The United States Cooperative Climate-Observing Systems: Reflections and Recommendations

David A. Robinson
Department of Geography
Rutgers University
New Brunswick, NJ 08903

Abstract

The system for the collection and archiving of climatic data from approximately 7000 cooperative observing stations across the United States is in need of improvement. Despite the efforts of many dedicated volunteers and professionals, suspect or incomplete data continue to enter the national climate archive. Cooperative observers need further education regarding the importance of collecting complete and accurate data. The transition to the Maximum/Minimum Temperature Systems (MMTS) from the former liquid-in-glass thermometers mounted in Cotton Region shelters needs to be better coordinated, particularly with respect to the continuity of temperature data. The role of the Cooperative Program Managers in overseeing the physical well being of the network and the quality of data emanating from its needs to be strengthened. A continuation of efforts recently begun at the National Climatic Data Center (NCDC) to improve the digitization and quality-control procedures employed in data archiving is required. These improvements might be fulfilled within the present National Oceanic and Atmospheric Administration (NOAA) infrastructure. Another avenue which might be explored is the amalgamation of all aspects of the United States climate-observing system under the supervision of a new department within NOAA's National Environmental Satellite, Data, and Information Service (NESDIS). Regardless of how this is achieved, now is the time to improve the system. There has never been a greater need for a national climatic database of superlative quality, whether it be for investigations of climate change, meteorological research, agricultural planning and assessment, engineering, environmental-impact assessment, utilities planning, or litigation.

1. Introduction

Along with several hundred National Weather Service (NWS) and Federal Aviation Administration (FAA) primary stations, a network of some 7000 volunteer cooperative observers collect daily climatic data for such variables as temperature and precipitation. Equipment to gather these data is provided and maintained by NWS personnel and data forms are sent monthly to the National Climatic Data Center (NCDC) in Asheville, North Carolina, where data are digitized, quality controlled, and subsequently archived. The NWS and NCDC are both part of the National Oceanic and Atmospheric Administration (NOAA). NCDC is part of NOAA's National Environmental Satellite, and Data Information Service (NESDIS).

These data are invaluable in learning more about the floods, droughts, heat and cold waves, etc., which inevitably affect everyone. They are also employed in agricultural planning and assessment, engineering, environmental-impact assessment, utilities planning, and litigation and play a critical role in efforts to recognize and evaluate the extent of human impacts on climate from local to global scales. Unfortunately, just when the need for a dense network of high-quality climatic data is greater than ever in the United States, the national system to acquire and archive these data is at its weakest point in years. One indicator of this is the significant decrease in the number of station-years of data gathered by observers (figure 1). Since the mid-1960s, decreases have totalled approximately 1800 precipitation (20%) and 700 (13%) temperature station-years.

A number of changes are required to improve the cooperative climate-observing system. This situation has recently been addressed in part by Redmond (1986) and Karl and Quayle (1988). Following is a discussion of some of the major problems with the current system, accompanied by recommendations on how they might be remedied. This essay evolved from a year-long sabbatical at the National Climatic Data Center.

2. Today's United States cooperative climate-observing system

Thousands of volunteers and hundreds of professionals are striving to make the United States cooperative climate-observing system a successful one. They are succeeding to a considerable extent; however, the system currently has various shortcomings that should and can be corrected. The need to have climatic data of a high quality has received increasing attention in recent years from scientists, United States federal agencies (i.e., Interagency Working Group on Data Management for Global Change) and international committees (i.e., International Geosphere/Biosphere Program Data Management Working

Group). However, concurrently, there have been major cuts in staffing at NCDC and an apparent deemphasis in the proper management and maintenance of cooperative observing stations.

The four major components of the cooperative climate observing system are 1) the observers; 2) the observing equipment; 3) Cooperative Program Managers (CPMs); and 4) the digitization, quality control, and archiving of observations at NCDC. Each are addressed individually below. Recommendations focusing on each of the components are made independent of the others.

a. Cooperative observers

One cannot fully appreciate the dedication of cooperative observers without personally making the effort to collect climatic data on a daily basis. The majority are doing a superlative job; however, there are numerous observers who fail to collect accurate and complete data. This is indicated by the absence of reports from as many as several hundred stations in any given month.

Among the data submitted, reports frequently contain missing days. Most disturbing are observing-forms containing incorrect observations. Subtle errors, perhaps due to a faulty instrument, may be beyond the control of the observer, and an occasional slip of the pen is only human. However, there are cases where temperature errors of some 5° or 10°C are noted day after day, or situations where impossible combinations of maximums and minimums are noted (i.e., one day's minimum exceeding the previous day's maximum). A recent evaluation of the quality of the snow data gathered by cooperative observers indicates that almost half fail to adequately record snowfall and/or snow on the ground (Robinson 1989). Forms submitted to NCDC with missing or incorrect station identification are another problem. Errors are costly in terms of the amount of time needed to correct them and with regard to the reliability of the archived data should they pass through the quality-control process.

How can we improve the situation?

1) *Observer education.* Emphasize correct data-acquisition methodology and instill observers with a heightened awareness of the necessity of gathering accurate and complete climatic data. Show observers where their data are being applied (i.e., for exploring the impact of humans on climate, investigating a flood, or for litigation). This can be accomplished through CPMs and the *National Cooperative Observer Newsletter.* Employ a clipping service to extract cases where cooperative data are being used and reported in the press and publish this information in the newsletter. Publish examples of research utilizing the data.

2) *Give credit where credit is due.* Institute a program that publicly recognizes superior observers for their efforts. This activity should be coordinated by the CPMs with input provided by NCDC, and accomplished through the local media, the Association of American Weather Observers newspaper, *Weatherwise,* etc. A variation of this idea is the current United States Centennial Cooperative Weather Station Program, which is designed to recognize observing stations with at least 100 years of virtually continuous observations.

3) *Eliminate those stations consistently collecting inferior data.* Quite often, no information is better than mediocre information. Formulate criteria for station elimination, and apply this on a nationwide basis.

b. Cooperative station equipment

Station relocations or changes in the physical environment surrounding stations are factors a climatologist must continually be concerned with when analyzing data. Until recently, one did not have to be too concerned with equipment changes influencing records over the past century. Liquid-in-glass maximum and minimum thermometers mounted in Cotton Region shelters (CRS, also known as Stevenson Screens) were the standard. The thermometers were not flawless; quality varied and they occasionally broke or drifted from their original calibration, but, overall, consistency reigned. This may have terminated some 6 years ago, with the introduction of the thermistor-based Maximum Minimum Temperature Systems (MMTS) mounted in small "beehivelike" structures. Presently, approximately half of the cooperative stations that observe temperature are

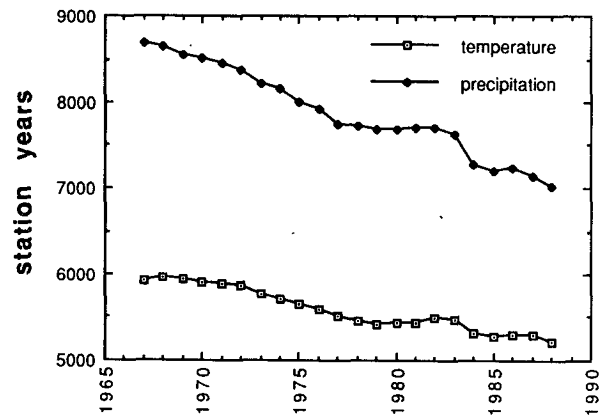

FIG. 1. The annual number of station-years of temperature and precipitation data from United States cooperative stations since 1967. The total number of monthly forms received at NCDC in a given year was divided by 12 to obtain these values. The number of station-years is approximately 5% lower than the actual number of stations, due to the omission of missing forms when calculating this value (data courtesy of NCDC).

operative stations that observe temperature are equipped with the MMTS (figure 2), and more will follow over the next several years.

Temporal continuity need not be sacrificed when introducing a new technology, but the climate-research community is concerned that this continuity may have been lost with the MMTS. Several informal studies have found differences between MMTS and liquid-in-glass/CRS readings (i.e., Baker and Ruschy 1989). It remains uncertain as to whether the dissimilarities are a function of instrument responses under various weather conditions (probably due to the different microenvironments within the shelters) or are random in nature. It is not too late to address this problem in a structured manner. A nationwide program must be established where approximately 100 stations, in an assortment of environments, simultaneously record MMTS and CRS/liquid-in-glass data for a period of years. A detailed analysis would follow and the results would be published. Coincident sky, wind, and moisture data from these or nearby stations should be incorporated into the analysis. The primary issue is not which temperature system is accurate or optimum, but the maintenance of comparable records (at least in terms of instruments and shelters) over decades-long intervals.

Apart from concerns over continuity, there have been problems with the durability of the MMTS sensors. Within the past 6 years there have been as many reported MMTS malfunctions as there are units installed (personal communication, G. Goodge). These malfunctions result from lightning strikes, burrowing animals cutting a cable, insects nesting in a shelter, etc. They often require a repair visit from a CPM and at times result in a backup unit (if available) needing to be installed and the original sent to the National Reconditioning Center in Kansas City for repair. Another problem with the MMTS results from their limited backup battery supply. When electrical power fails for more than 4 hours, the unit ceases to record and previous data are lost. At times when the temperature may be of great interest, for instance during an ice storm, this quite likely could lead to serious data gaps. Another example of this problem resulted from Hurricane Hugo in September 1989. Some 20 MMTS-equipped stations in South and North Carolina were without power for a number of days following the storm; thus, all temperature information was unavailable.

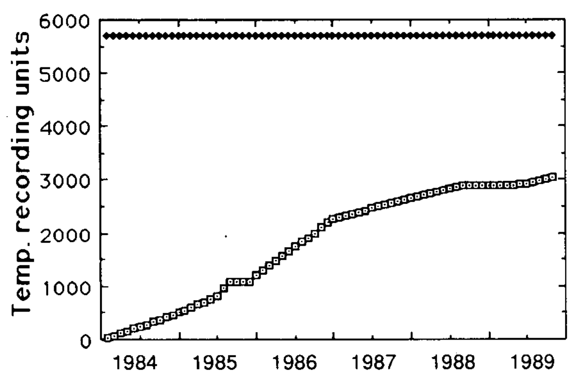

FIG. 2. The number of Maximum Minimum Temperature Systems (MMTS) installed at cooperative stations in the United States and the approximate number of stations recording temperature data (data courtesy of NCDC).

A photograph in the *National Cooperative Observer Newsletter* (1988) shows the MMTS shelter at the Homer 8NW, Alaska station to be situated approximately 2 m from the observer's home. The photo also shows a CRS, where it is assumed data were previously recorded, sitting some 20 m from the house. While the trenching and wiring required to install an MMTS is time-consuming, it is unfortunate that effort was not made to locate the shelter exactly where the CRS was situated. How many other stations has this occurred at and to what extent might such situations affect temperature data?

With the omission of a clock and a small amount of memory in the construction of the MMTS system, an excellent opportunity has been lost to standardize the time of temperature observations at all cooperative stations and to permit the occasional absence of an observer (only with regard to temperature readings). Also, at key cooperative stations, such as those in unique environments or distant from NWS or FAA stations, hourly temperature readings could be retrieved, stored, and later forwarded to NCDC on floppy disk. Some of these stations could also be connected by phone line to forecast offices. It is strongly recommended that these enhancements be implemented as soon as possible. These moves will suffice until an affordable technology is available to fully automate the data-gathering effort (whether the latter is a wise decision is open to debate).

NOAA must not be satisfied with the fact that the quantity of data missing from MMTS stations is no different (and at times worse) than the quantity missing at the liquid-in-glass/CRS stations. *The MMTS performance must be improved.*

c. *Cooperative Program Managers (CPMs)*

CPMs are responsible for maintaining the network of observer stations. This includes installing and maintaining equipment, educating observers, keeping accurate station histories, finding new observers, and eliminating poor ones. Presently, the approximately 50 CPMs are NWS personnel. Unfortunately, being a part of this operational branch of NOAA leads to

CPM duties not always being given highest priority. It is apparent that the CPM system is in need of reorganization and improvement. The following steps are recommended:

1) Assign CPMs the task of reviewing all observation forms from stations in their network prior to forwarding them to NCDC. The CPM should be keenly aware of the recent weather and overall climate of his or her region, knowledge that can be employed in the identification of suspicious data. Information regarding the latter should be forwarded with all forms to NCDC in a timely manner. This approach will also keep the CPM apprised as to the completeness of station observations and alert the manager to any undesignated changes in the time of observation (e.g., shifts from morning to afternoon observations). This, with subsequent feedback to the observers, will ultimately result in fewer data gaps, a more efficient quality-control process at NCDC, and a more expeditious elimination (and replacement) of poor observers. While presently implemented in some regions, the effort is hampered by the operational duties and excessive travel needs of CPMs.

2) Make CPMs more accountable regarding the accuracy and completeness of data emanating from stations under their domain. Detailed analyses of cooperative station performances (by CPM region) should be employed for purposes of evaluation.

3) Encourage CPMs to contact the numerous amateur and professional organizations of serious weather "buffs" around the nation when searching for observers to relieve those retiring, replace poor personnel, or fill gaps in the network. Examples of such groups include the American Meteorological Society (AMS), the National Weather Association, the Association of American Weather Observers, and local organizations, such as AMS chapters, the North Jersey Weather Observers, and television-observer networks.

4) Increase the number of CPMs by at least ten nationwide. This will permit the fulfillment of their stated duties, reduce their travel burden (a particular problem in the western United States), and allow for the lengthy process of trenching and wiring when installing MMTS units.

5) Consider removing the CPMs from NWS management and placing them under the jurisdiction of NCDC or within a new NOAA/NESDIS department responsible for coordinating the collection and archiving of climatic data (more on this in section 4).

d. National Climatic Data Center (NCDC)

A three-step effort must be undertaken at NCDC to improve the quality of data being archived. This includes a manual quality control (QC) of daily data as it arrives at the center, double keying all data, and a final interactive QC of the keyed data. Recent progress has been made on all fronts and will be discussed below.

1) Manual QC

More time must be spent with the cooperative data forms prior to their digitization. This will facilitate the identification of problems with the daily data which might not be picked up by the automated QC. At present, such difficulties include incorrect station numbers on forms, persistent errors in temperature and false zeros in precipitation, snowfall, and snow-on-ground data. To best accomplish this task, the personnel must be familiar with the station network they are quality controlling, including the characteristics of the climate within that region. Progress in this direction has recently been made at NCDC, with the introduction of function-specific assignments. This will eventually result in QC experts consistently reviewing forms from the same stations on a month-to-month basis.

2) Double keying

Double keying of all cooperative data has recently begun at NCDC. Earlier tests indicated that double keying eliminates approximately 98% of the keying errors which previously made it through the process, and is cost effective (personal communication, R. Quayle). This process greatly reduces the amount of hands-on quality controlling of keyed data, locates errors which would otherwise pass through the QC, and permits the QC emphasis to be placed on those data that are suspicious upon arrival at NCDC.

3) Interactive QC

The most effective means of quality controlling keyed daily data is to pass the data through an automated QC, followed by an interactive QC employing trained NCDC personnel and computer-generated spatial and temporal information.

The inclusion of a computerized QC into the overall process occurred in 1982. The computer program identifies suspicious data and, in the case of temperature, adjusts these values or fills in data gaps using an empirically derived algorithm. Both original keyed data and estimated edit values are retained in the database. Enough time has transpired to illustrate the benefits of this procedure and illuminate its limitations. Presently, a major effort is needed to fully evaluate and improve the automated QC. The appraisal must commence by examining and restructuring the lengthy QC computer program. This will permit a full understanding of just how certain QC steps are managed (something lacking at present) and allow for procedural improvements. Simultaneously, input from

CPMs, state climatologists, and regional climate center directors should be solicited. These people are most cognizant of the automated QC's present shortcomings and are well-qualified to make suggestions regarding its enhancement.

The problems to be remedied include the plethora of false zeros making it through the routine and into the digital file, and the frequent elimination and replacement of accurate temperature data with incorrect values. The latter is a common problem in the western United States, where the varied topography and reduced spatial density of stations exacerbate imperfections within the present processing system.

An example of an addition to the automated QC program that will help solve one aspect of the false zero problem has to do with snowfall. A new subroutine is needed to identify (flag) any day with measurable precipitation and a maximum temperature of less than −1°C, yet with no snowfall reported. Experience applying this to a sample dataset shows that while some freezing rain events are flagged, they usually can be identified from supportive data (i.e., many other stations in the region similarly flagging events or notes on the cooperative observer's form) and left unchanged. All other flagged events are considered incorrect and the snowfall is digitized as missing, unless it is determined that it is the precipitation report that is incorrect, in which case this value is replaced with a zero.

All daily data flagged as suspicious in the automated QC should be scrutinized by trained personnel. An educated judgement of its validity can best be achieved by charting a suspicious value on a map including data from surrounding stations. Such a procedure is currently in the developmental stage at NCDC. Known as the Geographic Edit and Analysis System (GEAS), it shows great promise as an effective evaluation methodology. Presently, the major obstacle to bringing the system on-line is uncertainty regarding station observation times. As common observation times are required for useful station intercomparisons, erroneous or missing information stymies this type of QC effort. NCDC must work with the CPMs and observers to assure that consistent and well-documented times are maintained.

In addition to daily data, totals and means (monthly, seasonal, and annual values) of climatic variables should be evaluated to facilitate the identification of subtle and persistent errors in the data. This undertaking could utilize the GEAS. It is not recommended that the outcome of this portion of the QC effort be the correction of suspicious data, rather that 1) those data be adequately flagged as suspicious or eliminated from the digital file, and 2) the CPM in charge of the station in question be notified that an observation problem apparently exists.

It bears repeating that the recommendations made in this and other sections are independent of any changes in other components of the cooperative observing system.

3. Discussion

Presently, at least several percent of the data being collected by observers and later archived are missing or inaccurate (precise numbers are not maintained at NCDC). This performance requires improvement for a number of reasons. Illustrations follow that fall under three broad categories: 1) case or episode specific, 2) long-term variations or trends, and 3) extreme value analyses.

1) A miskeying of one precipitation observation during the summer of 1988 almost cost a farmer a \$70,000 drought-insurance claim. A rainfall of 0.07" was keyed as 0.17", putting the seasonal total above the threshold for collecting on the policy. Only when the records were rechecked was the error noticed. In other cases, missing information can impede investigations or litigation involving traffic accidents or claims regarding construction contracts with weather-dependent completion clauses.

2) Investigations of long-term changes or trends in temperature, precipitation, or snow may be seriously weakened or flawed without proper station histories, including observation times, locations, and measurement systems.

3) Extreme value analyses are perhaps more affected by inaccuracies and missing data than any other studies. Their outcome influences decisions regarding flood protection, utility capacities, engineering studies with regard to snow loads, etc.

Finally, with the substantial reduction of stations in the cooperative observing network in recent years, verification of suspicious data or substitution of missing data at a given station with data from a neighboring station is presently less likely to be accurate than in the past.

4. Conclusions

Climatic data of a relatively high quality continue to be collected and archived from thousands of cooperative stations in the United States. However, more accurate and complete data could be realized. Here, recommendations have been made touching on all aspects of the program, including observer education, comparability and quality of station equipment, the duties of CPMs and the digitization and quality control of data at NCDC.

These recommendations might be fulfilled within the present NOAA infrastructure. However, another avenue that might be explored is the amalgamation of all aspects of the United States climate-observing system under the supervision of a new department within NOAA/NESDIS. This unit would coordinate and supervise all aspects of the observation and archiving of climatic data. This includes all facets of the cooperative observers program, as well as the climatic aspects of the primary observing program. The latter includes station instrumentation and maintenance, and the end-of-month quality control and archiving of data gathered at NWS and FAA stations. Regardless of how the situation is addressed, it is crucial that all possible efforts be made to assure the integrity of the United States cooperative-climate observing network.

Ackowledgments. Thanks to T. Blackburn, G. Goodge, T. Karl, and R. Quayle for their helpful comments. The author's visiting scientist appointment at NCDC was supported by National Science Foundation Grants ATM 87-19865 and ATM 89-96113 and the National Oceanic and Atmospheric Administration.

References

Baker, D. G., and D. L. Ruschy. 1989. Temperature measurements compared. *The State Climatologist* **13:** 2–5.

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

National Cooperative Observer Newsletter. 1988. **10**(5): 4.

Redmond, K. 1986. Thoughts on improvements in cooperative data. *The State Climatologist* 10: 28–36.

Robinson, D. A. 1989. Evaluation of the collection, archiving and publication of daily snow data in the United States. *Physical Geography* **10:** 120–130.