WORLD METEOROLOGICAL ORGANIZATION

COMMISSION FOR INSTRUMENT AND METHODS OF OBSERVATION
OPAG-SURFACE

EXPERT TEAM ON SURFACE TECHNOLOGY AND MEASUREMENT TECHNIQUES
First Session

Geneva, Switzerland, 13-16 October 2004

CIMO/OPAG-SURFACE/
ET ST&MT-1/Doc.6.1(1)

(20.IX.2004)

ITEM: 6

Original: ENGLISH ONLY

Quality monitoring procedures for data from AWS

Submitted by Igor Zahumenský (Slovakia)

Summary and purpose of document

The document contains a proposal for procedures suitable for quality monitoring of data from AWS.

Action proposed

The meeting is invited to consider the proposal for procedures suitable for quality monitoring data from AWS.

References:

1. Guidelines on Quality Control Procedures for Data from Automatic Weather Stations, CIMO/OPAG-SURFACE/ET ST&MT-1/Doc.6.1(2)
2. Guide to Meteorological Instruments and Methods of Observation, WMO-No. 8
3. Guide on WWW Data Management, WMO-No.788
4. Ingest, Integration, Quality Control, and Distribution of Observations from State Transportation Departments using MADIS, Miller, P.A., Barth, M.F., http://ams.confex.com/ams/pdfpapers/55867.pdf

I. Background

Quality control is the best known component of quality management systems. It consists of examination of data at stations and at data centres to detect errors so that data may be either corrected or deleted. A quality control system should include procedures for returning to the source of data to verify them and to prevent recurrence of the errors. Data quality control has to be applied as real time quality control performed at the station and at data processing centre. In addition, it has to be performed as near real time and non real time control at data processing centre.

Quality monitoring or performance monitoring is both real time and non-real time activity in which the performance of the network or observation system is examined for trends and systematic deficiencies. It is typically performed by the office which manages and takes responsibility for the network or system, and which can prescribe changes to equipment or procedures.

Besides instrument installation, maintenance and calibration, quality of data depends also on the real-time quality control procedures applied during data acquisition and processing, in order to eliminate the main sources of errors. These procedures are specific for each type of measurement but generally include plausible value checks, time consistency checks and internal consistency checks. In AWSs, special built-in test equipment can allow detection of specific hardware errors. Application of these procedures is most important because some errors introduced during the measuring process cannot be eliminated later.

As real-time quality control procedures have their limitations and some errors can go undetected, such as sensor drift or bias, as well as errors in data transmission, performance monitoring at the network level is required at meteorological data processing centers and by network managers.

A critical component of rigorous QC is the accurate and appropriate flagging of data. All AWS data should be flagged using appropriate QC flags. At basic QC level performed at AWS five data QC categories are enough:

- good (accurate; data with errors less than or equal to a specified value);
- inconsistent (one or more parameters are inconsistent);
- doubtful (suspect);
- erroneous (wrong; data with errors exceeding a specified value);
- missing data.

It is essential that data quality is known and demonstrable; data has to pass all checks in the framework of basic QC. In case of inconsistent, doubtful and erroneous data, additional information should be transmitted; in case of missing data the reason of missing should be transmitted.

It is important to establish effective liaison procedures between those responsible for monitoring and for maintenance and calibration, to facilitate rapid response to fault or failure reports from the monitoring system.

II. Proposal

Successful real time QC monitoring has to include checks of the following items:

- ✓ Completeness of observations at the meteorological station;
- ✓ Quality of data;
- ✓ Completeness and timeliness of the collection of observational data at the centre concerned.

QC monitoring is intended to identify deficiencies and errors, monitor them and activate appropriate remedial procedures. Some assessment can be and should be performed in real time, whereas other evaluations can only be accomplished after gathering of sufficient data over a longer period.

QC monitoring requires the preparation of summaries and various statistics. Therefore, it is necessary to built up QC Monitoring System which has to collect different statistics on observational errors of individual meteorological variables, through a series of flags indicating the results of each check, and at least generate hourly, daily, weekly, monthly and yearly summaries of:

- The total number of observations scheduled and available for each variable (completeness of data);
- The total number of observations which failed the QC checks for each variable (quality of data) in case of:
    - Plausible value check,
    - Time consistency check,
    - Check on a maximum allowed variability of an instantaneous value,
    - Check on a minimum required variability of instantaneous values,
    - Internal consistency check;
- The percentage of failed observations (quality of data);
- The error and threshold values for each failed observation (reason of failure);
- Root mean square (RMS) error / mean error / percentage failure for failed observations for each station (daily/weekly/monthly/yearly) (quality statistics).

Stations with large percentages of failed observations are most likely experiencing hardware or software failures.

QC Monitoring System has to keep station monitoring statistics on the frequency and magnitude of observation errors encountered at each station. The statistics provides information for the purpose of:

- ✓ Monitoring quality of station performance,
- ✓ Locating persistent biases or failures in observations,
- ✓ Evaluating improvement of quality of observation data, performance and maintenance of station/network.