Effects of Recent Thermometer Changes in the Cooperative Station Network

Robert G. Quayle
David R. Easterling
Thomas R. Karl
Pamela Y. Hughes
Global Climate Laboratory
NOAA/NESDIS/National Climatic Data Center
Asheville, NC 28801

Abstract

During the past five years, the National Weather Service (NWS) has replaced over half of its liquid-in-glass maximum and minimum thermometers in wooden Cotton Region Shelters (CRSs) with thermistor-based Maximum–Minimum Temperature Systems (MMTSs) housed in smaller plastic shelters. Analyses of data from 424 (of the 3300) MMTS stations and 675 CRS stations show that a mean daily minimum temperature change of roughly +0.3°C, a mean daily maximum temperature change of –0.4°C, and a change in average temperature of –0.1°C were introduced as a result of the new instrumentation. The change of –0.7°C in daily temperature range is particularly significant for climate change studies that use this element as an independent variable. Although troublesome for climatologists, there is reason to believe that this change (relative to older records) represents an improvement in absolute accuracy. The bias appears to be rather sharp and well defined. Since the National Climatic Data Center (NCDC) station history database contains records of instrumentation, adjustments for this bias can be readily applied, and we are reasonably confident that the corrections we have developed can be used to produce homogeneous time series of area-average temperature.

1. Introduction

The Cooperative Station Network of the National Weather Service (Fig. 1) was formally recognized as a nationwide federally supported system in 1890, but many of its stations began operation long before that time. Because of its many decades of relatively stable operation, high station density, and high proportion of rural locations, the cooperative network has been recognized as the most definitive source of information on U.S. climate trends for temperature and precipitation. Cooperative stations form the core of the U.S. Historical Climate Network (HCN) (Karl et al. 1990) and the U.S. Reference Climate Network (Quayle 1991).

A significant change in temperature measurement took place in the cooperative network in the mid- and late-1980s (Fig. 2). Because of rather fundamental differences between the liquid-in-glass (LIG) thermometers used in the wooden Cotton Region Shelter (CRS) and the new thermistor Maximum–Minimum Temperature System (MMTS) housed in cylindrical plastic instrument shelters, it was considered entirely possible that systematic biases could be introduced into the record. In addition, the introduction of the MMTS into the network occurred rather quickly, with the total units installed rising from virtually none at the beginning of 1984 to nearly 3000 at the end of 1988, and about 3300 at the end of 1990 (i.e., about 60% of all temperature stations). Century-scale analyses of large-scale temperature often reflect trends on the order of a few tenths of a degree (IPCC 1990) and can be very sensitive to systematic artifices in the data. This makes it extremely important that all systematic and random biases be removed from the station's measurements prior to developing any assessments of, or explanations for, climate change.

We have already expended considerable effort in developing corrections for temperature trends derived from the HCN dataset for a variety of causes. These include systematic biases introduced by changes in observation time (Karl et al. 1986), urban heat islands (Karl et al. 1988), as well as adjustments for random or near-random discontinuities resulting from changes of station location, exposure changes, instrument relocations, etc. (Karl and Williams 1987). When systematic biases are suspected in an observational network, it is most effective to make use of this information when developing a strategy to estimate the magnitude of the bias. Unlike the adjustment procedures developed by Karl et al. (1986) and Karl et al. (1988) for the time of observation bias and the urban heat-island bias, the technique by Karl and Williams (1987) does not make use of a priori information regarding systematic biases. For these reasons we considered it essential, if we were to maintain the integrity of the time series derived from the HCN or other cooperative stations, to collect sufficient data for a thorough investigation of any biases that may have been introduced with the changeover to the MMTS. Whether the new unit was more accurate in an absolute sense was not the issue. Rather, our intent was to investigate quantitative differences between the old and new systems so that biases could be removed prior to time series analyses.

The MMTS was introduced into the cooperative network for several reasons, among them: (1) high-

COOPERATIVE STATION NETWORK
(AS OF JANUARY 1990)

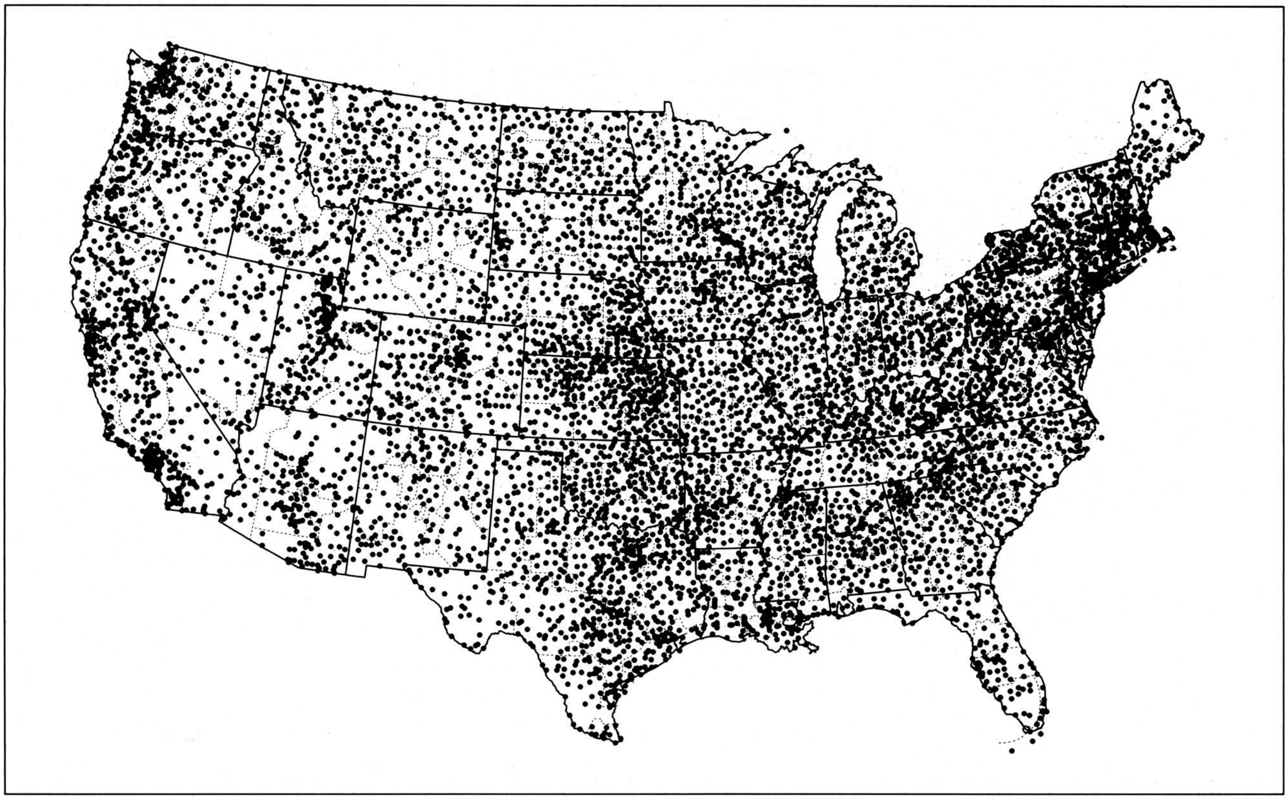

FIG. 1. The Cooperative Station Network of the National Weather Service for the 48 contiguous states.

quality LIG instruments are becoming expensive and difficult to procure; (2) the CRS is large, expensive, and costly to maintain; (3) the convenience of the semiautomatic indoor-reading MMTS makes the ob-

While changes in a climate network are often controversial...we should emphasize that many changes are both inevitable and desirable. We simply need to ensure that a proper scientific infrastructure exists to monitor these changes and develop adjustment factors so that a homogeneous climate record can be maintained.

serving chore easier (this is particularly important for this all-volunteer network); and (4) funding was available. While changes in a climate network are often controversial, as current literature indicates (Robinson 1990; Karl and Quayle 1988), we should emphasize that many changes are both inevitable and desirable. We simply need to ensure that a proper scientific infrastructure exists to monitor these changes and develop adjustment factors so that a homogeneous climate record can be maintained.

2. Experiment design and data analysis

In order to increase the robustness of our analysis, a very large sample of data was analyzed. The experimental strategy was quite straightforward:

(1) We selected two large samples of stations such that one group switched from CRS to MMTS in the 1980s (we call these MMTS stations), and one group remained CRS throughout the test period of 1980–1989 (we call these CRS stations). Stations were limited to those that had no moves or changes in observation time and had adequate instrument-type documentation for the time period in the NCDC station history database. The source of these metadata is NWS B-44 forms, which are prepared by Cooperative Program Managers for each station. The total number of cooperative daily observing stations is about 7750,

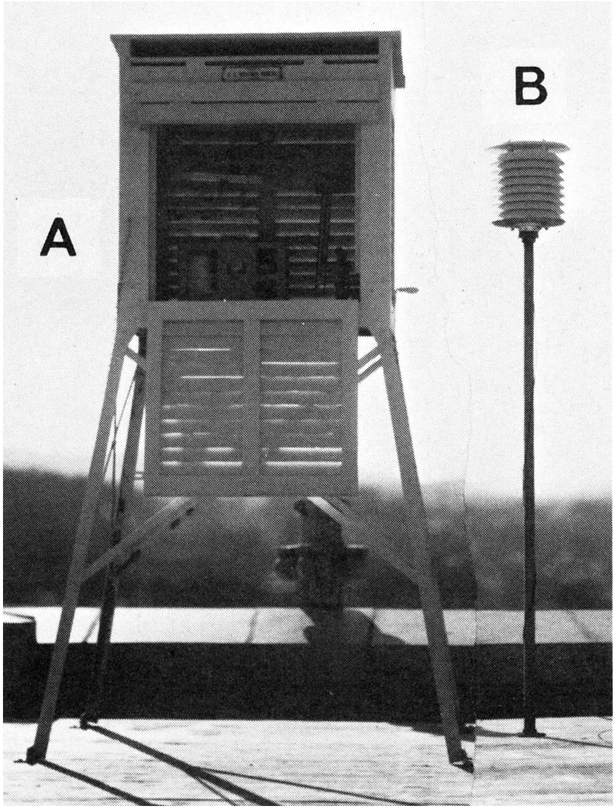

FIG. 2. (a) The standard NWS Cotton Region Shelter (CRS), and (b) the Maximum–Minimum Temperature System (MMTS) photographed at the same scale.

of which about 5300 observe temperature, with about 3300 of those using the MMTS as of the end of 1990. Monthly data summaries from the NCDC TD 3220 files were used in the experiment. This data source includes monthly mean maximum, minimum, and average temperature for stations with less than ten missing days in the month. Average temperature is defined as (maximum + minimum)2^{-1}.

(2) Correlation coefficients (r) of annual mean temperature were calculated for (MMTS, CRS) station pairs for each of the 40 CRS stations nearest each MMTS station during the pre-MMTS era (1960–1979). An MMTS station and its group of five CRS stations were retained for analysis only if the five most highly correlated CRS stations all had an $r > 0.60$. There were 424 MMTS stations selected in this manner, and they are shown in Fig. 3 along with the five CRS counterparts for each MMTS site. A total of 675 CRS stations were selected, many of them being paired with more than one MMTS station. Note that some stations were highly correlated with each other despite a substantial separation (hundreds of kilometers in some instances).

(3) Monthly and annual mean maximum, minimum, and average temperature differences were calculated for the (MMTS, CRS) station groupings in two ways: (a) the most highly correlated CRS station was paired with the MMTS station; and (b) a weighted mean of the five most highly correlated CRS stations (weighted by r) was paired with the MMTS station. It was difficult to assess which one of these methods is more appropriate, so we used them both to help test the sensitivity of our results to our experimental design.

(4) MMTS–CRS differences were calculated for each month of the 1980–1989 period for mean daily maximum, minimum, and average temperature, and for temperature range (maximum–minimum). The difference series for each station was indexed chronologically, with Month 0 defined as the month when the MMTS was installed. The average difference series for each element was calculated by centering each series on Month 0 and computing the average over all available stations for each month before and after the change (−1, +1, −2, +2, etc.). A systematic bias should be reflected in a change in the average temperature difference between CRS and MMTS stations, with this change occurring after the month that the MMTS was installed. Months 0 through 5 were not used in the analysis, since the exact date that the MMTS began full operation was not always precisely known (see section 3 for a more detailed explanation). Furthermore, this change in mean temperature difference could be tested for statistical significance via the Student's t-test.

(5) The differences formed in item (4) above are derived from an aggregate of over 400 stations; however, a question arose regarding what to do when there were incomplete data for stations depicted in Fig. 3. Due to the varying lengths of time the MMTS had been in place at each station, and since cooperative stations occasionally close or have missing data, the number of stations used to compute the average difference series for each element decreased away from the Month 0 (Fig. 4e). We tested the sensitivity of our design to the number of stations in the period prior and subsequent to the MMTS (Month 0) by varying the minimum number of stations that could be used to calculate the mean difference for a given month. Four tests were run with 100, 200, 300, and 400 (MMTS, CRS) station minimums. This was used with both types of MMTS–CRS station groupings listed in item (3) above. As a result, this design produced eight sets of differences for each test of the maximum, minimum, mean, and temperature range (maximum–minimum). This enabled us to produce an estimate of the uncertainty associated with any bias we detected.

3. Results and discussion

Over half of the individual station groupings showed a

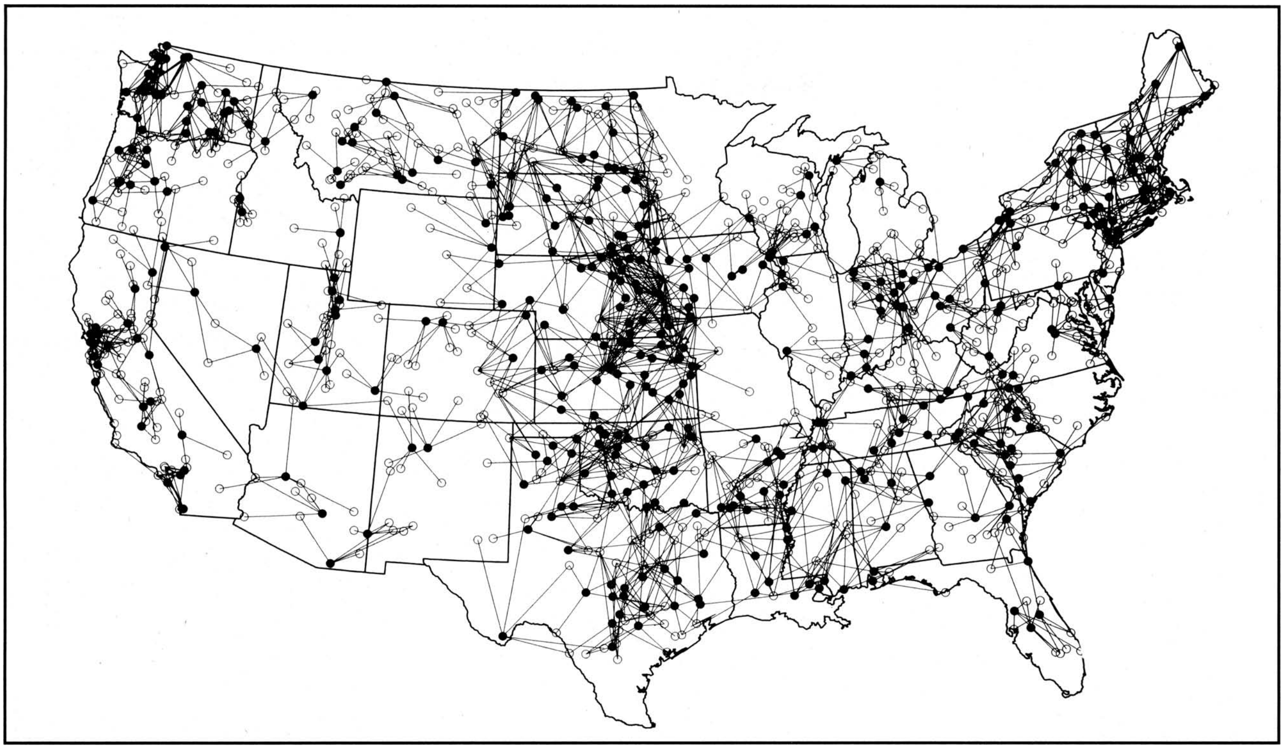

FIG. 3. The 424 MMTS stations selected for this analysis (closed dots) connected by thin straight lines with the 5 most highly correlated CRS stations (open dots). The 675 CRS stations were paired with multiple MMTS sites.

statistically significant change (0.05 significance level) of the (MMTS–CRS) mean difference in maximum and/or minimum temperature when the period before installation of MMTS was compared to the period following installation. The time series of the aggregated mean (MMTS–CRS) differences from both the multi- and single-station neighbors are averaged in Fig. 4. It should be noted from Fig. 4 that Month 0 is not any particular calendar month, but instead is the recorded month of MMTS installation, which occurred gradually throughout the middle and late 1980s. Fig. 4 shows a striking discontinuity for mean daily maximum and minimum temperature, and temperature range, beginning at precisely the time window of the instrument change. Although there is a strong discontinuity at Month 0, the time series did not show a full response to the MMTS until a few months after Month 0. This delay in a full response can be explained by discrepancies in station-history records and continued reliance on the CRS despite MMTS installation. Although we have found that it was not uncommon to dismantle the CRS immediately, many observers apparently continued to use the CRS for a few months after the installation of the MMTS. For this reason we did not consider the first five months after Month 0, despite the fact that our station histories indicated that the MMTS was operational (months 0 through 5 were excluded from the plots). When the data are analyzed in this manner and averaged for the eight different experimental designs, the results indicate that mean daily maxima decreased by 0.40°C, while the minima increased by about 0.28°C. The range decreased by 0.68°C and the average decreased by 0.06°C. Although the differences in average temperature are statistically significant and potentially important to climate-change studies, they may not necessarily be operationally significant for most weather and climate applications.

To investigate the seasonality of the bias, data were analyzed by season (Winter = December, January, February; Spring = March, April, May; etc.). The time series plots (not shown) were very similar to Fig. 4, with the biases shown in Table 1. While seasonality is

TABLE 1. Average temperature change caused by the changeover to MMTS, and the uncertainty (95% confidence limits) based on the experimental design.

	Maximum	Minimum	Average	Range
Winter	–.42±.06°C	+.24±.02°C	–.08±.04°C	–.66±.04°C
Spring	–.34±.02°C	+.23±.06°C	–.05±.02°C	–.57±.04°C
Summer	–.43±.04°C	+.33±.04°C	–.05±.04°C	–.76±.02°C
Fall	–.37±.04°C	+.35±.04°C	–.02±.02°C	–.73±.04°C
All Months	–.40±.02°C	+.28±.02°C	–.06±.02°C	–.68±.02°C

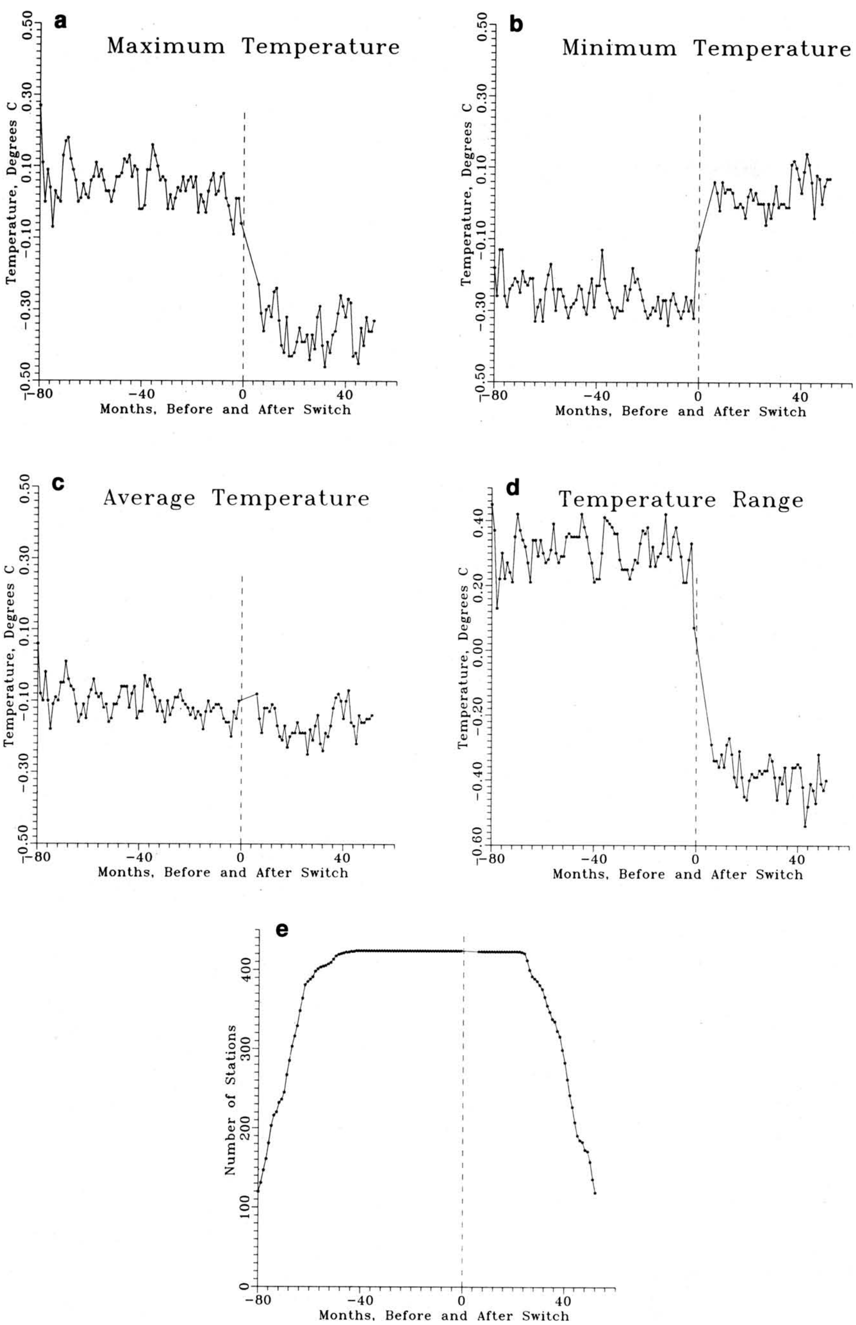

FIG. 4. Averaged time series of aggregated monthly mean (MMTS–CRS) temperature differences for the contiguous United States and total number of stations used to compute average differences for each month. The zero month is the month when the MMTS was installed. Note that months 0 through 5 are not included.

not pronounced, there is a tendency toward larger differences in the maximum in winter and summer, and larger differences in the minimum in summer and fall.

The robustness of the biases we obtained from our analysis was derived from the eight combinations of tests we performed. The standard deviations of the biases were calculated, and used with the standard normal distribution to define the confidence intervals depicted in Table 1. Our results were not overly sensitive to the type of design we chose (e.g., 100 versus 400 stations in the time period, multiple versus single station pairs).

The issue of what caused these apparent biases and which data are more correct (MMTS or CRS?) has not been settled. Blackburn (1991, personal communication) suggests that column separation in the LIG instruments is a likely culprit for at least part of the problem. The maximum LIG thermometer has a constriction in the base of the bore to prevent the mercury from re-entering the bulb when the temperature begins to decrease, thereby recording the highest temperature since the last setting (National Weather Service 1989). It is reset by spinning the instrument such that centrifugal force returns the mercury to the bulb. Column separation in the maximum thermometer (which sometimes occurs near the constriction) causes erroneous high readings. In addition to this problem, unpublished NWS testing has shown that the CRS tends to heat up more during the day than the MMTS shelter (which is consistent with the larger differences in the maximum that occur during the summer), and may admit some reflected sunlight under some conditions. All of these conditions help explain the apparent cooling of maximum daily temperatures that occurs when MMTS units are installed.

The situation is reversed for the minimum LIG thermometers, which use surface tension to drag an index within the alcohol column to the lowest temperature since the previous setting. In this instrument, column separations tend to occur above the index, thus introducing erroneous low readings in the instrument when this condition exists. Another possible cause is the radiation loss from the CRS through its single slatted bottom (the MMTS has a double bottom). These conditions help explain the apparent warming of minimum daily temperatures that occur when MMTS units are installed.

Several other factors have also been put forward to help explain the observed differences. Response times are different for LIG thermometers and thermistors. Shelter and instrument radiation balances probably also vary in different ways with respect to wind, sky cover, and ground cover (including snow). Furthermore, the MMTS units typically may be installed closer to the observer's building than the CRSs to reduce the length of cable that is needed, which may artificially reduce the daily range. Regional differences may also exist in installation, biasing, and maintenance requirements, which could lead to systematic data differences. However, it is beyond the scope of this study to conduct the field tests and research necessary to answer these questions definitively.

Overall, it appears that the MMTS may be a more precise instrument, and that older CRS data should be adjusted when high precision is required. [There is some anecdotal evidence that older (pre-1975) LIG instruments were of higher quality than newer LIG instruments, but there is not yet any quantitative evidence that adjustment factors would be different for older LIG thermometers as opposed to newer ones.] One example of an application requiring the MMTS adjustment is the work of Plantico et al. (1990). They calculate long-term trends for several elements and relate changes in the mean daily temperature range to changes in cloud cover. For that type of research, consistency is more important than knowledge of which data are correct in an absolute sense.

4. Conclusions

Our results indicate that when using maximum or minimum temperatures from the U.S. Cooperative Station Network, it is essential to employ yet another bias correction. Without adjustments to the data, large-scale area averages of mean maximum and minimum temperature could have biases as large as 0.4° and 0.3°C, respectively, and even larger for the mean diurnal temperature range (0.7°C). For the mean temperature, the recommended bias correction is considerably less (about 0.06°C), but this can still be quite significant if one considers the results of Plantico et al. (1990), who report changes of the mean temperature across the United States of a similar magnitude over the twentieth century. The adjustment factors we have derived are fairly robust on a seasonal and annual basis, and should prove useful toward the development of large-scale area-average homogeneous time series. They are most appropriate for use when time series of state or larger area averages are desired, and are not useful for application to daily station data. On any individual day, biases may be significantly different than reported here. The sign of the adjustment will depend upon the standard to which the data refer. The signs in Table 1 adjust CRS data to the MMTS standard.

Acknowledgments. Thanks to Andy Goss, who helped with the early stages of computer programming. Partial support for this work was provided under the DOE/NOAA Interagency Agreement DE-AI05-90ER60952.

References

Intergovernmental Panel on Climate Change, 1990: Observed climate variations and change. *Climatic Change—The IPCC Scientific Assessment,* Cambridge Univ. Press, 201–238.

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.

——, and R. G. Quayle, 1988: Climate change in fact and theory: Are we collecting the facts? *Climate Change,* **13,** 5–17.

——, 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 temperatures 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.

——, C. N. Williams, Jr., and F. T. Quinlan, 1990: United States Historical Climatology Network (HCN) serial temperature and precipitation data. ORNL/CDIAC-30 NDP-019/R1, 374 pp. [Available from Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN.]

National Weather Service, 1989: *NWS Observing Handbook No. 2, Cooperative Station Observations 1st Edition,* Government Printing Office, 83 pp.

Plantico, M. S., T. R. Karl, G. Kukla and J. Gavin, 1990: Is recent climate change across the United States related to rising levels of anthropogenic greenhouse gases? *J. Geophys. Res.,* **95D10,** 16 617–16 637.

Quayle, R. G., 1991: What sort of reference climate station network should the U.S. have? *Proc. of the 15th Climate Diagnostics Workshop.* Asheville, NC, NOAA Climate Analysis Center, NWS, Washington, D.C., 237–240.

Robinson, D. A., 1990: The United States cooperative climate observing systems: Reflections and recommendations. *Bull. Amer. Meteor. Soc.,* **71,** 826–831.