AUTOMATED QUALITY ASSURANCE OF DATA FROM THE OKLAHOMA MESONETWORK

by

Mark A. Shafer and Timothy W. Hughes
Oklahoma Climatological Survey
Norman, Oklahoma

1. INTRODUCTION

During any single month, the Oklahoma Mesonetwork records in excess of 13 million observations of meteorological and agricultural parameters. These observations come from a 111-station mesoscale network, which includes each of the 77 counties in Oklahoma. It has a 5-minute temporal resolution and a 32-km average spatial resolution. Data are collected and disseminated at 15-minute intervals. Each report consists of 13 'core' parameters and up to seven 'supplemental' parameters (Brock et al., 1995). All observations are archived.

Maintaining the integrity of the data for real-time distribution and assuring a research-quality archive data set are critical to the long-term success of the network. Data are used in several large research projects, including VORTEX (Verification of the Origins of Rotation in Tornadic storms EXperiment), ARM (Atmospheric Radiation Measurement program) and the CAPS (Center for Analysis and Prediction of Storms) model. To these ends, the Mesonet uses a variety of means to monitor data quality, including instrument calibration and replacement, visual inspections and an automated quality-assurance (QA) routine.

The Mesonetwork staff includes four full-time technicians responsible for calibration and maintenance of all instruments in the network. They are responsible for both routine maintenance and emergency response for damaged sensors. They also operate an instrument calibration laboratory, where all instruments are checked before being installed at remote locations, as well as verifying instrument calibrations when they are returned from the field.

Visual inspections of data occur on a nearly continuous basis. Observations which appear inconsistent with surrounding stations are noted in a log and checked. The primary means of identifying suspicious data is via standard station-model plots or color-filled contours. Weekly and monthly plots of averages and extrema also are produced. Sensors which are suspect relative to surrounding sites are recorded in a bad parameter (BADPARM) table and reported to technicians for further investigation.

The third means of data quality-assurance is an automated quality-assurance routine. The routine utilizes several approaches, varying in complexity from a range test to a Barnes objective analysis routine (Wade, 1987). The routines must be able to identify suspect data without erroneously flagging mesoscale features in the data. The QA routine also incorporates human input via the BADPARM table. The information from this table is combined with range, step, persistence and spatial tests for each of the 5-minute observations.

The results from each of these tests are combined to form a single data-quality flag for each observation. The individual tests are weighted according to their applicability for a particular parameter. For example, little weight is given to spatial analysis of rainfall or solar radiation, while more weight is given to temperature. Range and BADPARM failures automatically set the flag to failure. The driver program is versatile enough to add new quality assurance tests as they become available.

2. CALIBRATION LAB AND FIELD SUPPORT

Before any instrument is deployed remotely, it is checked in an instrument calibration laboratory. All instruments, when they arrive from the manufacturer, are checked to verify the manufacturer's specifications. Calibration tables for rainfall and solar radiation sensors are determined in the laboratory and then applied to data before they are distributed. Many of the techniques used for calibration were developed at the University of Oklahoma (Brock et al., 1995).

Technicians visit each site a minimum of three times per year. Each technician carries with him high-quality sensors for comparison to the readings in the field. If any sensor does not perform well in the intercomparison, it is replaced and returned to the laboratory. The sensor, if possible, is repaired, re-calibrated and returned to the field. If a sensor is known to be bad, a technician will respond within 48 hours, when practical considerations allow.

Several problems which affect data quality have been identified by technicians on their field visits. During growing season, some grasses grow quickly and can exceed the height of the raingage. Small particulates may fall through the protective wire mesh and clog the funnel. Although technicians carry equipment to maintain the site grounds, they cannot visit frequently enough to keep up with fast grass growth at some sites. Technicians also have noted problems with soil temperature readings due to erosion of the bare soil temperature plots, burrowing animals chewing through the wires, and accumulation of windblown debris over the plots, which then acts as an insulator. Damage has also occurred at several sites from lightning and large hail.

Laboratory calibration tests and field evaluation of sensors have led to design enhancements of instrumentation to improve sensor's performance. Intercomparison data and fluid flow modeling studies (Richardson and Brock, 1995), are producing a deeper insight into the limitations of naturally aspirated solar radiation temperature shields. Shield flow-through design and sensor housing radiation characteristics are much more critical than previously thought. Several improved designs are being tested. Experiments with new designs of the

Table 1. List of Mesonet parameter id's and QA test threshold values. The first group represents 5-minute parameters, the second group represents 15-minute parameters.

				RANGE		PERSIST.	MAX	SPATIAL
PARAMETER	**UNITS**	**ID**	**TIME**	**MIN**	**MAX**	**DELTA**	**STEP**	**STD. DEV.**
Relative Humidity	percent	RELH	5 min	5	103	0.1	20	10.0
1.5 m Air Temperature	degrees C	TAIR	5 min	-30	50	0.1	10	2.5
10 m Wind Speed	m/s	WSPD	5 min	0	60	0.0	40	5.0
10 m Wind Vector	m/s	WVEC	5 min	0	60	0.0	40	5.0
10 m Wind Direction	degrees	WDIR	5 min	0	360	0.0	360	60.0
Wind Direction Std. Dev.	degrees	WDSD	5 min	0	180	0.0	90	60.0
Wind Speed Std. Dev.	m/s	WSSD	5 min	0	20	0.0	10	5.0
Maximum Wind Speed	m/s	WMAX	5 min	0	100	0.0	80	10.0
Rainfall	mm	RAIN	5 min	0	508	0.0	25	50.0
Pressure	mb	PRES	5 min	800	1050	0.1	10	1.5
Solar Radiation	W/m^2	SRAD	5 min	-1	1500	100.0	800	400.0
9 m Air Temperature	degrees C	TA9M	5 min	-30	50	0.1	10	2.5
2 m Wind Speed	m/s	WS2M	5 min	0	60	0.0	40	5.0
10 cm soil temp, sod	degrees C	TS10	15 min	-30	50	0.1	3	2.5
10 cm soil temp, bare	degrees C	TB10	15 min	-30	50	0.1	3	2.5
5 cm soil temp, sod	degrees C	TS05	15 min	-30	50	0.1	5	2.5
5 cm soil temp, bare	degrees C	TB05	15 min	-30	50	0.1	5	2.5
30 cm soil temp, sod	degrees C	TS30	15 min	-30	50	0.05	2	2.5
Leaf Wetness	-	LEAF	15 min	0	100	0.0	50	30.0
Battery Voltage	volts	BATV	15 min	10	16	0.0	3	2.0

raingages led to requirements for modifications, which were adopted by the manufacturer. Retro-fitting of all raingages in the network is still in progress.

Evaluation of sensors in the laboratory and frequent field visits account for a high quality of data being archived from the field observation sites. Comparisons in an operational setting also are possible from an intercomparison site in Norman, located 100 m from an operational Mesonet site. The intercomparison site provides an opportunity to compare instruments deployed at remote sites directly with high-quality sensors, whose cost prohibited their deployment in the network. The intercomparison site also allows evaluation of proposed additional sensors. Mesonet data also can be compared directly to observations from ASOS. Two sites, one in western Oklahoma and the other in eastern Oklahoma, are situated 100 m from ASOS sites. This provides a comparison in different air masses and climatic zones (Crawford et al., 1995).

3. QA FLAGS

All data which are reported from the remote sites are stored in archive files unaltered. To compensate for known or assessed errors, a quality-assurance flag accompanies each value in the database. Flags applied to the data are as follows:

0 = passes all QA tests

1 = suspect data; probably acceptable but be cautious

2 = warning; observation is likely in error

3 = failure; sensor known to be in error or failed

range or multiple tests
8 = no instrument installed
9 = missing data record

The QA program runs each day on archive files.

Each entry contains a station id, time of observation and a string of 13 flags, corresponding to each parameter from the input file (RELH, TAIR, WSPD, etc.). A list of parameters is shown in Table 1. For example:

ADAX 0755 1000000002308

The BADPARM test utilizes a manually-edited table. The table consists of the same flags described above, where parameters at any given station are flagged with 1, 2 or 3 if they have been visually determined to be suspect or bad. Changes to BADPARM flags are checked at each time increment, so that flags are altered during the day as errors are noted or technicians repair or replace instruments.

BADPARM is intended as a manual override. If a site/parameter is listed as suspect, it will be flagged as suspect, warning or failure in the final QA file, depending upon the results of other QA tests. A flag will never be set to a value less than that indicated in the BADPARM file. However, a sensor could be suspect or bad for a while before it is reported and set as such in the BADPARM table. In such cases, other QA routines must indicate the concern.

5. RANGE TEST

The range test is based upon a combination of performance specifications for the instruments and the annual climatology across Oklahoma. Each parameter has set limits (Table 1). Any observation which exceeds the maximum or minimum allowable values is flagged as failure (3). No observations are flagged as suspect or warning. Range and BADPARM are the only tests capable of indicating failure by themselves.

6. STEP TEST

The step test compares the change over a 5 or 15-minute period, depending upon the data reporting time. If a reading changes more than an allowed value, distinct for each parameter, the observation is flagged as 2 (warning). If either one of the data points used in the comparison are missing, the station is flagged as 9. If there is no instrument, it is flagged as 8. The maximum allowable step for each parameter is shown in Table 1. This test has proven useful for catching erroneous readings due to loose wires.

7. PERSISTENCE TEST

In this example, relative humidity (RELH) is flagged as suspect, pressure (PRES) has a warning, Solar radiation (SRAD) is a failure and 2-meter wind speed (WS2M) is not observed at the site. All other data at this observation time (0755 UTC) passed the QA tests. The 15-minute parameters (soil temperatures, leaf wetness and battery voltage) are similarly fashioned.

4. BADPARM TEST

The persistence test checks data in two ways. First, the program retrieves data for a single station over an entire day, one parameter at a time. The mean and standard deviation for that parameter are calculated. The standard deviation is compared to an acceptable minimum. If the calculated standard deviation is less than the acceptable, then the parameter is flagged. The flag is passed back to the subroutine along with the standard deviation.

The subroutine then calls the delta program, which determines the maximum change of a parameter for a single station over six hours. The maximum change is compared with a minimum acceptable change. If the maximum change is less than the minimum acceptable change, then the station parameter is flagged. The value and flag are both returned to the main subroutine.

The persistence subroutine receives the flags from the standard deviation and delta tests. A final flag, which combines the flags from the two subroutines, is derived and written to an output file. The output file contains the combined flag, flags from each subroutine and the standard deviation and maximum change. These tests are useful for finding damaged instruments or those 'stuck' at some particular reading (e.g., due to icing conditions).

8. SPATIAL TEST

The spatial test utilizes a Barnes objective analysis routine to find an estimated value for each valid observation. It is based on an exponentially-weighted observation:

Install Equation Editor and double-click here to view equation.

where z_e is the estimated value of a parameter at a particular station, z_i is each observation and w is the weight applied to the observed value, based on distance from the estimated value being determined (r_i). The weight decreases exponentially with distance from the station:

The weight function, k_o is determined automatically based on the mean station spacing of the network. The radius of influence is approximately 100 km for the Mesonet.

Each separate observation is sent to a Barnes subroutine, which uses every valid observation, except the station which is being analyzed. An estimated value is returned, along with the mean, median and standard deviation of observations within the radius of influence. In the central part of the Mesonet, 20-25 stations are typically included in the estimates. The difference between the observed and estimated value is compared to the standard deviation

Install Equation Editor and double-click here to view equation.

Any difference exceeding twice the standard deviation is flagged as suspect (1) and any over four times is flagged as warning (2). If a station has only 5 or fewer observations, no flag is set. Thus, stations in the Panhandle are not included in the spatial analysis. A second pass is made which does not include observations flagged as warnings in the first pass.

The standard deviation is used rather than absolute thresholds to allow increased variability during situations of large contrasts across the network (Keith Brewster, personal communication). For example, more variability is given to stations along a frontal boundary than for stations in a nearly uniform air mass. In order to avoid flagging stations where the standard deviation is small, a minimum standard deviation is applied to each parameter (Table 1). For example, if the standard deviation of air temperature is 0.5 ^oC, stations departing by more than 1.0 ^oC would be flagged; however a minimum standard deviation of 2.5 ^oC assures that only

Install Equation Editor and double-click here to view equation.

stations differing by more that 5.0 ^oC are flagged. If a standard deviation of 5.0 ^oC is noted across a frontal boundary, stations would have to differ by more than 10.0 ^oC before being flagged.

9. MASTER QA ROUTINE

After each of these QA tests is completed, a master QA subroutine is called. Flags from BADPARM and range tests are included unaltered. Step test results are slightly altered. The step test will not by itself indicate an observation as a failure; it will only assign a warning (2). The warning will be reset to zero if a flag is for rainfall (RAIN) at 0005 UTC, when accumulated rainfall totals are reset. A warning will be reset to suspect (1) if it fails the step test but passes the spatial test. This often occurs when a sensor has spiked, and is returning to a 'normal' value. Wind direction (WDIR) flags are not included because of the discontinuity in wind direction at 360 degrees. Analysis of WDIR from objectively-analyzed u and v wind components is currently under development. Rainfall (RAIN) is not checked at all due to it's extreme spatial variability.

Each of the flags, thus altered, are combined to a single flag by adding the results of each test. Any total over 3 is set to 3. For example, if the persistence test puts a warning on TAIR but no other test noted an error, the flag is 2. If the persistence test warned and spatial indicated a suspect reading, the flag is set to 3 (2+1). If both the persistence and spatial tests issue a warning, the flag is 3 (2+2). If an observation is flagged as 8 or 9 by the range test, the flag is incorporated unaltered in the final flag. These individual flags are then assembled into strings of 13 1-digit integer values for 5-minute data and 7 1-digit integer values for 15-minute data.

10. PERFORMANCE

Quality Assurance flags should be examined closely in the context of other data collected in the area. The automated QA tests will sometimes flag data which are valid, usually in cases where there is a BADPARM flag from a prior report or a small-scale event affecting only one site. Events which affect multiple sites are more likely to pass spatial QA tests.

Thresholds indicated in Table 1 were selected based on experience accumulated during development of the individual routines. These thresholds appear to reasonably screen suspect observations while allowing passage of most valid mesoscale variations. However, some local effects may still get flagged by one or more of the QA routines. The Lahoma windstorm of August 17, 1994 provides a good example. A microburst occurred north of the Mesonet site near Lahoma. Wind speeds (5-minute average) rose quickly to 35 m/s and maximum wind speed peaked at 50.5 m/s. The air temperature fell from 35 ^oC to 12 ^oC. These anomalies triggered warning flags on both the max step and spatial routines, and were indicated as failures by the QA routine.

Three notable events were identified in the May 1995 data set. In each case, one or more observations were flagged as warning or failure, although the data appear to be valid. On May 15, a heat burst was recorded at the Ardmore Mesonet site. Temperature and pressure increased and relative humidity decreased during a period from 0850-0925 UTC. Temperature and relative humidity were flagged with warnings or failures and pressure was flagged as suspect. A dryline passage at Erick on the evening of May 7 caused relative humidity values to be flagged low for approximately five hours. On May 24, a thunderstorm outflow boundary passed the Mesonet site at Ada. At 0235 UTC average wind speed and maximum wind speed were both flagged as suspect, reading too high. Relative humidity was flagged as high for the following 20 minutes. Rainfall was not recorded at the site until 0255 UTC.

Some considerations should be given when using the QA flags. Large spatial variability in solar radiation and rainfall result in frequent flags. Flags in the BADPARM table account for frequent flags on individual observations, despite the appearance of valid data. These flags are often precautionary, and are not removed until a technician had investigated the reported problem.

Despite some instances of erroneous flags, the routines performed well. During the review of VORTEX data, we did not identify instances where a reading in error was not flagged. As is true with any data quality-assurance program, the final decision upon whether or not to include data in an analysis or application rests with the user. Because of the volume of data produced by the Mesonet, a degree of automation is necessary. The aim of this QA development was to provide guidance to those using Mesonet data. To this end, it has met this requirement well.

ACKNOWLEDGEMENTS:

The authors note the contributions of many of the Mesonetwork staff. Ken Meyers, David Grimsley, Gary Reimer and Bill Wyatt maintain the integrity of the instruments in the field. Dr. Fred Brock, Sherman Fredrickson and Scott Richardson test instrumentation enhancements in the calibration laboratory. David Shellberg, Derek Arndt and Dale Morris reviewed suspicious data and maintain visual inspection of data. The QA software development team included Justin Greenfield (step test), Putnam Reiter (persistence test) and Chris Fiebrich (BADPARM test). We also recognize Billy McPherson and Keith Brewster for suggestions during development of the QA concepts.

REFERENCES:

Brock, F.V., K.C. Crawford, R.L. Elliott, G.W. Cuperus, S.J. Stadler, H.L. Johnson and M.D. Eilts, 1995: The Oklahoma Mesonet: A Technical Overview. *J. Atmos. Oceanic Technol.*, **12**, 5-19

_______ and S.J. Richardson, 1995: Passive Multiplate Solar Radiation Shields. *Preprints, 9th Symposium on Meteorological Observations and Instrumentation,* Charlotte, NC, March 27-31

Crawford, K.C., D.S. Arndt, D.J. Shellberg and B.M. Brus, 1995: A Comparison of Differences in Meteorological Measurements Made by the Oklahoma Mesonetwork at Two Co-Located ASOS Sites. *Preprints, 11th International Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, Dallas, TX, January 15-20.

Richardson, S.J., 1995: Passive Solar Radiation Shields: Numerical Simulation of Flow Dynamics. *Preprints, 9th Symposium on Meteorological Observations and Instrumentation*, Charlotte, NC, March 27-31

________ and F.V. Brock, 1995: Passive Solar Radiation Shields: Energy Budget-Optimizing Shield Design. *Preprints, 9th Symposium on Meteorological Observations and Instrumentation*, Charlotte, NC, March 27-31

Wade, C.G., 1987: A Quality Control Program for Surface Mesometeorological Data. *J. Atmos. Oceanic Technol.*, **4**, 435-453.