Data Management: Your Date with Data

Mark Stoelting
Quality Assurance Officer
Bay Area AQMD
14 January 2015
The Life of Data
The Life of Data

Data Acquisition System (DAS)

Data Management System (DMS)

Data Validation Process
Generic Data Flow Diagram

- EPA Quality Assurance Handbook Vol. II
- Chapter 14: Data Acquisition and Information Management
Data Management System

**QA Handbook definition:** Data management is the “development, execution and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets”

**Practical definition:** A set of tools for collecting, transmitting, tracking, storing, and reporting *data* from origination through a series of validation steps to the *Users* (‘customers’).
Six Data Management Principles

- Data integrity (internally consistent)
- Accurate and appropriate
- Chain of custody
- Consistent and appropriate change controls
- Documented procedures (SOPs)
- Failure tolerant (Disaster Recovery)
Presentation Objectives

• What is Data Management
• What are the tools of Data Management
• What functions do they provide
• What management tools do Air Districts in CA use
• How are they used to:
  • Reduce workload
  • Improve data quality
  • Get better dates (with data)
Data Acquisition Phase

Data Acquisition System (DAS)

- **Collects** raw data
- **Pushes** data to DMS – OR – waits for **Polling**
- **Controls** local functions
Two Basic DAS Types

- **Analog**
  - one data channel per voltage
  - limited resolution

- **Digital** (much preferred)
  - numeric data transfer (exact resolution)
  - includes instrument metadata
  - requires digital instruments
Popular Ambient Monitoring Dataloggers

- **Agilaire** (ESC) 8872
- **DR DAS** Envidas Ultimate (EnvidasFW)
- Campbell, Omega, EcoTech, etc.
- Custom (CARBLogger, BAAQMD DAS)
Telemetry options

- Broadband (DSL, cable): inexpensive, urban areas

- Cellular: moderate cost, coverage dependent
Telemetry options

- Satellite: expensive, global coverage
- Plain Old Telephone Service (POTS): cheap, slo-o-o-o-o-w
Telemetry options

- SneakerNet: no added cost; site visits must happen anyway...
- Exercise bonus
- Not fault tolerant!
Data Management Phase

Data Management Systems (DMS)

- Exports real-time data to web pages, AQMIS, AirNow, etc.
- Provides data review/validation resources for staff
- DMS systems generally do the following:
  - Tracks data from incoming sources
  - Automatically performs:
    - Level 1 validation
    - Auto QC checks
  - Tracks regulatory data uploaded to AQS
Data Storage and Validation Tools

Current CA Data Management Solutions:

- AirVision (Agilaire)
- Envista ARM (DR DAS)
- DMS (BAAQMD, SCAQMD, CARB)
- Custom (Microsoft Access)
- LEADS/MeteoStar
- EcoTech
- EMC, Quick Link, AirPlus, WINDS, Excel
Data Management use in CA

19 Agencies Responding

- 7 use AirVision (Agilaire)
- 3 use Envista (DR DAS)
- 3 use DMS (agency-owned)
- 1 uses custom/Access
- 1 uses EMC
- 4 use mixture of Excel, WINDS, Air Plus, Quick Link, ESC
Data Management System Functions

Automated tasks

• Track data promptness
Station Reporting Status

![Station Reporting Status](image)
Data Management System Functions

Automated tasks

• Track data promptness
• **Apply autoQC checks**
• **Alerts staff** (email, text, etc.) when data validation criteria are exceeded or otherwise invalidated
• Records data chain-of-custody events and edits
AutoQC in Action

• At 10:05 AM, DMS ingested the hourly data file from the San Rafael Air Monitoring Station (ID#3005) DAS

• DMS recorded:
  ➢ A 9AM 985 µg/m³ ambient value from a PM2.5 femBAM
  ➢ A BAM ‘M’ status flag at 9:19AM
AutoQC in Action

- At 10:06, the new data triggered a Range Check for values > 980 µg/m³
- QC Check invalidated the value with QC Code 113
- QC code 113 generates a Null Code ‘BA’ (Maintenance and Routine Repairs), a typical cause for a 985 µg/m³ BAM value and indicated by the ‘M’ flag

### DM5 - Chain of Custody

<table>
<thead>
<tr>
<th>QC Date</th>
<th>User Name</th>
<th>Site Name</th>
<th>Parameter Unit</th>
<th>Model</th>
<th>Date</th>
<th>New Value</th>
<th>New QC Code</th>
<th>New OP Code</th>
<th>Old Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/21/2014</td>
<td>System</td>
<td>San Rafael</td>
<td>PM25_fem (µg/m³)</td>
<td>BAM1020_fem</td>
<td>10/21/2014 9:00 AM</td>
<td>985.00000</td>
<td>113</td>
<td>0</td>
<td>985.00000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Old Op Code</th>
<th>QC Comment</th>
<th>AutoQC Check</th>
<th>AutoQC Test Site</th>
<th>AutoQC Test Parameter Unit</th>
<th>AutoQC Test Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Automatic Import: 3005 2014 294.MIN</td>
<td>Range: &gt; 980</td>
<td>San Rafael</td>
<td>PM25_fem (µg/m³)</td>
<td>985.00000</td>
</tr>
</tbody>
</table>
AutoQC in Action

- Alert emails sent to AirMon staff at 10:08

  Test: (3005) San Rafael / PM25_fem (ug/m3_LC) / Range: > 980
  Triggers: 1 from Oct 21 2014 10:00AM to Oct 21 2014 10:00AM
  Target/QC: (3005) San Rafael / PM25_fem (ug/m3_LC) / 113 – Maint/Repair

- Station Operator reviewed data at 8:43 AM the following morning without change

<table>
<thead>
<tr>
<th>Reviewed On</th>
<th>User Name</th>
<th>Note</th>
<th>Site Name</th>
<th>Parameter [units]</th>
<th>Data Date</th>
<th>Locked</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/22/2014 8:43 AM</td>
<td>Chris Rumm</td>
<td>Reviewed.</td>
<td>San Rafael</td>
<td>PM25_fem (ug/m3_LC)</td>
<td>10/21/2014 9:00 AM</td>
<td></td>
<td>BAM1020_fem</td>
</tr>
</tbody>
</table>

- Final data status in DMS:

<table>
<thead>
<tr>
<th>R</th>
<th>L</th>
<th>C</th>
<th>Local Standard Time</th>
<th>Data</th>
<th>Status</th>
<th>QC Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td></td>
<td></td>
<td>10/21/2014 08:00</td>
<td>7.0</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>R</td>
<td>C</td>
<td></td>
<td>10/21/2014 09:00</td>
<td>985.0</td>
<td></td>
<td>113</td>
</tr>
<tr>
<td>R</td>
<td></td>
<td></td>
<td>10/21/2014 10:00</td>
<td>6.0</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>
Data Management System Functions

Automated tasks

• Track data promptness
• Apply autoQC checks
• Alerts staff (email, text, etc.) when data validation criteria are exceeded or otherwise invalidated
• Records data chain-of-custody events and edits
• **Tracks Operational Criteria** (shelter temperature)
• **Monitors Instrument Metadata**
NOx Instrument Pressure

Detects and alerts staff on rapid pressure changes > 1 mmHg

- AutoCalibration
- Particulate Filter Change
- Pump Diaphragm Failure, Filter debris?
Data Management System Functions

Data Review

• Display/review ‘strip chart’ data (1-min plots)
Data Review

![Graph showing NOx concentrations over time](image-url)
Better Data Review

![Graph showing NOx concentrations over time with data points for San Francisco - Arkansas St. NO2 (ppb) for 1-Hr and 1-Min intervals. The graph includes a time scale from 12/26/2011 to 12/27/2011.](image-url)
Data Management System Functions

Data Review

- Display/review ‘strip chart’ data (1-min plots)
- **Pollutant relationships**
- **Meteorological effects**
- **Local Sources**
Comprehensive Data Review

Concord - 2956-A Treat Blvd - m_WSpm (m/s) vs Concord - 2956-A Treat Blvd - m_WDir (deg)

z2036 - Concord

Concord - 2956-A Treat Blvd - NO (ppb) - 1 Min
Concord - 2956-A Treat Blvd - NCO2 (ppb) - 1 Min
Concord - 2956-A Treat Blvd - O3 (ppb) - 1 Min
Concord - 2956-A Treat Blvd - StaT (degC) - 1 Min
What makes a good date (with data)?

- Electronic (digital) data management
- Data tracking
- Automated QC checks
- Instrument metadata monitoring
- Prompt problem resolution
- Know your instrument and pollution sources
- Frequent, timely, and efficient data review opportunities (Data Dates)
Data Management System Functions

Data Validation

- Log entries
- Calibration tracking
- Meeting QA Handbook Validation Criteria
  - QA Handbook Volume II, Appendix D
  - Data validation templates for all criteria monitoring Programs: O3, CO, NO2, SO2, PM2.5, PM10c, PM10 (several), Lead
Thank You!

from your DMS.

Questions?
Contact

Mark Stoelting
Bay Area AQMD
mstoelting@baaqmd.gov
415-749-4619