Trinidad	1.63 °C	1.63 °C	1.46 °C	1.53 °C
Cheyenne Wells	1.87 °C	1.87 °C	1.66 °C	1.91 °C
Las Animas	1.82 °C	1.85 °C	1.51 °C	1.86 °C
Eads	1.96 °C	1.81 °C	1.67 °C	1.89 °C
Lamar	1.87 °C	1.86 °C	1.55 °C	1.89 °C

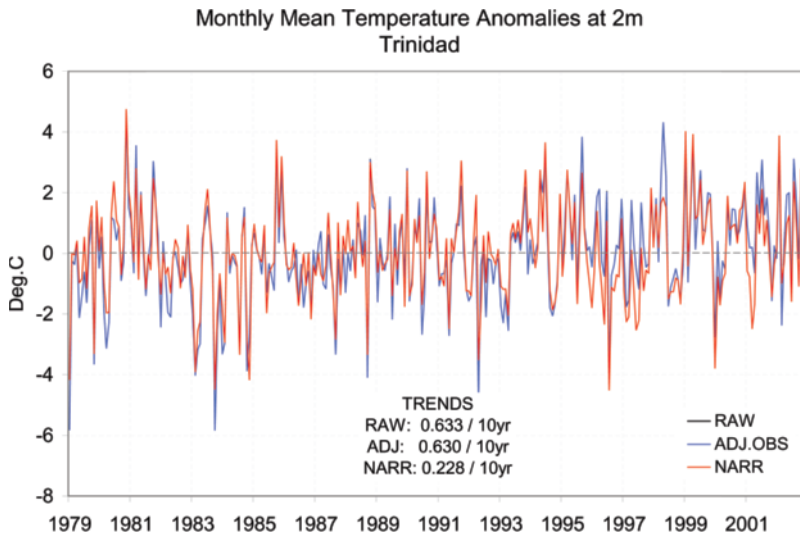


FIG. 5. Monthly mean temperature anomalies (curves; unit: °C) at Trinidad.

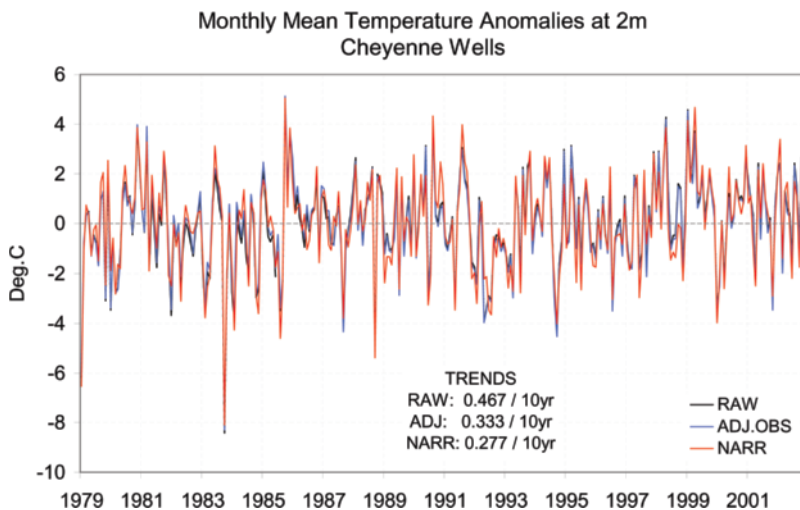


FIG. 6. Same as Fig. 5, except for the Cheyenne Wells station.

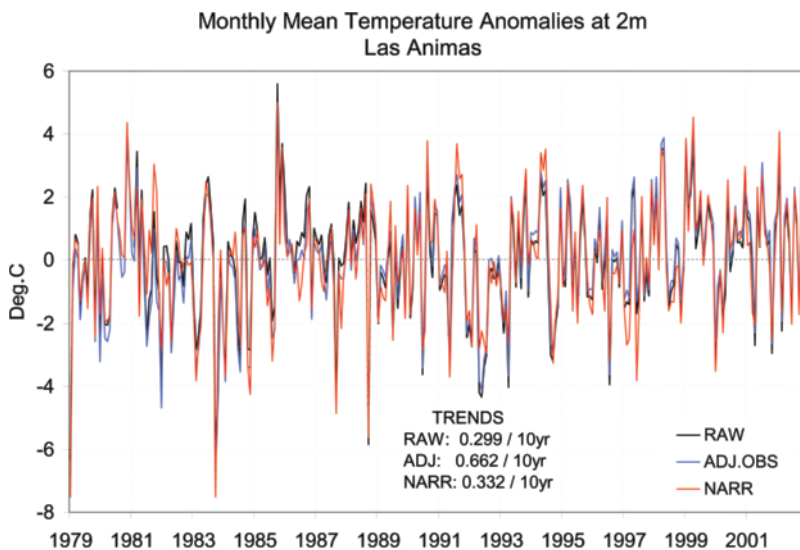


FIG. 7. The same as Fig. 5, except for the Las Animas station.

Figures 5–9 clearly show that the NARR faithfully captures the intraseasonal and interannual variability of the station observations (the station locations are shown in Fig. 10). Nevertheless, the difference between the NARR and (unadjusted) observations for the poorly sited inhomogeneous stations is larger than that for Cheyenne Wells, a well-sited station that is the nearest to the poorly sited stations with a similar elevation, consistent with the NNR results.

Therefore, the NARR can also provide valuable information about the poorly sited inhomogeneous effect on individual station observations. In addition, it is obvious that the adjustments to the observations at the poorly sited stations improve the station data because the adjusted temperatures are better correlated with the NARR and present smaller RMS differences. Therefore, we conclude that the adjustments indeed correct a large portion of nonclimatic biases in these poorly sited stations as far as the difference between the NARR/NNR and station data are concerned.

Now, let us turn our attention to the climate trend comparison. Table 5 lists the trends of the unadjusted, adjusted, NNR, and NARR monthly temperature anomalies at these five stations. All trends are calculated for the period of 1979–99, the longest common period of all of the data, because our NNR collection does not include any data after 1999. We also estimate the significance of these trends and the difference between these trends at each station. The statistical significance test on the trend in a time series is evaluated by calculating the ratio between the estimated trend and its standard error. The standard error can be evaluated

by the standard deviation of the residuals about the regression line. In general, the residuals are not statistically independent because of the serially correlated time series. The “effective sample size” suggested in Zwiers and Von Storch (1995) is used to adjust the standard error of the trend. The adjusted standard error is then used to assess the significance of the individual trends and the difference between two trends. It is seen that the trends derived from the NNR dataset are all significant at the 5% level, except at Trinidad whose trend is small compared to the other four stations. It is of importance to note that the NNR trends that are significant vary little from one station to another, reflecting the fact that the NNR in general only captures the trend on the large scale, and reveals little information about the local effects. The NNR trend in this region is about $0.45^{\circ}\text{C} (10 \text{ yr})^{-1}$.

The NARR trend for this region during the period of 1979–99 is quite small, and none of the linear trends derived from the NARR dataset is significant, even at the 10% level, because the NARR trends are smaller compared to their temporal variability.

The fact that the trends derived from the NNR analysis for the nearby stations are quite uniform indicates that more spatially representative trends are obtained from the reanalyses. Moreover, because there is a relatively large variance in the trends derived from (unadjusted/adjusted) station data, a “regional” trend inferred from just one or even a few sites is merely a sample from a statistical distribution of the trend values.

Below we focus only on the comparison between the NNR and station data trends. The trends derived

Monthly Mean Temperature Anomalies at 2m
Eads 2S

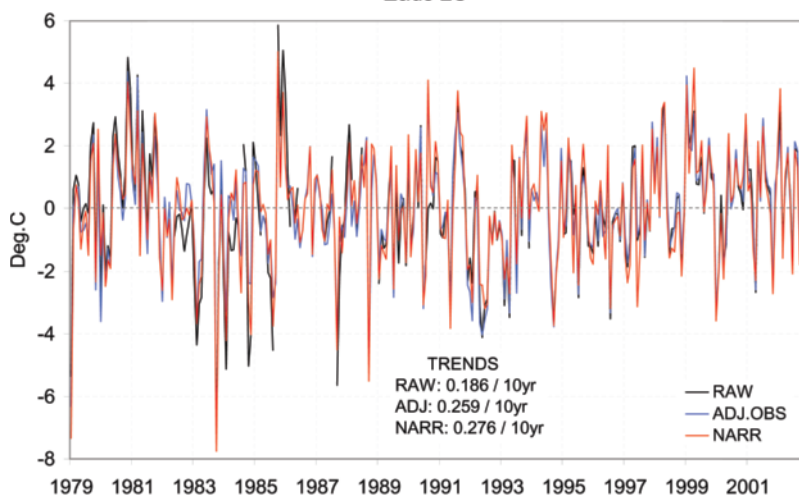


Fig. 8. The same as Fig. 5, except for the Eads station.

Monthly Mean Temperature Anomalies at 2m
Lamar

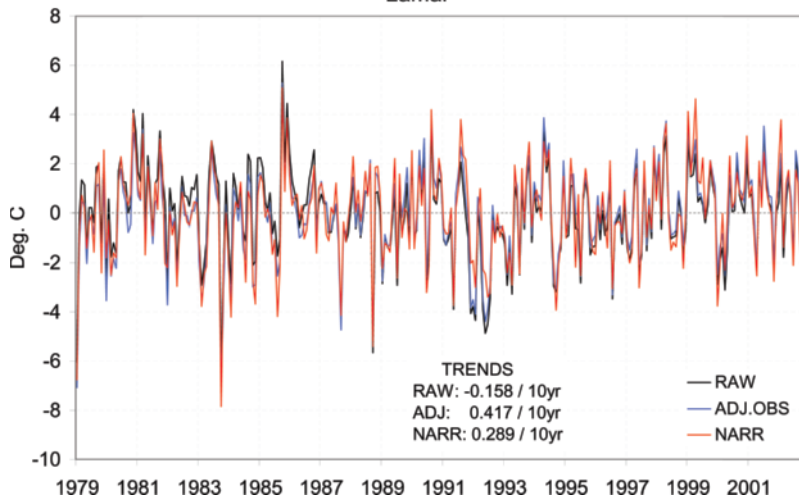


Fig. 9. The same as Fig. 5, except for the Lamar station.

Fig. 10. Station locations for Figs. 5–9. The elevations of each station are Trinidad: 1838 m (6030 ft), Cheyenne Wells: 1295 m (4250 ft), Las Animas: 1186 m (3890 ft), Eads: 1285 m (4215 ft), and Lamar: 1105 m (3627 ft).

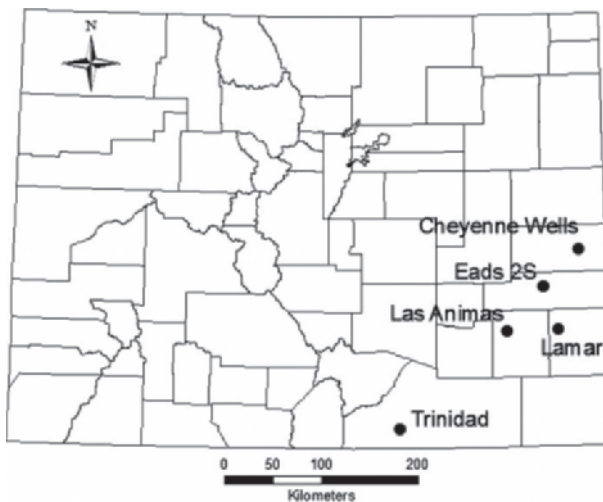


TABLE 5. Linear regression trends (1979–99) of the monthly mean temperature anomalies [$^{\circ}\text{C}$ (10 yr) $^{-1}$]. The bold number indicates the trend is significant at the 5% level. The trend with superscript “*n*” (*n* = 1, 2, 3, 4, representing the column index) indicates the difference between that trend and the other trend at the *n*th column is significant at the 5% level (superscripts are only assigned to those trends that are significant at the 5% level).

Station	Raw obs	Adjusted obs	NNR	NARR
Trinidad	0.48 ^{3,4}	0.48 ^{3,4}	0.23	0.04
Cheyenne Wells	0.40 ^{2,4}	0.23	0.47 ^{2,4}	0.09
Las Animas	0.10	0.60 ^{1,4}	0.41 ^{1,4}	0.12
Eads	0.0	0.04	0.49 ^{1,2,4}	0.06
Lamar	-0.47	0.27	0.43 ^{1,4}	0.08

from the unadjusted station data vary greatly from one station to another, ranging from -0.47° to 0.48°C per 10 yr, while the adjusted data vary from 0.04° to 0.48°C (10 yr) $^{-1}$. At Trinidad, the NNR trend is smaller and not significant, perhaps reflecting the fact the NNR’s elevation is higher than the elevation at that site. At Cheyenne Wells, the NNR trend is very close to the trend of the original station observation. The adjustment made at Cheyenne Wells reduces the trend of the unadjusted data by about 40% and makes it insignificant. Together with the fact that the adjustments at Cheyenne Wells also move the adjusted data further away from the NNR (see row 2 of Table 2), this seems to suggest that the adjustments could introduce some inconsistencies to the observations at Cheyenne Wells, a well-sited station.

At the three poorly sited stations, none of the trends in the unadjusted data are significant. The adjustments increase the trends at the three poorly sited stations. One of them (at Las Animas) is significant at the 5% level and equals 0.6°C (10 yr) $^{-1}$, which is about 20% stronger than the NNR trend over this region. Furthermore, the difference between the adjusted and the NNR trends at this station is not statistically significant. Therefore, the adjustment at Las Animas helps to improve the correlation with the reanalysis data not only at monthly and yearly time scales, but also in terms of the long-term trend.

LAND USE/LAND COVER CHANGE ISSUES. There are three primary issues related to land use/land cover (LULC) and changes in LULC related to placement of climate stations. First, a station may be initially placed in what might be considered a poor LULC environment (e.g., near a highway or other man-made environment that could influence the observed temperature based on day of week, holiday, etc.). Second, a station may have been initially located at what might be considered a good

LULC environment only to have that environment change over time. And third, possibly due to one of the above situations, a station may be moved from one LULC environment to another.

Peterson (2006) examined an admittedly “small subset” of USHCN stations (5 of >1,200 stations, i.e., less than 0.5% of the total stations) and concluded that this is evidence for which “if poor siting causes a bias, homogeneity adjustments account for the biases.” Other evidence, however, shows that LULC differences or changes can introduce issues not addressed in the routinely applied USHCN adjustments (e.g., Peterson 2006), or other adjustments designed to account for horizontal or vertical differences in station locations (e.g., Peterson 2003).

Routinely made climate station adjustments often include adjustments for station moves, as discussed earlier in this paper. These adjustments can be based on the temperature records of other stations within the vicinity of the candidate station (the one that moved) that have “no documented changes in the five or more years on either side of that date” of the move (Peterson 2006). Additionally, proposed adjustments associated with differences in station locations with respect to each other (locational differences) have included adjustments for horizontal (latitude) or vertical (elevation) differences (e.g., Peterson 2003). These methodologies generally do not include any adjustments that reflect environmental differences that may exist between station locations or differences resulting from station moves from one land cover type to another [indeed, Peterson (2003) was attempting to examine just such differences with his adjustments].

Adjustments are not typically made for the environmental (and related temperature) changes that can occur at a station that has a constant location, yet experiences changes in LULC over time. Gallo et al. (1996) pointed out how the potential impact of

changes from rural to urban LULC would be expected to result in decreases in the diurnal temperature range (DTR), and how these LULC-induced changes might confound temperature change analyses.

The Trends Project temperature analysis (Hale et al. 2006) examined “normals” (National Climatic Data Center 2002) temperature data for stations near sample blocks in which LULC has been determined for five dates during the period from 1973 to 2000. The normals temperature data have been adjusted for time-of-observation biases based on results of Karl et al. (1986) and have also undergone quality control (Peterson et al. 1998a). Within this dataset, inhomogeneities in the temperature data have been addressed based on recommendations of Peterson and Easterling (1994) and Easterling and Peterson (1995). Hale et al. (2006) examined temperature trends at the normals stations before and after periods of dominant LULC change. Temperature trends were primarily insignificant prior to the period during which the greatest single type of LULC change occurred around normals stations. Additionally, those trends that were significant were equally divided between warming and cooling trends.

However, after periods of dominant LULC change, significant trends in minimum, maximum, or mean temperature were far more common, and 95% or more of these significant trends were warming trends. Although the LULC changes have not been identified as the causative factor in the exhibited temperature trends, there is substantial evidence for such speculation. This issue is relevant to the Peterson (2006) analysis because the photographs in Davey and Pielke (2005) suggest that the landscape (and thus the microclimate) around the poorly-sited measurement location (and even the well-sited locations) is not likely to be static.

The general application of adjustments of temperature data to individual stations based on relationships developed over a large sample of stations should also be cause for some concern, as suggested by Gallo (2005). Gallo (2005) applied the adjustments for locational differences (horizontal or vertical; Peterson 2003) to five pairs of Climate Reference Network (CRN) stations that had no differences in instruments or observation times. The distance between the pairs of stations ranged from 5 to 30 km. The expected differences in mean annual temperatures for the stations based on the locational adjustments differed from those actually observed by -0.37° to 1.35°C . These results suggest that microclimate influences, including differences in LULC, are potentially greater than what might be anticipated from differences in

station location. The results may further suggest that at least some adjustments applied to station temperature data for locational differences (e.g., Peterson 2003) are not applicable in all situations. Additionally, the adjustments derived for the most part from analysis of a large number of stations might only be expected to be appropriate when applied (and error assessed) over a large number of samples, rather than individual station pairs. Based on these results, assessment and potential inclusion of adjustments for microclimate influences within USHCN adjustments is recommended for consideration.

Runnalls and Oke (2006) also present a methodology for the detection of inhomogeneities in temperature records associated with changes in LULC (e.g., “vegetation growth, or encroachment by built features such as paths, roads, runways,” etc.), and related factors that can be “microscale and subtle.” Gallo et al. (1996) recommended that the land use/land cover at climate stations be monitored like any other variable that might introduce an inhomogeneity in the data record. A future solution to the LULC influences on station temperature records may exist with the increased resolution of satellite-based remote sensing systems and the products under development from these systems. The National Land Cover Database (Homer et al. 2004) provides LULC information for 29 classes of land cover for the conterminous United States at a spatial resolution of 30 m for 1992 and 2001. It is anticipated that future versions of this database will be available.

This database could potentially be used to monitor LULC change at all climate stations from 1992 forward, and provide recommendations for those stations that might be candidates for temperature record adjustments based on LULC change at or near the stations [in addition to those identified by the methodologies of Runnalls and Oke (2006)].

This database could also potentially provide assessment of regional LULC change associated with station locations such that if the LULC change observed at or near a station was, in reality, a change that is taking place on a regional basis, then the station temperature record might be considered indicative of the true climate change of the region and not an anomalous change in a trend specific to an individual station. Thus, temperature adjustments may not be appropriate for a station that truly represents the LULC change that has occurred within a region, rather than LULC that is site specific at or near a single station.

CONCLUSIONS. As Davey and Pielke (2005) documented and Peterson (2006) acknowledges,

several USHCN stations are poorly sited or have siting conditions that change over time. These deficiencies in the observations should be rectified at the source, that is, by correcting the location and then ensuring high-quality data that are locally and, in aggregate, regionally representative. Station micrometeorology produces complex effects on surface temperatures, however, and, as we show in this paper, *attempting to correct the errors with existing adjustment methods artificially forces toward regional representativeness and cannot be expected to recover all of the trend information that would have been obtained locally from a well-sited station.*

The comparison of the reanalysis with the unadjusted and adjusted station data indicates that the reanalysis can be used to detect the inhomogeneity of individual station observations resulting from non-climatic biases. In general, the adjustments indeed correct a large portion of nonclimatic biases in these poorly sited stations as far as the difference between the NARR/NNR and station data is concerned. The NNR yields a relatively uniform and statistically significant trend in this region, which is statistically similar to two of the four station trends. However, we found that there are some inconsistencies in the trends of the adjusted data. Among the four stations that have been subjected to adjustments, only the adjusted trend at Lamar is consistent with the NNR trend (being statistically similar). The other three adjustments either make the consistent trend (Cheyenne Wells) statistically inconsistent, produce a statistically significant larger trend than for the surrounding stations (Las Animas), or cause little change in the trend (Eads). This leads us to conclude that, whereas the adjustments do improve the consistency among the nearby station data and reduce the differences with respect to the reanalysis at the monthly and yearly scales, the trends of the adjusted data are often inconsistent among closely located stations.

Peterson's approach and conclusions, therefore, provide a false sense of confidence with these data for temperature change studies by seeming to indicate that the errors can be corrected. For instance, the dependence of the corrections on other information (such as regional station moves, which in itself has been found on occasion to be inaccurate) can be considered an indication of the uncertainty and limitations of the "corrective approach" that is being sought. As a requirement, the statistical uncertainty associated with the effect of the adjustments on the regional temperature record needs to be quantified and documented.

Temperature adjustments such as those resulting from change in instrumentation are, of course, necessary. However, the results shown in this paper demonstrate that the lack of correctly and consistently sited stations results in an inherent uncertainty in the datasets that should be addressed at the root, by documenting the micrometeorological deficiencies in the sites and adhering to sites that conform to standards such as the Global Climate Observing System (GCOS) Climate Monitoring Principles (online at http://gosc.org/GCOS/GCOS_climate_monitoring_principles.htm). A continued mode of corrections using approaches where statistical uncertainties are not quantified is not a scientifically sound methodology and should be avoided, considering the importance of such surface station data to a broad variety of climate applications as well as climate variability and change studies.

ACKNOWLEDGMENTS. This study was in part funded by the F/DOE/The University of Alabama in Huntsville's Utilization of Satellite Data for Climate Change Analysis (DE-FG02-04ER 63841 through the University of Alabama at Huntsville), National Science Foundation (ATM-0233780, ATM-0403211, DEB 0217631), National Aeronautics and Space Administration (NNG05GQ47G, NNG04GL61G, NNG04GB87G), USGS Grant 04CRAG0032, Mary S. Rice Fund/Purdue University. Any opinions, findings, and conclusions are those of the authors and do not necessarily reflect the views of the Illinois State Water Survey. Dallas Staley did her usual outstanding job in editing the paper for its final version. The authors would also like to thank Dr. Thomas C. Peterson and the two anonymous reviewers for their detailed review and comments.

REFERENCES

- Alexandersson, H., 1986: A homogeneity test applied to precipitation data. *J. Climatol.*, **6**, 661–675.
- Andersen, U., E. Kaas, and P. Alpert, 2001: Using analysis increments to estimate atmospheric heating rates following volcanic eruptions. *Geophys. Res. Lett.*, **28**, 991–994.
- Cai, M., and E. Kalnay, 2004: Response to the comments by Rose et al. and Trenberth. *Nature*, **427**, 214.
- , and —, 2005: Can reanalysis have anthropogenic climate trends without model forcing? *J. Climate*, **18**, 1844–1849.
- Christy, J. R., 2002: When was the hottest summer? A State Climatologist struggles for an answer. *Bull. Amer. Meteor. Soc.*, **83**, 723–734.

- , W. Norris, K. Redmond, and K. Gallo, 2006: Methodology and results of calculating central California surface temperature trends: Evidence of a human-induced climate change. *J. Climate*, **19**, 548–563.
- Davey, C. A., and R. A. Pielke Sr., 2005: Microclimate exposures of surface-based weather stations—Implications for the assessment of long-term temperature trends. *Bull. Amer. Meteor. Soc.*, **86**, 497–504.
- DeGaetano, A. T., 2006: Attributes of several methods for detecting discontinuities in mean temperature series. *J. Climate*, **19**, 838–853.
- Ducré-Robitaille, J.-R., L. A. Vincent, and G. Boulet, 2003: Comparison of techniques for detecting of discontinuities in temperature series. *Int. J. Climatol.*, **23**, 1087–1101.
- Easterling, D. R., and T. C. Peterson, 1995: A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol.*, **15**, 369–377.
- , T. R. Karl, E. H. Mason, P. Y. Hughes, D. P. Bowman, R. C. Daniels, and T. Boden, Eds., 1996: United States Historical Climatology Network (U.S. HCN). ORNL Rep. ORNL/CDIA-87 NDP-019/R3, 214 pp.
- Frauenfeld, O. W., T. Zhang, and M. C. Serreze, 2005: Climate change and variability using European Centre for Medium-Range Weather Forecasts reanalysis (ERA-40) temperatures on the Tibetan Plateau. *J. Geophys. Res.*, **110**, D02101, doi:10.1029/2004JD005230.
- Gallo, K. P., 2005: Evaluation of temperature differences for paired stations of the U.S. Climate Reference Network. *J. Climate*, **18**, 1629–1636.
- , D. R. Easterling, and T. C. Peterson, 1996: The influence of land use/land cover on climatological values of the diurnal temperature range. *J. Climate*, **9**, 2941–2944.
- Hale, R. C., K. P. Gallo, T. W. Owen, and T. R. Loveland, 2006: Land use/land cover change effects on temperature trends at U.S. Climate Normals stations. *Geophys. Res. Lett.*, **33**, L11703, doi:10.1029/2006GL026358.
- Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan, 2004: Development of a 2001 national land-cover database for the United States. *Photogramm. Eng. Remote Sens.*, **70**, 829–840.
- Kalnay, E., and M. Cai, 2003: Impact of urbanization and land use on climate change. *Nature*, **423**, 528–531.
- , and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- , M. Cai, H. Li, and C. J. Tobin, 2006: Estimation of the impact of land-surface forcings on temperature trends in eastern United States. *J. Geophys. Res.*, **111**, D06106, doi:10.1029/2005JD006555.
- Karl, T. R., and C. N. Williams Jr., 1987: An approach to adjusting climatological time series for discontinuous inhomogeneities. *J. Climate Appl. Meteor.*, **26**, 1744–1763.
- , —, 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.
- , H. F. Diaz, and G. Kukla, 1988: Urbanization: Its detection and effect in the United States climate record. *J. Climate*, **1**, 1099–1123.
- Kistler, R., and Coauthors, 2001: The NCEP/NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247–267.
- Lim, Y.-K., M. Cai, E. Kalnay, and L.-M. Zhou, 2005: Observational evidence of sensitivity of climate changes to land types and urbanization. *Geophys. Res. Lett.*, **32**, L22712, doi:10.1029/2005GL024267.
- , —, and —, 2006: Impact of vegetation types on surface climate change. *J. Appl. Meteor. Climatol.*, submitted.
- Mahmood, R. L., S. A. Foster, and D. Logan, 2006: The GeoProfile metadata, exposure of instruments, and measurement bias in climatic record revisited. *Int. J. Climatol.*, **26**, 1091–1124.
- Menne, M. J., and C. N. Williams, 2005: Detection of undocumented changepoints using multiple test statistics and composite reference series. *J. Climate*, **20**, 4271–4286.
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360.
- Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.*, **25**, 693–712.
- National Climatic Data Center, 2002: Data documentation for data set 9641C (DSI-9641C). Monthly station normals of temperature, precipitation, and degree days, and precipitation probabilities and quantiles: 1971–2000. 21 pp. [Available online at <http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td9641c.pdf>.]
- Peterson, T. C., 2003: Assessment of urban versus rural in situ surface temperatures in the contiguous United States: No difference found. *J. Climate*, **16**, 2941–2959.

- , 2006: Examination of potential biases in air temperature caused by poor station locations. *Bull. Amer. Meteor. Soc.*, **87**, 1073–1089.
- , and D. R. Easterling, 1994: Creation of homogeneous composite climatological reference series. *Int. J. Climatol.*, **14**, 671–679.
- , and Coauthors, 1998a: Homogeneity adjustments of *in situ* atmospheric climate data: A review. *Int. J. Climatol.*, **18**, 1493–1517.
- , R. Vose, R. Schmoyer, and V. Razuvaev, 1998b: Global Historical Climatology Network (GHCN) quality control of monthly temperature data. *Int. J. Climatol.*, **18**, 1169–1179.
- Pielke, R. A., Sr., and Coauthors, 2002: Problems in evaluating regional and local trends in temperature: An example from eastern Colorado, USA. *Int. J. Climatol.*, **22**, 421–434.
- 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. Amer. Meteor. Soc.*, **72**, 1718–1723.
- Runnalls, K. E., and T. R. Oke, 2006: A technique to detect microclimatic inhomogeneities in historical records of screen-level air temperature. *J. Climate*, **19**, 959–978.
- Trenberth, K. E., 2004: Rural land-use change and climate. *Nature*, **427**, 213.
- Vose, R. S., C. N. Williams Jr., T. C. Peterson, and T. R. Karl, 2003: An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network. *Geophys. Res. Lett.*, **30**, 2046, doi:10.1029/2003GL018111.
- , T. R. Karl, D. R. Easterling, C. N. Williams, and M. J. Menne, 2004: Impact of land-use change on climate. *Nature*, **427**, 213–214.
- Zhou, L., R. E. Dickinson, Y. Tian, J. Fang, Q. Li, R. K. Kaufmann, C. J. Tucker, and R. B. Myneni, 2004: Evidence for a significant urbanization effect on climate in China. *Proc. Nat. Acad. Sci.*, **101**, 9540–9544.
- Zwiers, F., and H. Von Storch, 1995: Taking serial correlation into account in test of the mean. *J. Climate*, **8**, 336–351.