Data Screening & Panel Discussion

AQS Conference

August 23, 2012

What We Will Cover

- Data Screening currently performed by AQS
  - Robert Coats, 15 mins
- Data Screening performed outside AQS (30 mins)
  - AirNow Data Screening: Hilary Hafner, Sonoma Technology, Inc., 10 mins
  - Wisconsin: Grant Hetherington, 10 mins
  - Bay Area Air Quality Management District, Duc Nguyen, 10 mins
- Discussion / Q&A (15 mins)

Background

- Why are we having this discussion?
  - AQS implements a set of data screening processes for submitted sample measurements (raw data). The statistical tests date to a paper published in 1978. Recently, the following questions has been asked:
    - Are these tests still valid?
    - Are there other tests that should be performed instead of, or in addition to the present set of tests.
    - AirNow has a different set of tests, and several states are reported to have a different set of tests, both of which are reported to be useful.

Purpose

- The quality of AQS data is important for both economic reasons, and human health. The purpose of data screening is to improve data quality.
- The purpose of data screening is not to find or exclude all “bad data”, but to find data points that are statistically unlikely, and therefore should be manually reviewed.
- Data screening does not replace monitoring QA, but only adds an additional validation.

AQS Data Screening

- Presently AQS applies the following screening tests to submitted sample measurements (RD transactions):
  - Limit Tests
  - Pattern Tests
  - Gap Test
  - Shewhart Test
  - Critical Review Tests
- The results of these tests are reported to the user via the Stat_CR report, which is generated as part of the AQS Load Process.

AQS Limit Tests

- AQS performs two limit tests on sample measurements – an absolute limit test and a relative limit test.
  - Absolute test: Each AQS Sampling Methodology has an Absolute Maximum and an Absolute Minimum defined. If a sample measurement is above the Absolute Maximum or below the Absolute Minimum, it is rejected and not loaded at all.
  - Relative test: Each AQS parameter has a Relative Maximum and a Relative Minimum. If a sample measurement is above the Relative Maximum or below the Relative Minimum, it is flagged as an anomaly to be investigated.

AQS Pattern Tests (1):

- AQS performs a set of pattern tests for the following pollutants: Ozone, CO, SO2, and NO2. The thresholds for failure vary by pollutant and climate. Exceptional Event flagged data is excluded.
  - The Dixon test (not applied to CO) scans each day's values and determines the highest, second highest, and lowest values in that day. It then computes the Dixon ratio, defined as (max – second_max)/(max -low). If this value is greater than 0.55, the day fails the Dixon test and all hours are marked as failing.
  - The max hour test compares each value in the month to a constant to determine if the value is too high. If so, it is marked as failing.

AQS Pattern Tests (2):

- The high difference test compares the value at each hour in the month to the previous hour and the subsequent hour. If the difference between any two hours is greater than allowable, then the hour under inspection is marked as failing.
- The spike test works much like the high difference test, except that both differences must be greater than allowable for the test to fail. Also, if the percentage difference between the hour in question and both its adjacent hours is greater than allowable, the test to fails.
- The high consecutive values test looks at each hour and the subsequent three hours. If all four values are greater than allowable, then all four hours are marked as failing the test.

AQS Gap Test:

- The gap test is performed on data for Ozone, CO, SO2, and NO2. The gap test is so named because it looks for gaps in the Frequency Distribution Table for a month's values. For each pollutant, the program builds a Frequency Distribution Table and computes constants associated with the frequency distribution. Having determined the constants, a largest reasonable gap is estimated. Then the largest actual gap in the data is determined and compared to the largest estimated gap to determine whether the month passes or fails. Exceptional events are excluded.

AQS Shewhart Test:

- The Shewhart test is performed on daily data for Lead, PM 10, and PM 2.5. The program counts the number of valid samples for the current month and each of the three previous months. The program computes the mean and range for the current month. It then computes the historical mean and range, from the mean and range of the data for the three historical months. The mean and range for the current month are compared against the historical values to determine whether the current month passes or fails. Exceptional event data is excluded.

AQS Critical Review Tests:

- AQS also implements a set of screening tests called “critical review tests”. They are as follows:
  - Any value that exceeds the 3-year historical maximum is reported as an anomaly.
  - The first maximum of of a monitor-year is greater than second maximum by a specified threshold it is reported as an anomaly.
  - Any value exceeds of the NAAQS standard level by a state-specific factor is reported as an anomaly.