P1.10 COMPARISON OF AUTOMATED AND MANUAL QUALITY CONTROL TECHNIQUES FOR SURFACE OBSERVATIONS OVER COMPLEX TERRAIN

David T. Myrick[1], Joshua Watson[2] and Judith B. Pechmann[3]
[1]NOAA/NWS Western Region, Scientific Services Division, Salt Lake City, UT
[2]NOAA/NWS Eastern Region, Scientific Services Division, Bohemia, NY
[3]University of Utah, Salt Lake City, UT

1. INTRODUCTION

Numerical models and objective analysis systems use a variety of methods to quality control (QC) observations. As observing networks have grown, manual (or human) QC methods have been replaced with automated techniques to accommodate more observations.

In the western United States, National Weather Service (NWS) weather forecast offices (WFOs) conduct manual QC on mesonet surface observations located within each office's area of responsibility. Bad observations are added to a reject list that is used to create local surface objective analyses. The reject list is organized by variable (i.e., by temperature, dewpoint, wind speed, wind direction, and wind gust) so that only the bad variable is withheld and not the entire observing station. The manual QC process is largely subjective and the methodologies used to evaluate observational quality vary by WFO.

In late 2007, NWS WFOs in Western Region (Fig. 1) began providing lists of flagged (i.e., rejected) observations to the Environmental Modeling Center at the National Centers for Environmental Prediction (NCEP) on a quarterly basis to be withheld from the Real-Time Mesoscale Analysis (RTMA). The RTMA (De Pondeca et al. 2007) is the first step in a multi-year project to build an "Analysis of Record" (Horel and Colman 2005). RTMA surface analyses of temperature, dewpoint, wind and pressure are generated using the NCEP Gridpoint Statistical Interpolation Analysis System run in 2D-var mode for domains centered over the conterminous United States, Alaska, Hawaii, Puerto Rico and Guam.

Corresponding author address: David T. Myrick, NOAA/NWS Western Region, 125 S. State St., Room 1217, Salt Lake City, UT 84138, David.Myrick@noaa.gov

In addition to the manual QC lists provided by NWS WFOs, the RTMA also withholds observations on an automated reject list which is based on QC flags from the Meteorological Assimilation Data Ingest System (MADIS; Miller et al. 2005; Miller et al. 2007). MADIS provides a QC flag for each observation based on automated methods (e.g., validity, internal consistency, temporal consistency, statistical spatial consistency, and spatial "buddy" checks). If an observation fails the MADIS automated QC checks more than 25% of the time during the past month, the observation is withheld from the RTMA.

2. OBJECTIVES

The goal of this study is to compare the reject lists generated by human and automated QC methods. Specifically, this study will address the following questions:

• Are NWS WFOs in the western United States rejecting observations at the same rate? Are some offices stricter when it comes to quality control?

• Are particular mesonet observational networks "flagged" more frequently than others?

• Do the automated and manual QC reject lists used by the RTMA capture the same observations? How do they differ?

3. METHODOLOGY

Statistics on the number of flagged observations were calculated based on reject lists from a six month period (Dec 2007 – May 2008). The manual reject lists from Western Region WFOs were collected at the beginning of each month by Western Region Headquarters, Scientific Services Division. The automated reject lists (using the MADIS QC flags) were created at the beginning of each month by NWS Eastern Region Headquarters, Scientific

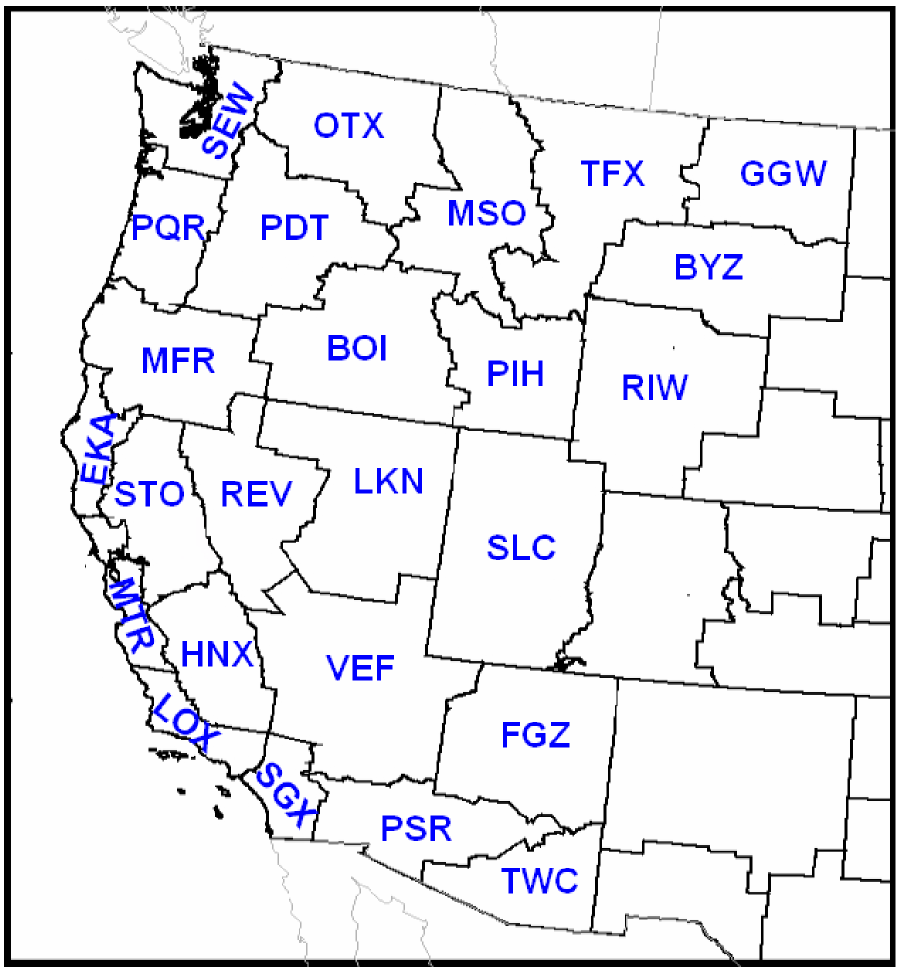

Fig. 1. NWS forecast offices in the western United States.

Services Division. Observations that failed the MADIS QC at least 25% of the time in a given month were included on these lists.

The statistics were calculated on a monthly basis by variable (temperature, dewpoint, wind speed and wind direction) for NWS Western Region as a whole and by WFO. Regional counts were also calculated for the following major observational networks: Automated Surface Observing Stations (ASOS; http://www.nws.noaa.gov/asos/), Remote Automated Weather Stations (RAWS; http://www.fs.fed.us/raws/), SNOpack TELemtry (SNOTEL; http://www.wcc.nrcs.usda.gov/snow/), Union Pacific RailRoad (UPRR; http://www.up.com/), Citizen Weather Observer Program (CWOP; http://www.wxqa.com), and observations not included in the aforementioned networks (OTHER).

The counts presented in this paper are an average of the monthly counts from the six month study period. Monthly trends are not shown because the number of observations flagged by each WFO did not change much from month to month.

Many WFO's leave inactive stations on their local reject lists. The "active flag" from MesoWest (Horel et al. 2002) was used to filter inactive stations. MesoWest is a database maintained at the University of Utah that collects surface weather data from over 150 government agencies and commercial firms across the United States. An observing station must report at least 1 observation in the previous 60 days to be considered "active" by MesoWest.

4. RESULTS/DISCUSSION

4.1 *Manual QC Trends by WFO*

The number of observations on the manual reject lists varied considerably by WFO, suggesting that the methods/criteria used to flag surface observations are not consistent across NWS Western Region. For example, the average number of observations on the manual reject list for temperature ranged from 1 to 200 (Fig 2a). The variability in the number of rejected observations was reduced when the counts only considered active MesoWest stations. For example, the Sacramento, CA WFO (STO) had more than 325 observations flagged for dewpoint on their manual reject list. This number drops to 51 when considering only active MesoWest stations (Fig. 2b). The number of observations from active MesoWest stations on the manual reject lists was generally less than half of the total number of observations flagged due to the inclusion of inactive stations (Fig. 2).

Approximately twice as many observations were flagged for temperature and dewpoint than for wind speed and direction. Western Region WFOs rely heavily on local analyses of temperature and dewpoint to create bias-corrected model guidance. This may explain why temperature and dewpoint observations are flagged more frequently. Wind observations are also considered to be more difficult to quality control and tend to only be rejected when they report egregious speed and direction errors.

Tabulating the manual reject list counts based on active MesoWest stations allows more direct comparisons to be made about the frequency at which different WFOs flag observations. Compared to the total number of active MesoWest stations, the number of active observations that are flagged is relatively small (Fig. 3). The percentage of observations from active MesoWest stations flagged differs by variable and by WFO (Fig. 4). For Western Region as a whole, the percentage of observations flagged is 11% for temperature, 15% for dewpoint and 5% for wind speed and direction. The percentage of observations flagged by individual WFOs is as little as 0% for wind speed and direction to as much as 41% for dewpoint.

The frequency at which WFOs flag observations is related to a number of factors including observational density. WFOs with dense observational networks in metropolitan areas (such as SGX near San Diego, CA and SLC near Salt Lake City, UT) tend to flag more observations than WFOs comprised of mostly rural areas (such as GGW and LKN). One can hypothesize that it is easier to perform "buddy checks" in dense observational networks and thus be able to more easily justify flagging an observation. In rural areas, WFOs may be inclined to not flag an observation that is a bit suspect because they need it to fill a hole in their observing network. The data density argument, however, does not hold for all WFOs. In discussions with the WFOs during this project, it became apparent that the observational QC philosophy varies across Western Region such

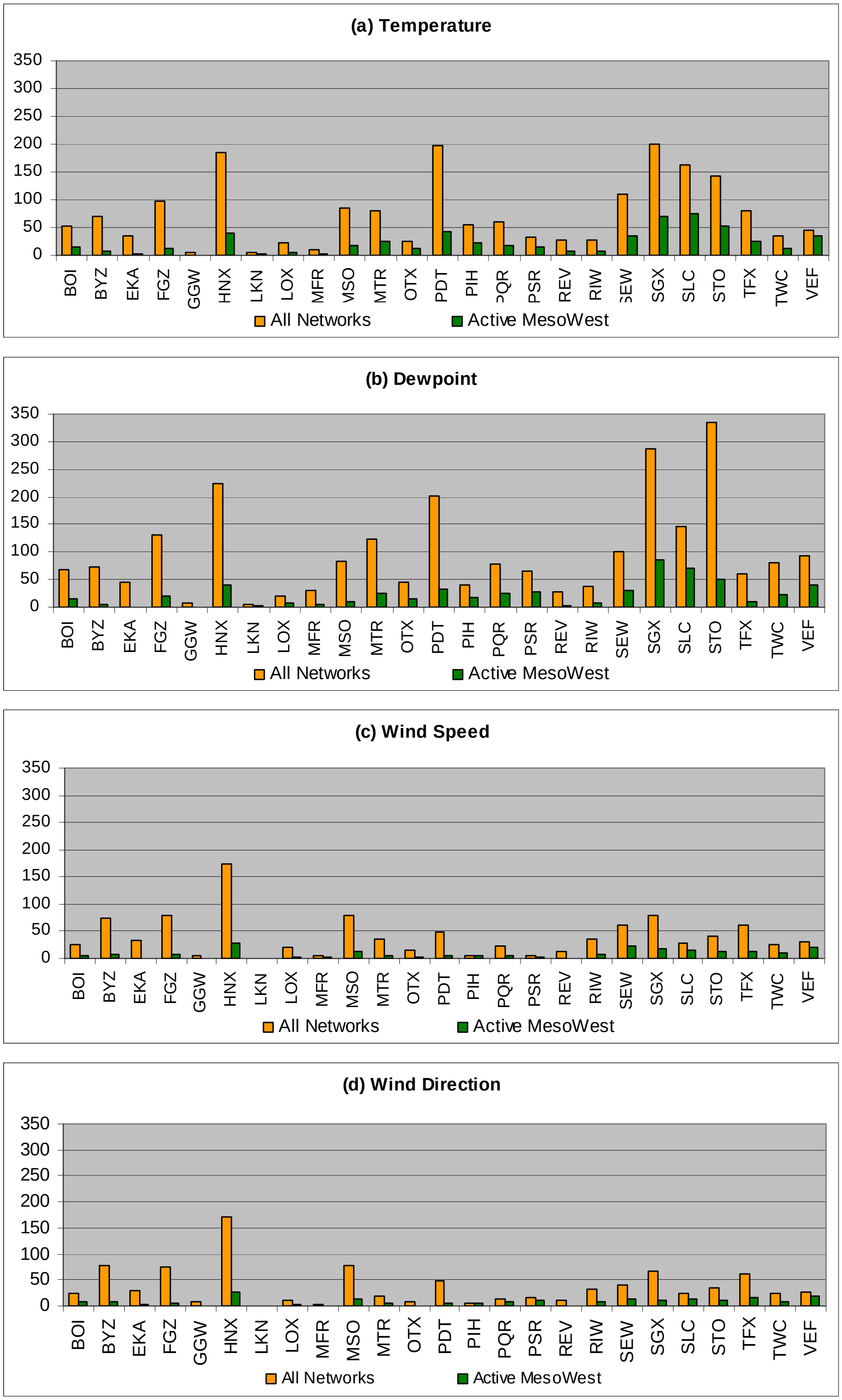

Fig. 2. Average number of observations on WFO reject lists between Dec 2006 and May 2007 for (a) temperature, (b) dewpoint, (c) wind speed, and (d) wind direction.

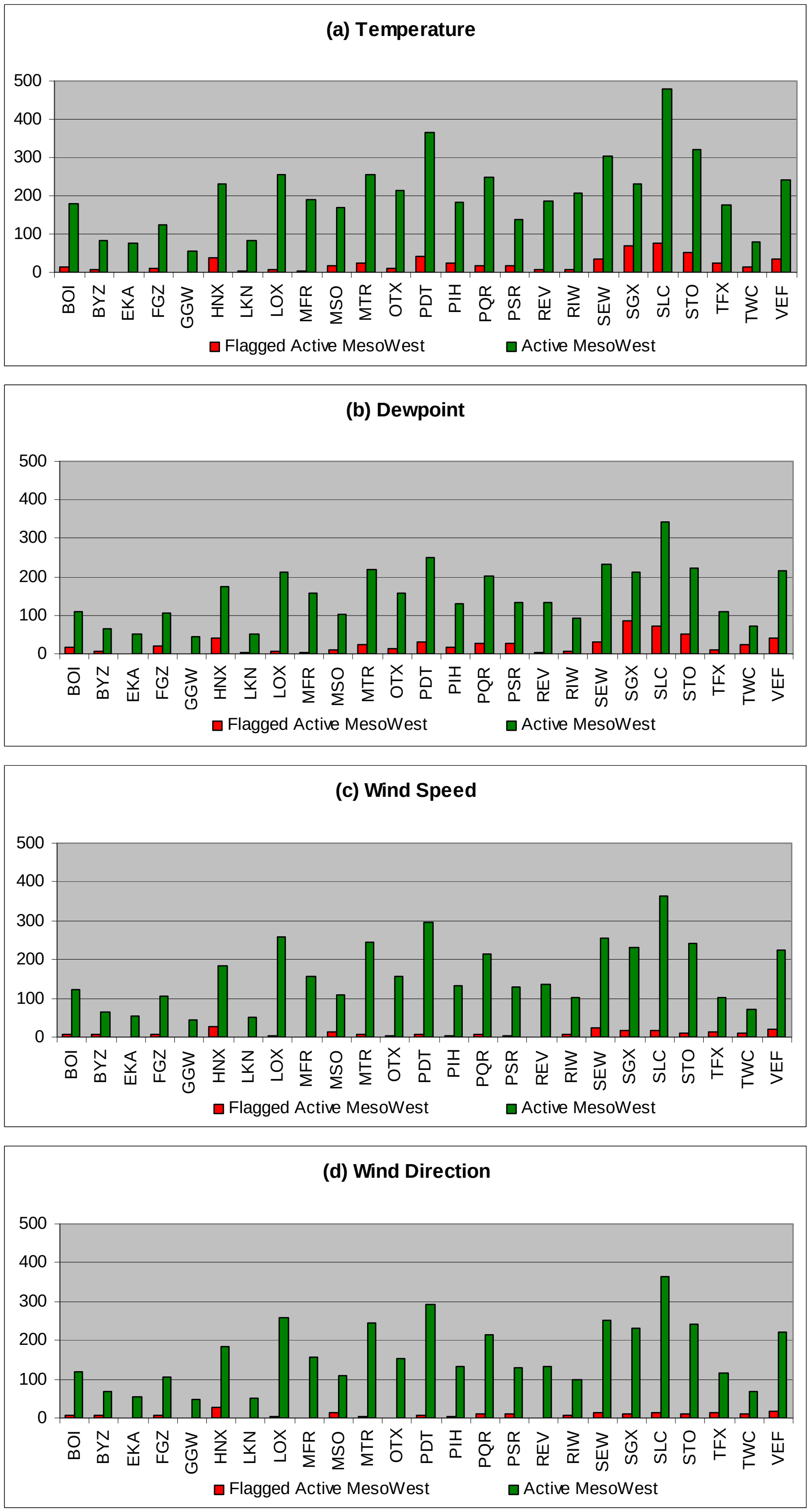

Fig. 3. Average number of active MesoWest observations and flagged active MesoWest observations between Dec 2006 and May 2007 sorted by WFO for (a) temperature, (b) dewpoint, (c) wind speed, and (d) wind direction.

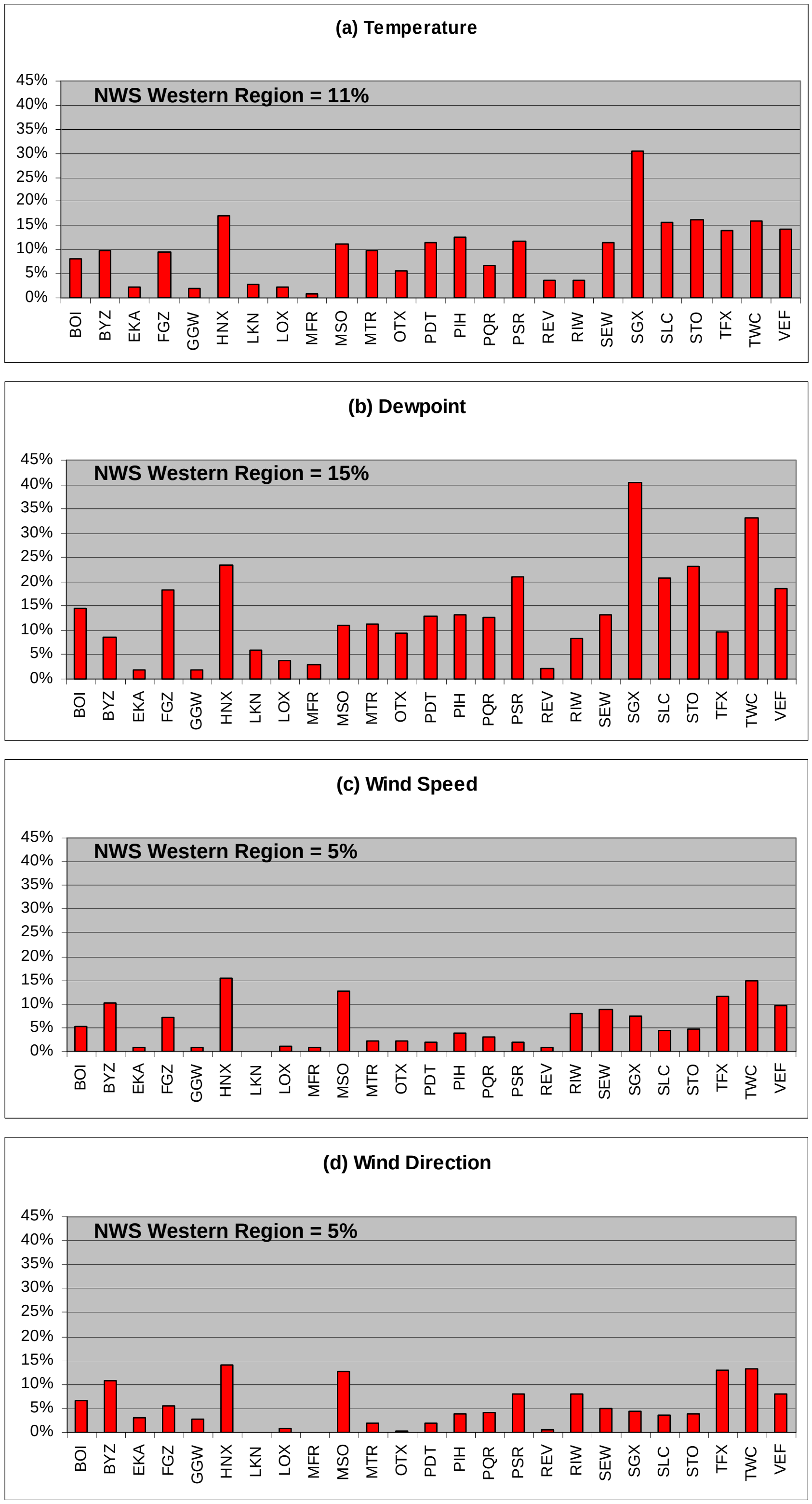

Fig. 4. Percent of active MesoWest observations that were flagged between Dec 2006 and May 2007 sorted by WFO for (a) temperature, (b) dewpoint, (c) wind speed, and (d) wind direction.

that some larger, data dense WFOs tend to only flag stations that exhibit obvious errors while others will tend to flag stations that are not representative of the conditions in the nearby area in an effort to improve local objective analyses.

4.2 *Manual QC Trends by Network*

Western Region WFOs rejected observations from all of the major observational networks (Fig. 5). ASOS observations had the lowest rejection rate (less than 5%) for all of the variables (Fig. 6). UPRR had the highest rejection rate for temperature (34%), wind speed (8%) and wind direction (12%). SNOTEL was the most frequently flagged network for dewpoint (26%). It should be noted that UPRR stations do not report dewpoint.

4.3 *Manual QC Compared to Automated QC*

The number of observations flagged by automated QC was small compared to the number of observations flagged by manual QC. For active MesoWest and non-MesoWest stations on the Western Region WFO manual reject lists, the number of observations flagged by automated QC was only 9% for temperature, 11% for dewpoint, 4% for wind speed, and 3% for wind direction. This suggests that automated QC is flagging the "worst offenders", but has trouble capturing observations with more subtle problems.

The automated QC may not be rejecting enough observations. For an observation to fail automated QC, it must fail MADIS QC at least 25% of the time during a given month. Future work includes testing lower cutoffs (such as 10 or 15%) to determine how many more observations would be flagged.

The process of manual QC is complex and the methods used to flag an observation vary by WFO. Most "bad" observations are flagged when a noticeable error appears in locally generated objective analyses. After tracking down the offending observation, WFOs will add the observation to the manual reject list. Many WFOs use a graphical user interface (GUI) to monitor surface observations. The GUI allows a user to add (subtract) an observation to (from) the local manual reject list.

The process of determining when a flagged observation is no longer in error is a difficult one. Forecast offices will often leave "suspect" observations on the manual reject list for months before they become confident enough to remove them. It is often up to the verification focal point or Science and Operations Officer to check the list of flagged observations on a case by case basis to determine whether any of the observations should be taken off the reject list. There is also the human tendency to flag all 5 observation variables from a station, even though only one observation (e.g., dewpoint) is bad. Thus, the number of observations on the manual QC lists is likely too high.

5. IMPLICATIONS FOR THE RTMA

The differing standards used to reject observations for the RTMA may affect the quality of the analysis. Automated QC seems to reject too few observations, while manual QC seems to reject too many observations and is geographically uneven. An effort should be made to standardize the QC criteria that are used to reject observations from the RTMA.

6. REFERENCES

De Pondeca, M., and Coauthors, 2007: The status of the Real Time Mesoscale Analysis at NCEP. Preprints, *22nd Conf. on WAF/18th Conf. on NWP*, Park City, UT, Amer. Meteor. Soc., 4A.5.

Horel, J., and B. Colman, 2005: Real-time and retrospective mesoscale objective analyses. *Bull. Amer. Meteor. Soc.*, **86**, 1477–1480.

Horel, J., and Coauthors, 2002: Mesowest: Cooperative mesonets in the western United States. *Bull. Amer. Meteor. Soc.*, **83**, 211–225.

Miller, P. A., M. F. Barth, and L. A. Benjamin 2005: An update on MADIS observation ingest, integration, quality control, and distribution capabilities. *21st Conf. on IIPS*, San Diego, CA, Amer. Meteor. Soc., J7.12.

Miller, P. A., M. F. Barth, L. A. Benjamin, R. S. Artz, and W. R. Pendergrass, 2007: MADIS Support for UrbaNet. *14th Symp. on Meteorological Observations and Instrumentation*, San Antonio, TX, Amer. Meteor. Soc., JP2.5.

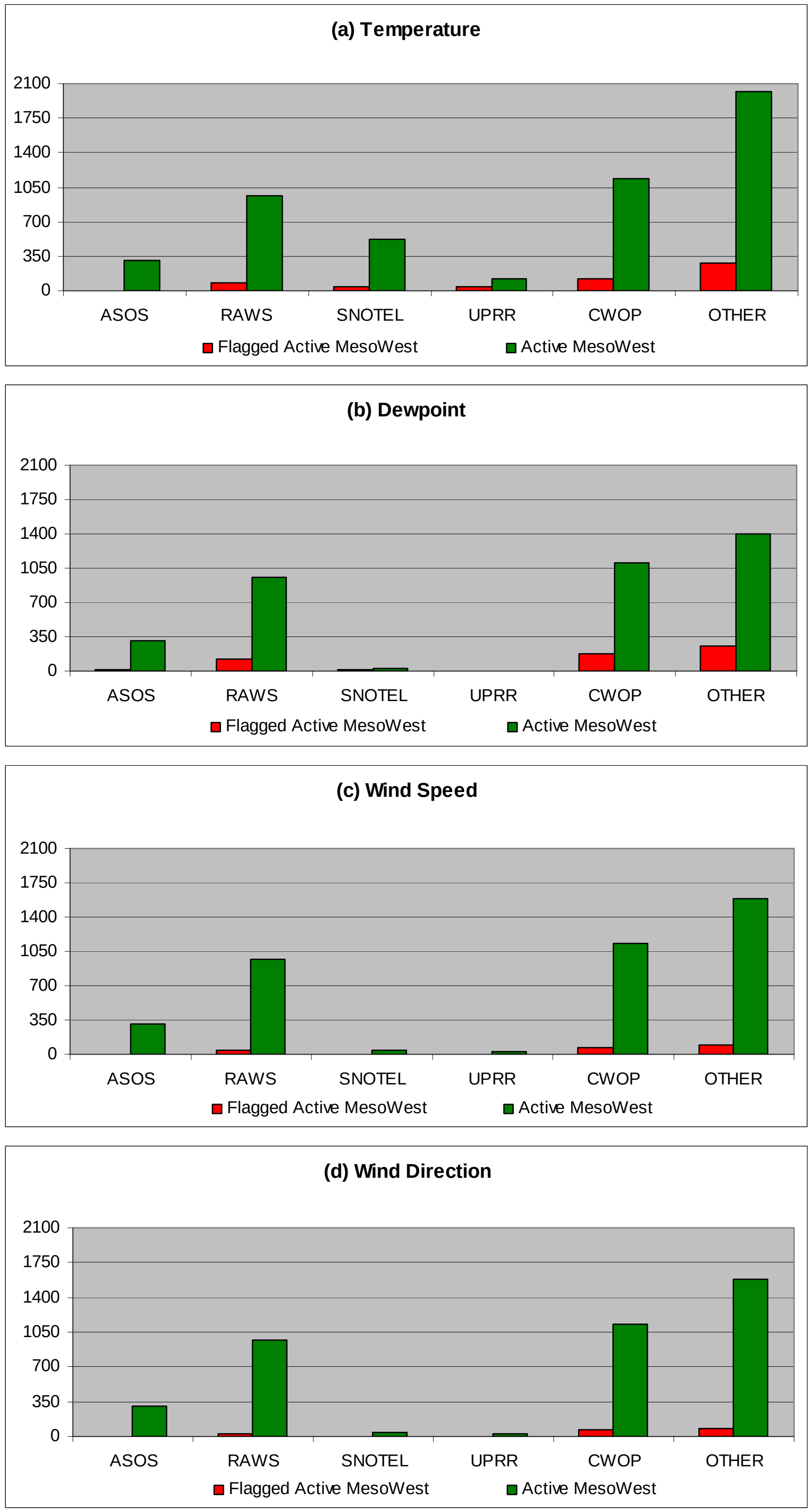

Fig. 5. Average number of active MesoWest observations and flagged active MesoWest observations in NWS Western Region between Dec 2006 and May 2007 sorted by observational network for (a) temperature, (b) dewpoint, (c) wind speed, and (d) wind direction.

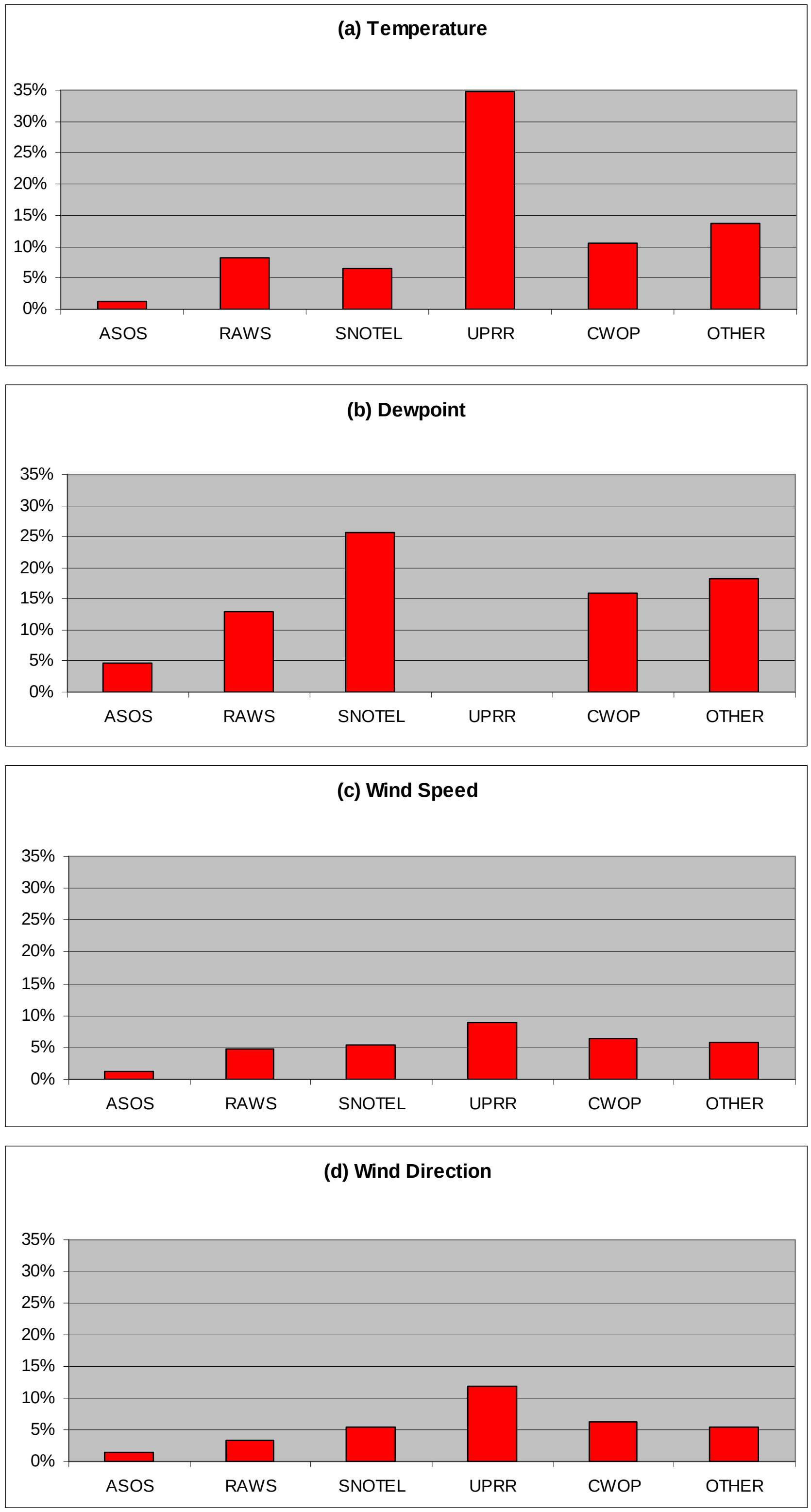

Fig. 6. Percent of active MesoWest observations flagged by manual QC in NWS Western Region between Dec 2006 and May 2007 sorted by observational network for (a) temperature, (b) dewpoint, (c) wind speed, and (d) wind direction.

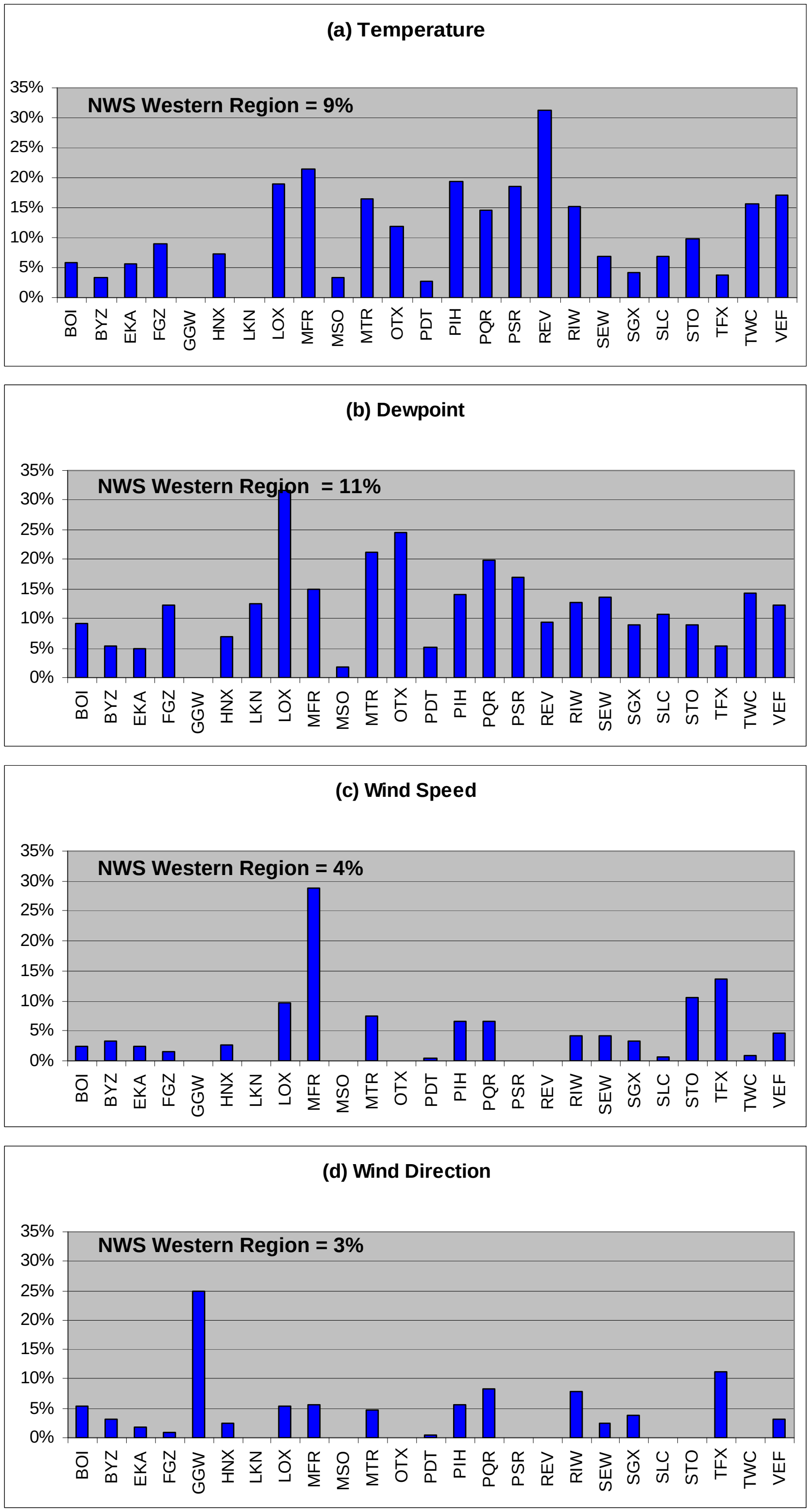

Fig. 7. Percent of observations flagged by manual QC that were also flagged by automated QC between Dec 2006 and May 2007 sorted by WFO for (a) temperature, (b) dewpoint, (c) wind speed, and (d) wind direction.