6A.9 INTERACTIVE QUALITY ASSURANCE PRACTICES

Karsten A. Shein*
NOAA National Climatic Data Center, Asheville, NC 28801.

1. INTRODUCTION

Historically, most post-observation quality assurance (QA) of federally-collected climate data has been administered by NOAA's National Climatic Data Center (NCDC). These efforts have achieved great success – appropriately identifying the validity of over 95% of in-situ data ingested by NCDC. However, growing interest from the atmospheric science community in ensuring the utmost quality of the historical climate record has led NCDC to both develop and participate in a number of interactive QA practices. These practices include the development of the Health of the Network (HoN) tool and the use of the Datzilla error reporting and tracking tool. These tools and their use ensure that stakeholders such as NOAA field personnel, regional climate centers and state climatologists are able to more readily detect observational irregularities, communicate more effectively with NCDC regarding possible errors they identify in the climate data archive, and track quality flags and data value estimates set by NCDC's quality control (QC) processes.

In particular, a great deal of QC is devoted to evaluating the validity of meteorological observations reported to NCDC from the NOAA Cooperative (COOP) volunteer weather observation program. The COOP network is the longest-lived meteorological network in the United States, and has provided daily observations from over 12,000 stations throughout its history (c.f., Root, 1930, Robinson, 1990). The importance of maintaining the highest quality of data from this network for use in long-term climate analysis cannot be understated.

In addition to a number of automated and quasi-interactive QC processes run internally on COOP data, NCDC also utilizes two fully interactive QC tools to further ensure the quality of the data as well as the performance of NCDC's QC processes. These tools are the Health of the Network tool and the Datzilla error reporting and tracking tool, and they facilitate the active participation of data managers and other trained personnel from outside NCDC in NCDC's QA process.

2. QA AND QC OF COOP DATA AT NCDC

A keystone component to NCDC's mission is "… to provide access and stewardship to the Nation's resource of global climate and weather related data and information …." As a steward of these data, NCDC actively develops and applies a comprehensive suite of tools that evaluate the quality of climate data received by NCDC from the several NOAA-supervised weather observation networks and other sources (e.g., FAA, DOD). Such quality control ensures that observations from these weather stations are thoroughly checked for logical inconsistencies (e.g., minimum temperature exceeding maximum temperature) and unexpected incongruities (e.g., outliers, spikes). The entire array of these QC tools applied to the data is the backbone of NCDC's quality assurance process.

There exist a number of reasons why an observation might be in error. Errors can be made at the time of observation (e.g., sensor read incorrectly), at the time the observation is transcribed (e.g., transposition of digits, date shifting), and when the data is transmitted (e.g., computer errors, digitizing errors). Each of these errors must be recognizable by the QA process and include objective solutions that may be applied. While many of these errors will eventually be prevented by NOAA's transition toward a paperless observation reporting environment for its weather networks (c.f., Shein and Owen, 2007), many will continue to occur and must be identified and corrected by NCDC's operations.

The variety of QC checks applied by NCDC to NOAA in-situ data is discussed extensively in numerous reports and articles (e.g., Kunkel et al., 2005, Eischeid et al., 1995, Reek et al., 1992, Guttman and Quayle, 1990), and include among others:

- Logical impossibilities
- Temporal spikes
- Flatliners
- Date shifting of observations
- Outliers
- Excessive range
- Spatial anomalies
- Change point detection

NCDC's QC checks are necessarily conservative, and will not flag a value as being definitively erroneous unless it is logically or meteorologically impossible. When an unmistakably erroneous observation is encountered by automated QC checks, it is flagged as invalid. If a unique correction can be identified by an objective method, a subsequent observational estimate is derived to accompany (not overwrite) the original (invalid) value. Given the large quantity of in-situ data received by NCDC each month, most of these QC routines are automated. Only when a value is identified by an automated QC check as being unlikely, but not impossible, will it be set aside for manual review.

* *Corresponding author address:* Dr. Karsten A. Shein, NOAA National Climatic Data Center, 151 Patton Ave., Asheville, NC 28801; E-mail: Karsten.Shein@noaa.gov.

Contrary to claims that excessive incorrect flagging is occurring (e.g., Hubbard et al., 2007), an overwhelming majority of observations received by NCDC are appropriately flagged by NCDC QC. To provide example, in 2006, NCDC received over 10 million daily temperature (maximum, minimum, and 'at observation') and precipitation (incl. snow and snow depth) observations from the COOP network. All of these elements were subjected to NCDC's QA process, and 96.7% passed all QC checks as being within acceptable tolerances and logical consistency. The remaining 331,777 values (3.3%) failed one or more of the automated QC checks and were flagged as invalid. NCDC's automated QC was able to provide unique, plausible estimates for 58.9% of those invalid observations.

Secondary, interactive QC is focused on the TempVal program for temperature (Angel et. al., 2003), and the PrecipVal program for precipitation (Urzen et. al., 2004). These programs compare the suspect value against a reference grid derived from automated (e.g., ASOS, AWOS) and remotely sensed (e.g., Nexrad radar) meteorological observations that have passed all respective automated QC checks. If a suspect value exceeds an objectively-prescribed threshold deviation from the grid box within which it resides, its estimated value is adjusted to bring it within an expected range. In 2006, TempVal estimates accounted for less than 1% of all COOP observations evaluated by NCDC QC. Furthermore, of the 5% of maximum, minimum and 'at observation' values declared invalid and given corrective estimates, only 35% of those estimates were provided by TempVal, while 64% were the result of simple date shifting.

Despite the robustness and comprehensiveness of the NCDC QA process, occasionally an unbiased observation may be incorrectly invalidated by a QC check. For example, a legitimate -60° F minimum temperature in Montana in January might be falsely flagged as invalid if the prior day's minimum temperature was -29° F and the subsequent day's minimum temperature was -30° F. In general, such "false positive" occurrences are minimized by virtue of the value having to pass through several checks. In addition, the flag would trigger manual QC review that would correct the erroneous flag. Most truly erroneous values are sufficiently anomalous to fail several QC checks (e.g., a -60° F minimum temperature for Miami in July might trigger both the sign and the extremes checks), and thus reasonable confidence can be ascribed to the objective judgment of the automated QC process.

More difficult to identify are false negatives, or erroneous values that are not sufficiently anomalous to trigger QC invalidation. These values, while rare, may occasionally pass through the QA process and be flagged as valid data. Such errors might include a temperature of 41° F that is digitized as 47° F due to the similarity between 1 and 7 in the observer's handwriting. Fortunately, the frequency of errors due to difficulty transcribing an observer's handwriting is diminishing as more and more observers transition to electronic reporting options (e.g., IV-ROCS, WxCoder III) that incorporate basic QC checks (Shein and Owen, 2007).

Although continued refinement and redevelopment of QC practices at NCDC have minimized the incidence of false QC outcomes, NCDC recognizes the possibility that such outcomes may occasionally occur given the millions of in-situ observations that are processed each month. In addition, due to the resources needed to reprocess the entire historical record, archived climate data more than a few decades old may not immediately benefit from current QC practices, and data and QC errors may persist in the historical record.

Thus, there are two QC issues that must be addressed. The first is preventing erroneous QC determinations from being passed into the climate record. The second is the identification and repair of data or QC errors that exist in the historical climate record. To address these issues, NCDC utilizes two primary interactive QC tools – Health of the Network and Datzilla.

3. HEALTH OF THE NETWORK

The need for a tool or process to monitor the operational performance of a global climatological data collection network was identified by the World Meteorological Organization (WMO, 1992). The rationale was to ensure, for long-term climate analysis, the quality and completeness of data being produced from the network. NCDC outlined such a process in 1998 (Steurer, 1998), and the Health of the Network (HoN) tool was subsequently developed. HoN became operational around 2002 and presently monitors several climate networks, including the Global Climate Observing System (GCOS) Upper Air Network (GUAN), the GCOS Surface Network (GSN), and for the United States the COOP network and Climate Reference Network (CRN). Monitoring of the Automated Surface Observing System (ASOS) network and the U.S., Historical Climate Network (USHCN) will be added to HoN in the future.

Over the past five years, the HoN has served as the primary NCDC tool to aid NOAA National Weather Service (NWS) field personnel and NCDC quality control experts in monitoring and evaluating the performance of QC routines applied to the included networks.

After the receipt by NCDC of climate observations, a process requiring a few hours for automated observation systems (e.g., ASOS) but up to 60 days for manual observations (paper forms that must be mailed to NCDC and transcribed into digital format [e.g., COOP]), the data are compiled and evaluated by NCDC's QC process. The complete QC process from ingest to archive takes approximately 90 days, although preliminary data are available much sooner.

During this process QC flags are set and estimated alternate values for invalid observations are produced. The results of the process are parsed into the object-oriented, relational HoN database where they may be viewed by the public via the Internet (see http://www.ncdc.noaa.gov/oa/hofn/index.html). These results are displayed in a variety of formats, including tabular flag reports, graphs and charts, and summary statistics.

The primary purpose of the HoN tool is to provide data users with a way to quickly and efficiently examine the results of NCDC-applied QC on data they may be utilizing (Fig. 1). Through the tabular or graphical output features, a user can, for example, identify a station at which an instrument may be malfunctioning, an observer was not correctly reporting observations, or a station move went unreported. Such issues become clearly evident through the tabular display of systematically applied QC flags, or through graphical indicators such as change points, trends, or spikes in the data.

In addition to individual stations, the HoN provides queries that allow a user to examine the data completeness and quality of all stations in a specific area, such as a NWS region or a particular state. This feature can be especially important to those users who require the data for areal analysis.

Although useful to all data users, since its inception HoN has proven especially important to the NWS, which is responsible for maintenance and supervision of the COOP network. In some cases, an NWS field office may have responsibility for the oversight of several hundred cooperative stations, and station visits are only practical and mandated once or twice per year (NWS, 2005). Through the HoN tool, the NWS employee responsible for COOP station oversight (e.g., Data Acquisition Program Manager, Observing Program Leader) can perform several important tasks.

Figure 1. *Health of the Network COOP station selection screen, showing the various display options for evaluating the QC processes for the particular station.*

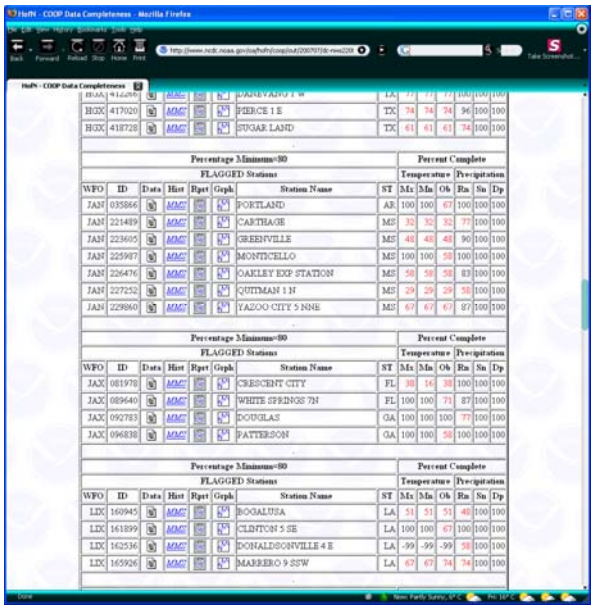

Figure 2. *Tabular display of data completeness for COOP stations in the NWS Southern region for July 2007.*

First, the NWS employee can access the data completeness and data quality information for all observation stations within their jurisdiction for any given month (with a 3-month lag due to the aforementioned processing). In the resulting table (e.g., Fig. 2), the viewer can easily identify stations that, for the month in question, had not reported their observations, failed to take observations, or had observations that were invalidated by NCDC QC.

Second, once identified, HoN provides the viewer with options for graphing data from the station in question, examining it in a tabular format, or examining the available metadata for the station. Visualization of data quality and completeness is accomplished by the display of graphs for particular elements (e.g., maximum temperature, Fig. 3) or, in the case of several stations, by mapping.

All of the available options in HoN are designed to maximize the ability to identify stations that are not operating at their fullest capacity. In doing so, an NWS employee can be alerted to the need for further investigation of the cause for any noted bias. At the same time, the HoN reveals stations at which few or no reporting or QC discrepancies are present, and in which the data user or NWS supervisor can have high confidence.

In the course of HoN-based analysis, or through the use of the data for climate analysis, a value may be identified that was erroneously reported, incorrectly transcribed from an analog source (e.g., paper observation form), or inappropriately flagged by NCDC's QC process. In such cases it is imperative that the

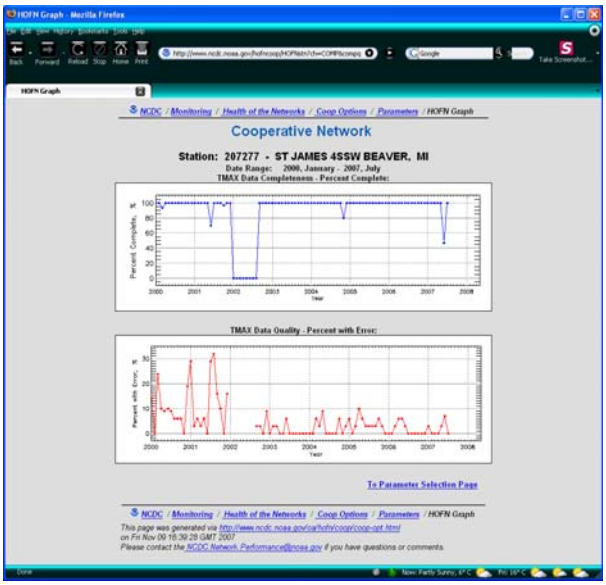

Figure 3. *Graphical display of data completeness and data quality for maximum temperature at the St. James 4 SSW (MI) COOP station from Jan 2000 – July 2007.*

person identifying such an error be able to effectively communicate it to NCDC, and have a way of ensuring the error is satisfactorily corrected. Datzilla is used for this purpose.

4. DATZILLA

Datzilla is a Web-based QA tool that was introduced in early 2005 and is used by authorized account holders to report suspected errors in NOAA-held climate data, and to track the resolution of those errors by the appropriate data managers. Datzilla is based upon Mozilla Project's popular Bugzilla bug tracking system (Robbins, 2005).

A wide variety of errors may be reported in Datzilla, including errors not only in archived climate data, but also in NOAA data delivery and display systems. However, the greatest use of Datzilla has been in the role of reporting errors to NCDC's historical daily climate data.

While HoN is designed to identify suspect data that has been received and processed by NCDC, the purpose of Datzilla is to allow responsible parties (e.g., NWS climate focal points, regional climatologists, network program managers, climate services partners) to notify appropriate data managers of suspected errors in archived climate data. Datzilla also differs from HoN in that it facilitates active communication between error reporters and data managers in order to monitor the error investigation and resolution process from start to finish.

Suspected errors in archived climate data can include inappropriate QC flags or estimated values, filling in of missing data, transcription errors from paper forms (e.g., number transposition, illegibility), and errors in station metadata, to name a few.

Reporting a suspected error via Datzilla is a simple process. A series of query-response type Web forms are displayed that prompt the reporter to enter the source system and data product in which the error was encountered (e.g., Climate Data Online, in-situ summary of the day), the type of error (e.g., flag, value, metadata), and a description of the error (Fig. 4). Additional options allow the reporter to attach files or images to the report. Such attachments are usually scans of original observation forms.

Once a suspected error has been reported, it is assigned to the appropriate NCDC or climate services partner data manager for investigation. Because Datzilla is interactive, the investigator is able to request additional information from the reporter if needed, and the reporter is able to follow the investigation and provide input as desired. In general, the investigator will retrieve and examine the original observation forms or other relevant station reports (often requesting the information from the error reporter if the requested change requires an amended observation form). If a value or metadata in the digital climate archive does not match the reported observation or station information, it is corrected. Likewise, if the investigator determines that the QC process inappropriately flagged or estimated a value, the flag and/or estimate is manually corrected.

Due to the nature of record retention mandated for NCDC and other climate services partners, it should be noted that the ability to implement changes to the

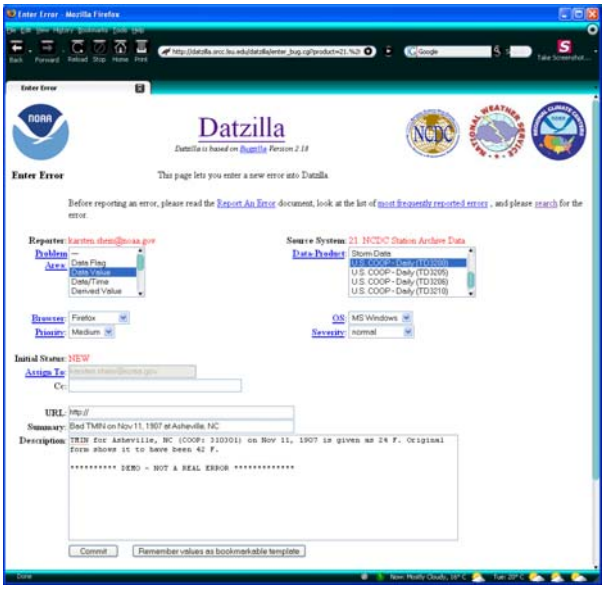

Figure 4. *Screenshot of Datzilla's Error Report entry page, showing a hypothetical error entry. The investigator will be assigned based upon the selected Data Product.*

archives is limited to ensuring the digital data match the original observations as reported or amended, regardless of the verity of the observation; that QC has been appropriately and objectively applied according to acknowledged standards; and that information regarding a station matches what has been officially reported to NCDC by the NWS. If an error cannot be verified in an absolute and objective sense, it is not considered an archive error.

As of 15 January, 2008, 865 error reports have been submitted through Datzilla. Of these, 564 directly concerned possible errors to NCDC archived climate data, and just 109 of those error reports remain open. Further, 317 of the 455 resolved errors were verified and have resulted in corrections to the digital archive.

5. CONCLUSIONS

While the quality assurance of federally-collected climate data by NCDC and its climate services partners (e.g., Regional Climate Centers) is excellent, it can be further improved by involving relevant field personnel and others in NCDC's quality control process. To this end, over the past five years, NCDC and its climate services partners have developed and implemented the Health of the Network and Datzilla interactive quality assurance tools.

The Health of the Network tool allows observation program supervisors at the NWS to review the results of NCDC's QC with respect with the observations at stations they supervise. QC errors noted by these field personnel can be brought to the attention of NCDC via the Datzilla error reporting and tracking tool. In addition, value, QC or metadata errors in the long-term historical record of climate observations can be reported through Datzilla.

With the increased use of digital observation reporting tools such as IV-ROCS and WxCoder III, client-side, baseline QC will further reduce the introduction of errors and invalid observations into the climate record. However, the historical record will continue to benefit from the advantage provided by interactive QC tools such as Health of the Network and Datzilla.

6. REFERENCES

Angel, W.E., M.L. Urzen, S.A. Del Greco, M.W. Bodosky, 2003: Automated validation for Summary of the Day temperature data. 19th AMS Conference on IIPS, Paper 15.3., 4 pp.

Eischeid, J.K., C.B. Baker, T.R. Karl and H.F. Diaz, 1995: The quality control of long-term climatological data using objective data analysis. *J. Applied Met.*, 34, 12, 2787-2795. DOI: 10.1175/1520-0450(1995)034<2787:TQCOLT>2.0.CO;2

Guttman, N.B., and R.G. Quayle, 1990: A review of cooperative temperature data validation. *J. Atmos. Oceanic Technol.*, 7, 2, 334-339. DOI: 10.1175/1520-0426(1990)007<0334:AROCTD>2.0.CO;2.

Hubbard, K.G., N.B. Guttman, J. You, and Z. Chen, 2007: An improved QC process for temperature in the daily cooperative weather observations. *J. Atmos. Oceanic Technol.*, 24, 2, 206-213. DOI: 10.1175/JTECH1963.1

Kunkel, K.E., D.R. Easterling, K. Hubbard, K. Redmond, K. Andsager, M.C. Kruk and M.L. Spinar, 2005: Quality control of pre-1948 Cooperative Observer Network data. *J. Atmos. Oceanic Technol.*, 22, 11, 1691-1705. DOI: 10.1175/JTECH1816.1.

NWS, 2005: Visitation Procedures – Cooperative Observing Stations. National Weather Service Instruction 10-1309. http://www.nws.noaa.gov/directives.

Reek, T., S.R. Doty, and T.W. Owen, 1992: A deterministic approach to the validation of historical daily temperature and precipitation data from the Cooperative network. *Bull. Amer. Meteor. Soc.*, 73, 6, 753-762. DOI: 10.1175/1520-0477(1992)073<0753:ADATTV>2.0.CO;2.

Robbins, K., 2005: Overview of Datzilla. http://datzilla.srcc.lsu.edu/datzilla/documents/DatzillaOverview.html.

Robinson, D.A., 1990: The United States Cooperative climate-observing systems: Reflections and recommendations. *Bull. Amer. Meteor. Soc.*, 71, 6, 826-831. DOI: 10.1175/1520-0477(1990)071<0826:TUSCCO>2.0.CO;2.

Root, C., 1930: The cooperative observer. *Mon. Weather Rev.*, 58, 11, 451-453. DOI: 10.1175/1520-0493(1930)58<451:TCO>2.0.CO;2.

Shein, K.A. and T.W. Owen, 2007: A new paradigm in near real-time cooperative data ingest at NOAA's National Climatic Data Center. 14th AMS Symposium on Meteorological Observation and Instrumentation, Paper J3.20, 3 pp.

Stuerer, P. M., 1998: Assessing the Health of NOAA's Observing Networks. U.S. Department of Commerce publication. 36 pp. (http://www.ncdc.noaa.gov/oa/hofn/coop/ahonon.pdf)

Urzen, M.L., S.A. Del Greco, and S. Ansari, 2004: Automated validation for Summary of the Day precipitation data. 21st AMS Conference on IIPS, Paper 18.13, 5 pp.

WMO, 1992: Guide to WWW data management. WMO-No. 788.

Acknowledgement: Special thanks to Helen Frederick (NCDC) for providing the QC output statistics.