P2.1 QUALITY ASSESSMENT OF MONTHLY TEMPERATURE DATA

David A. Robinson* and John Parlagreco
Office of the New Jersey State Climatologist
Rutgers University, Piscataway, New Jersey

1. INTRODUCTION

The availability of high quality climatological datasets is an important requirement for applied climate studies, as well to address increasing concerns regarding the potential impacts of humans on the climate system (P. Robinson, 1989; D. Robinson, 1990; 1997; Shea et al., 1995; Karl, 1996). Quality assessment of these datasets is a must, as invalid data can lead to erroneous, misleading results. Therefore, it is necessary to screen data prior to them being deemed accurate and complete. Here, we report on a method for evaluating the accuracy of air temperature data gathered at stations within a regional network. The assessment employs monthly means of daily data in order to identify subtle errors that likely may escape quality control procedures run on daily observations. Examples of this method are discussed for several stations in New Jersey.

2. THE NEED FOR MONTHLY ASSESSMENTS

Quality assessment of daily surface air temperature observations may involve internal comparisons of variables gathered at a specific station on a given day, to adjacent days, or to a station climatology. A comparison of an observation to others gathered simultaneously at nearby locations may also take place. This daily assessment is the type of quality control most often done at climate data centers (Reek et al, 1992; Peterson, 1993). However, daily evaluations of surface air temperature observations are unable to recognize subtle, sometimes not so subtle, errors in reported temperatures. This may be a result of having to accept a fair degree of daily variance between a station observation and others in the region, or even to past observations at that station, thereby avoiding the possibility of flagging accurate observations as erroneous, "correcting" accurate observations, or even rejecting them from an archival database.

We have found that some temperature errors are best identified when examining monthly means. By comparing station means of maximum and minimum temperatures to those from nearby, climatologically similar sites, errors appear in one to several months. If followed by rapid remedial action by individuals responsible for the maintenance of an observing network, the need for the identification and adjustment of data at a later date is obviated and the employment of unrecognized erroneous data in studies is avoided.

3. SCREENING METHODOLOGY

A Visual Basic program has been created to assess the quality of station temperature observations at National Weather Service first-order and Cooperative Observing stations in New Jersey. The program is straightforward and utilizes basic statistics. First, a forty-year (1961-2000) database of monthly mean maximum and minimum temperatures was created for stations with records of that length. Months with more than 5 days of missing data were excluded from the analysis, as were stations with an excessive number of months missing during the 40 years. Stations with known moves or changes in observation time, observer or instruments were not excluded from the study. To do so would eliminate from this evaluation virtually every station in New Jersey with the requisite length of record. Rather, we felt that influences associated with the incorporation of any of these changes would be minimized by the use of the 40 years of record. The only exception to this will be when a station change occurred close to or following the end of the mean base period. Thus base period values may not be appropriate to use when evaluating observations at the new site.

Once the monthly database was established, stations were grouped into 4 climatic regions. Differences in mean maximum and minimum

*Corresponding author address: David A. Robinson, Office of the State Climatologist, Rutgers University, 54 Joyce Kilmer Ave., Piscataway, NJ 08854; e-mail: drobins@rci.rutgers.edu

temperatures were calculated for every combination of stations within a region for each month where data were available for both paired stations. Several different measures of dispersion were examined, and given that the distribution of the differences around the mean was fairly normal, standard deviations (SD) were selected as a simple and appropriate statistic for the station evaluation process. Thus, SDs of differences were next calculated for all station pairs using up to 40 years of differences for a given month and variable.

With monthly standard deviations of differences established for all pair combinations within a region, the next decision to be made was how to use this information when evaluating mean temperatures within the region for a given month. Through trial and error, we found that using a threshold of 1.5 SD permitted the identification of suspect stations. Since there is a 13.4% chance in any month that differences will fall outside of the 1.5 SD envelope, with neither station in the pair exhibiting a problem, it is necessary to evaluate each station with respect to every other station in the region. Should a pattern emerge where the difference between a station and most others in the region exceeds 1.5 SD for a period of several months, then you can be rather certain that it is that particular station that has a problem. Just how many stations must exceed the 1.5 SD threshold and for how many months this must occur in order for some remedial attention to be paid to a station will depend on the magnitude of the problem. Thus there is a degree of subjectivity in evaluating the results. However, this is not time consuming and the time is well spent if long-term subtle station errors can be avoided by operational scrutiny of the temperature means.

4. ANALYSES

To exemplify the utility of our approach, graphs have been generated that depict results between a given station (home) and four others (neighbors) in a climatological region. The number of cases in which the difference between the monthly mean at the home station versus neighbors exceeds the long-term mean difference by more than 1.5 SD are plotted for three-year intervals (figures 1-4). Results for both maximum and minimum means are shown. For instance, on only three occasions from 1991 through 1993, did the maximum temperature difference between the Long Valley, NJ station and a neighbor exceed the 1.5 SD threshold (October 1992, August and November 1993) (figure 1). During this interval the Long Valley maximum observations were stable with respect to their neighbors. This was not the case for minimum observations from fall 1991 through summer 1992, when the threshold was exceeded for two to four neighbors in each month but two. Early in fall 1992, Long Valley was visited by a NWS technician, who found that the liquid in glass minimum thermometer was faulty. Once it was replaced, the figure illustrates the elimination of threshold exceedences.

Another problem station is shown in figure 2. Lambertville, NJ minimum temperature observations were well out of line from their long-term relationship with neighboring stations from fall 1998 through 2000, especially in 1999. An evaluation of the daily observations from this site identified some gross errors that were inconsistent with time. Apparently, one of the several observers at this station, operated by the Lambertville Sewerage Authority, did not know how to read the minimum liquid in glass thermometer properly. NWS personnel were contacted and the hope was that this problem would be remedied. However, subsequent suspect months in 2000 show that it did not promptly disappear. Unfortunately, the problem continues to emerge periodically at Lambertville.

One of the neighbors to which Lambertville is compared is the New Brunswick, NJ Cooperative station. Figure 3 shows the 1998-2000 time series of threshold exceedences for New Brunswick. Figure 4 shows the same, except that Lambertville is removed from the evaluation, leaving three neighbor stations for comparison. It is apparent that most of the threshold exceedences in figure 3 are the result of the errors at the Lambertville station. Without Lambertville, only the occasional exceedence is seen for New Brunswick. Thus we are confident that the New Brunswick station is not having any problems, and we can begin to see that the other three stations in the region are likely in good shape. However, the latter should not be assumed until analyses are conducted with each station as the "home" site.

5. CONCLUSION

Within the Office of the New Jersey State Climatologist (ONJSC), we have begun to operationally implement this threshold approach for identifying suspicious station observations to all NWS monthly data. While quantitative in nature, there is a degree of subjectivity in deciding when a station should be deemed suspect. Once such a determination is made, we will report our suspicions to personnel at the NWS Forecast Offices responsible for the maintenance of stations and the training of observers.

For NWS stations without the requisite length of record for establishing standard deviations of differences with neighboring stations, we have selected a surrogate station and use the SDs established between the surrogate and other long-term stations to check the new station for inconsistencies. We will be taking a similar approach for all stations within the New Jersey Weather and Climate Network (NJ WxNet) (http://climate.rutgers.edu/stateclim/index.html). The NJ WxNet is a real-time network under development within the ONJSC. It will eventually include hourly data from over 100 stations, gleaned from a quilt work of smaller electronic networks across the Garden State.

Accurate climatic data is an extremely valuable resource. Considerable efforts are necessary and justified to assure that information of the highest quality is collected and made readily available to the ever-growing user community. This begins by having accurate and properly situated instruments, well-educated observers, and diligent management of observation locations. This must be followed by careful scrutiny of the data gathered, and documentation of station particulars, including known problems with the data. We are confident that our efforts described within this paper, as well as those of others within the NWS and elsewhere will result in the availability of climatological information of the highest quality for applied studies.

Acknowledgments: Thanks to Johnnie Woo and Brian Cosgrove for preliminary efforts toward developing this methodology. The Office of the NJ State Climatologist is situated within the Center for Environmental Prediction, and is partly supported by the New Jersey Agricultural Experiment Station, Rutgers University.

REFERENCES

Karl, T.R. (ed.), 1995: *Long-Term Climate Monitoring by the Global Climate Observing System*. Kluwer Academic Publishers, Dordrecht, 518pp.

Robinson, D.A., 1990: The United States cooperative climate-observing systems: reflections and recommendations. *Bull. Am. Meteor. Soc*., **71**, 826-831.

Robinson, D.A., 1997: Gathering climatic data of the highest quality. *Proceedings of the 10th Conference on Applied Climatology*, Reno, NV, American Meteorological Society, 305-308.

Robinson, P.J., 1989: Climate Information Needs in the Southeast United States. *Final report: NOAA Grant No. NA88AA-D-CP041 to the Southeast Regional Climate Center*, 36pp.

Shea, D., R. Jenne, and W. Spangler, 1995: Daily and monthly maximum and minimum temperature datasets archived at NCAR and some applications. *Atmospheric Research*, **37**, 11-17.

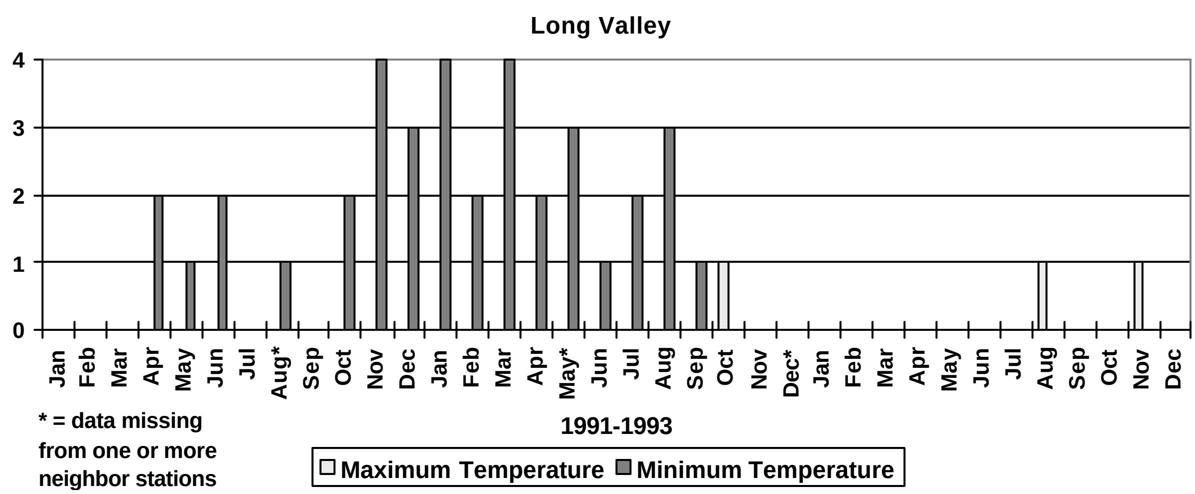

Figure 1. Time series of cumulative threshold exceedences for Long Valley compared to four neighbor stations. See text for additional details.

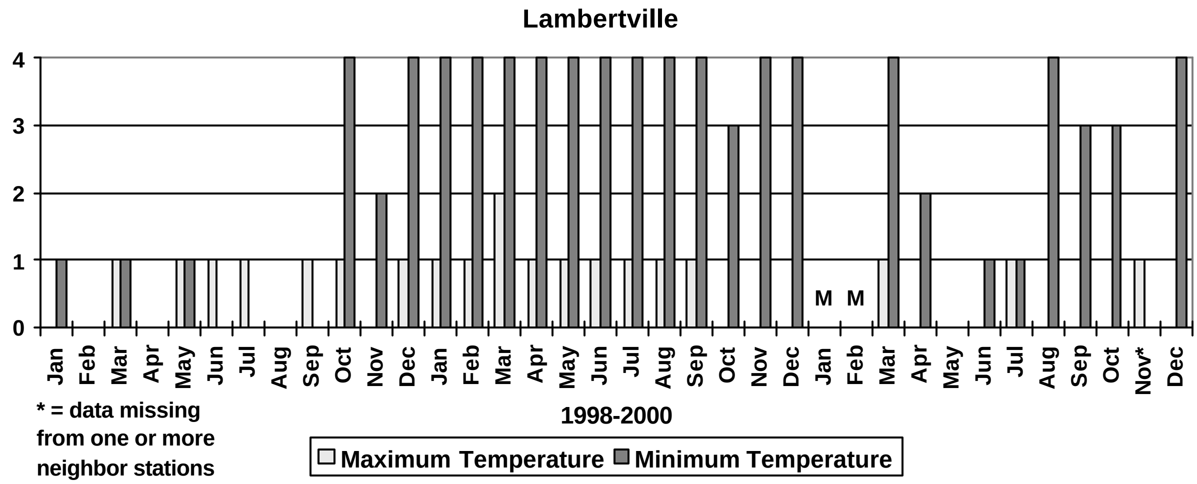

Figure 2. Same as figure 1 except for Lambertville. Temperature data are missing from Lambertville for the period January 2000-February 2000 (denoted by M above).

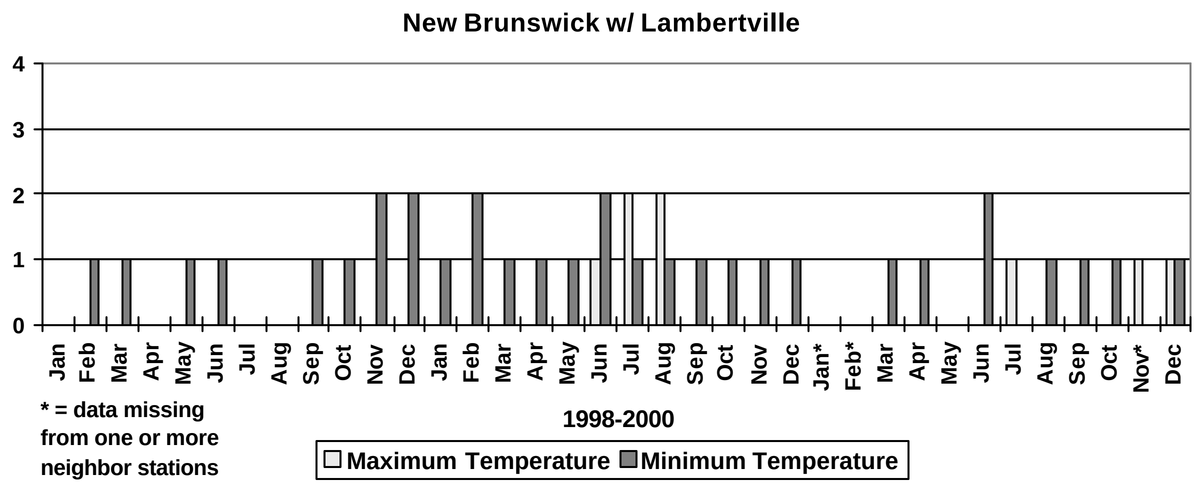

Figure 3. Same as figure 1 except for New Brunswick.

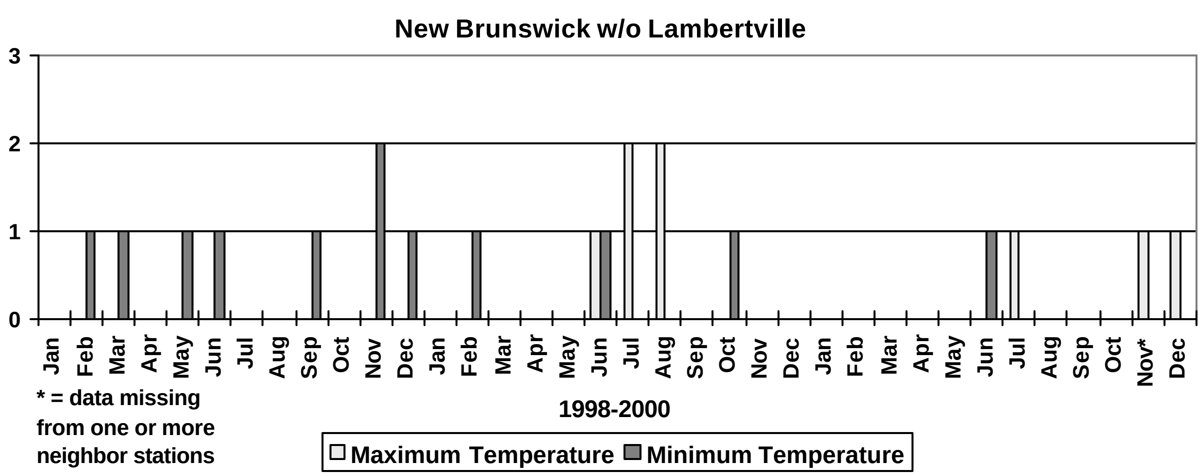

Figure 4. Same as figure 1 except for New Brunswick without Lambertville.