A cable car is the central focus, positioned on a hillside. In the background, a large body of water (likely San Francisco Bay) is visible, with a cityscape and mountains in the distance. The overall scene is a classic San Francisco view.

AutoQC and Graphical Data Review, Keys to Efficient Data Management

National Air Quality Conference
Denver, CO, May 14, 2012
Mark Stoelting
Bay Area AQMD

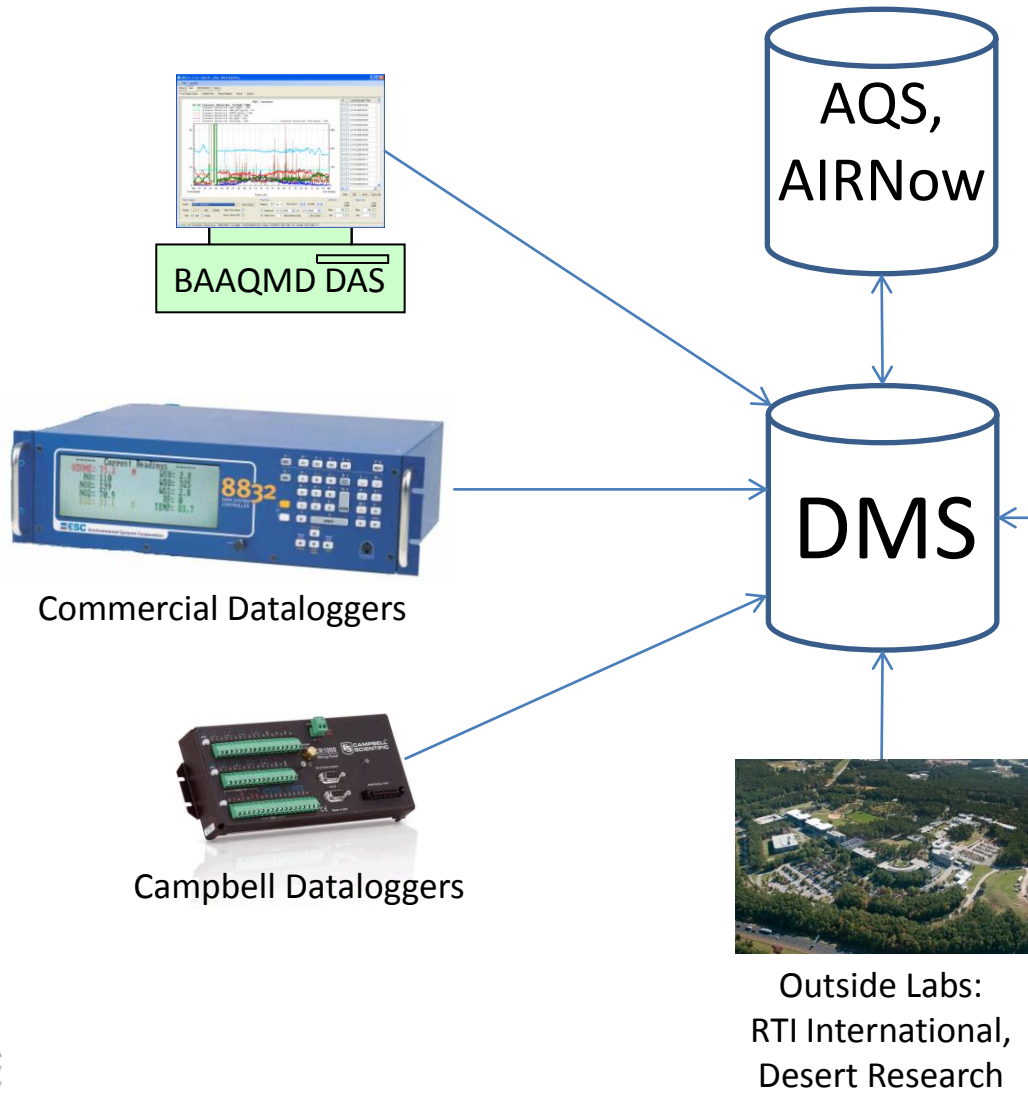
Presentation Topics







- How Bay Area AQMD collects and manages data
- Optimizing data collection at 1-minute resolution
- OpCoding data for efficient data management
- Data collection redundancy
- Managing field operations with Scripts (autocalibrations, maintenance, instrument changes, etc.)
- **Automated data Quality Control**
- **Graphical review**

Bay Area AQMD Data System

- Uses a custom, PC-based Data Acquisition System (DAS) at Air Quality monitoring sites
- Uses Campbell Scientific dataloggers at meteorological towers
- Uses a range of methods for importing laboratory data (toxics, carbonyls/aldehydes, XRF metals, gravimetric PM filters, OC/EC, ion concentrations, PM2.5 speciation, etc.)
- Uses a custom Database Management System (DMS) that can manage many different types of environmental data

DMS Manages Agency Data From Many Sources



- In-house Lab:**
-  Agilent 7890 GC/MS - toxics
 -  Varian HPLC - carbonyls - aldehydes
 -  Thermo QuantX XRF - metals
 -  PM Lab - gravimetric filters
 -  DRI Model 2001 - OC/EC
 -  Dionex ICS 5000 - ions

Custom PC Datalogger

DAS features critical for autoQC and graphical data review (common to other, commercial dataloggers as well)

- 1-min data resolution based on instrument averages
- Powerful scripting capability (autocal, maintenance, etc.) with data OpCoding
- File-based (unlimited) data storage

Optimized Data Collection

- Collect at the finest resolution available from the instrument (generally about one minute)
- Approximates most hardware data resolution
- Required by NCORE
- Optimize data collection:
 - Set instrument averaging interval to 1 minute
 - Set instrument digital record reporting interval to 1 minute
 - collect 1-minute records from the instrument every minute
 - maximize instrument data storage (if possible)
- Allows graphical data interpretation and review (approximates strip charts)

TECO 49i (ozone) example

Set instrument operating mode:

- Command to set the averaging interval to 60 s
 - ‘set low avg time 3’
- Command to set the record format (lrec)
 - ‘set lrec format 1’ sets format to ASCII with text
- Command to set the electronic reporting interval to 1 minute
 - ‘set lrec per 1’ sets logging interval at one minute
- Command to poll the latest electronic record = ‘lrec’; (some instruments automatically send new data without polling)
- Send instrument setup commands when DAS starts
- Resulting 1-min data record collected by the DAS:

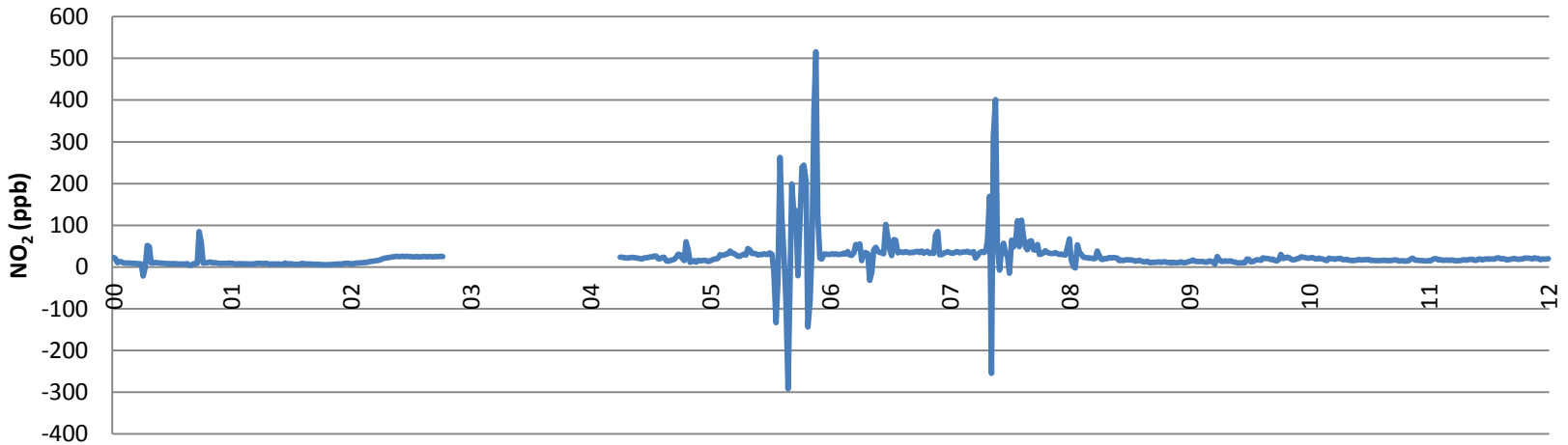
```
00:10 05-16-11 flags 1C100000 o3 17.433 noisa 2.400 cellai  
99806.000 cellbi 99782.000 bncht 29.761 lmpt 53.929 o3lt 76.789  
flowa 0.773 flowb 0.861 pres 719.143 noisb 2.200
```

Make instruments do the first phase of data collection!

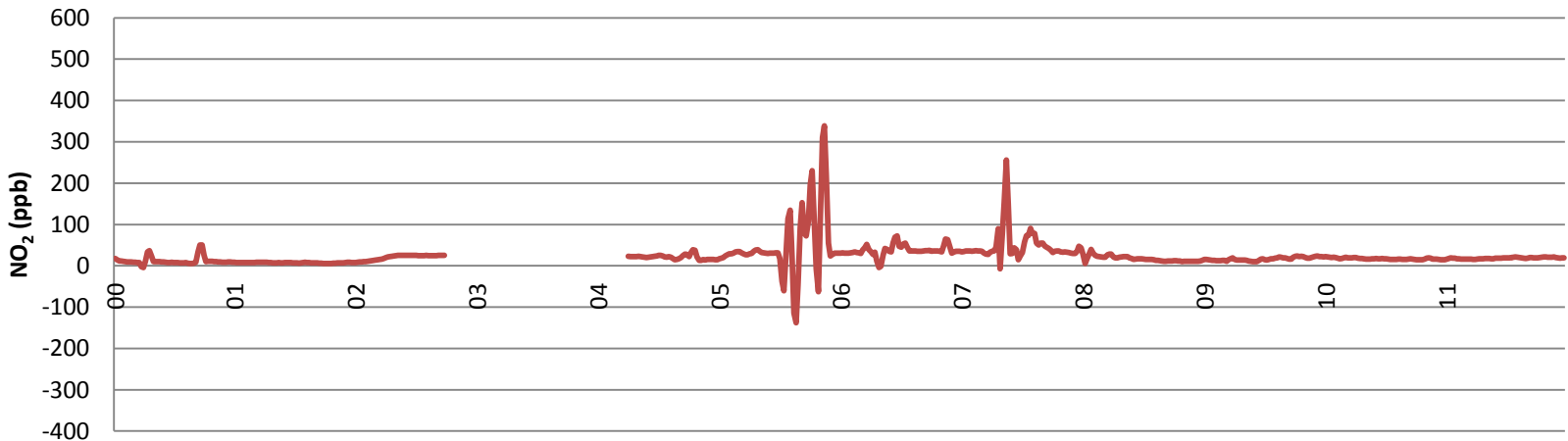
Instrument Averaging Times

Longer averaging times reduces data resolution

Instrument 60 s Averaging



Instrument 180 s Averaging



DAS Data Collection

- Locally archive all 1-minute data from instruments (first-level data file backup)
- Translate data into a database-ready input format:

<site> | <param> | <time> | <OpCode> | <value> | <status>

8004 | flags | 05-16-11 00:10 | 0 | | 1C100000

8004 | o3 | 05-16-11 00:10 | 0 | 17.433 |

8004 | noisa | 05-16-11 00:10 | 0 | 2.400 |

8004 | cellai | 05-16-11 00:10 | 0 | 99806.000 |

8004 | cellbi | 05-16-11 00:10 | 0 | 99782.000 |

8004 | bncht | 05-16-11 00:10 | 0 | 29.761 |

8004 | lmpt | 05-16-11 00:10 | 0 | 53.929 |

8004 | o3lt | 05-16-11 00:10 | 0 | 76.789 |

8004 | flowa | 05-16-11 00:10 | 0 | 0.773 |

8004 | flowb | 05-16-11 00:10 | 0 | 0.861 |

8004 | pres | 05-16-11 00:10 | 0 | 719.143 |

8004 | noisb | 05-16-11 00:10 | 0 | 2.200 |

- Archive into daily data files (2nd-level data file backup)

OpCodes

- Attaches a function and implied data validity to each measurement
- Are assigned by the datalogger when raw instrument data is reformatted into database records
- Are derived from automatic or manual scripts (programs) that control calibration functions, instrument status, etc.

All data collected by the DAS is uploaded to DMS regardless of data function or validity

OpCodes for Gaseous Data

DAS assigns an OpCode

DAS	Op Code function	DMS
Op Code	Name	QC Code
0	ambient	0
1	zero	0
3	gas precision	0
4	gas mid low	0
5	gas mid high	0
6	gas span	0
9	invalid	9
13	gpt precision	0
14	gpt mid low	0
15	gpt mid high	0
16	gpt span	0
23	ozone precision	0
24	ozone mid low	0
25	ozone mid high	0
26	ozone span	0

DMS automatically assigns an initial QC code to raw data:

0 = valid

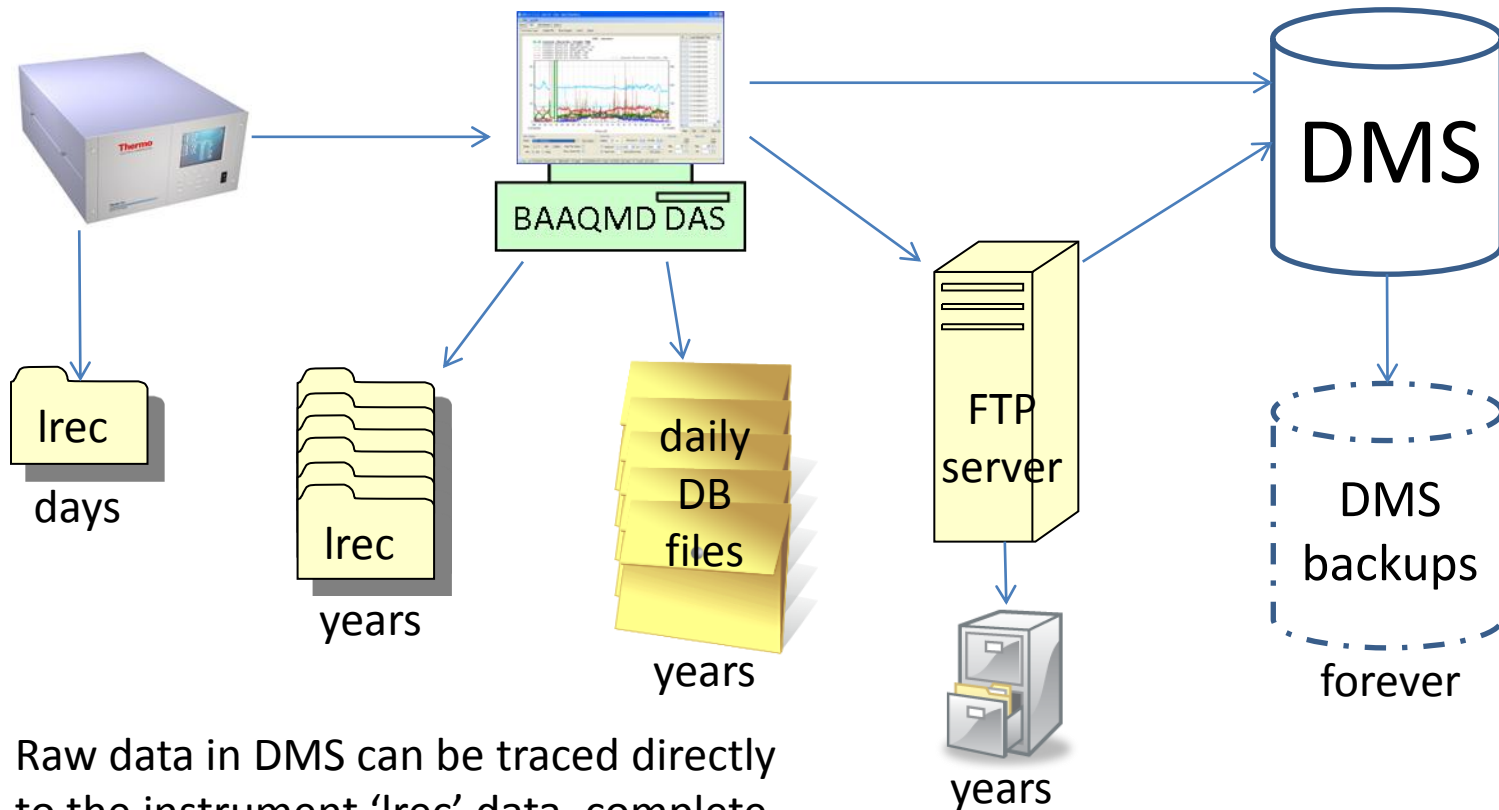
9 = invalid

DAS Communications

- Every hour, each DAS pushes (uploads) the daily data file to DMS (data collected since midnight)
 - Uses standard internet protocols like FTP
 - Simultaneous data collection from all sites
 - Redundant data collection since midnight
 - Very simple data transfer method
- Upload (FTP) server can provide a third level of collective data file backup plus firewall security
- Every day after midnight, DAS pushes past days' files to fill communication gaps (4 days for BAAQMD)
- Data gaps are rare and easily filled

DAS collects, formats, uploads all raw 1-min data without changes

Data Collection Redundancy and Traceability



Raw data in DMS can be traced directly to the instrument 'lrec' data, complete with date/time stamps.

Who's Reporting??

A simple, visual tool reports network reporting status

DMS v1.1.7.0 - dms-01\dms - dms - Mark Stoelting

File Help

Status Data Administration Setup

About Stations Event Log QC Log

Settings

Agency: [Dropdown] Tardy At: 4 hrs. Late At: 8 hrs. Level of Detail: Station Instrument Parameter Refresh

Station Reporting Status

Agency	Station ID	Station Name	Last Data Received	Hours Elapsed
CA2	9903	m_Sonoma Baylds	05/08/2012 07:00	14
CA2	2903	m_Concord	05/08/2012 19:00	2
CA2	3902	m_Valley Ford	05/08/2012 19:00	2
NDA	5803	m_Mission Bay	05/08/2012 19:00	2
CA2	2037	San Ramon	05/08/2012 19:13	2
CA2	1804	m_Oakland STP	05/08/2012 20:00	1
NDA	1805	m_Berkeley Lab	05/08/2012 20:00	1
CA2	1903	m_Chabot	05/08/2012 20:00	1
CA2	1905	m_Pleasanton	05/08/2012 20:00	1
CA2	1907	m_Livemore Rincon	05/08/2012 20:00	1
CA2	1908	m_Patterson Pass	05/08/2012 20:00	1
CA2	2901	m_Bethel Island	05/08/2012 20:00	1
CA2	2905	m_Kregor Peak	05/08/2012 20:00	1
CA2	2907	m_Pt. San Pablo	05/08/2012 20:00	1
CA2	2908	m_San Ramon	05/08/2012 20:00	1
CA2	4902	m_Napa	05/08/2012 20:00	1
NDA	5801	m_San Francisco STP	05/08/2012 20:00	1
CA2	5905	m_Fort Funston	05/08/2012 20:00	1
CA2	6901	m_San Carlos	05/08/2012 20:00	1
NDA	7803	m_Stanford	05/08/2012 20:00	1
CA2	7901	m_San Martin	05/08/2012 20:00	1

DAS Scripts

- Scripts are text files that execute on the DAS to send digital commands to instruments and calibrators, control relays, solenoids and pumps, and **assign OpCodes**
- Scripts can conduct calibrations, change instrument settings, invalidate data, set instrument clocks, etc.
- Can be scheduled by the DAS for automatic execution

ABORT=85

***0=

***1=\B3sel gas a span 4\0D

|8|0000 0000 0011 1111 0000 0001 0000 0111|prec gas a cal***

***24=

| |0000 0000 0000 0000 0000 0001 0000 0111|Measure precision***

***26=\B3sel gas a span 4\0D::\B3sel ozon 4\0D

|8|0000 0000 0011 1111 0000 1011 0000 0111|Precision GPT***

***49=

| |0000 0000 0000 0000 0000 1011 0000 0111|Measure GPT precision***

***51=\B3sel ozon 4\0D::\B3sel gas a span 0\0D

|8|0000 0000 0011 1111 0000 1010 0000 0011|Precision ozone***

***74=

| |0000 0000 0000 0000 0000 1010 0000 0011|Measure Precision Ozone***

***76=\B3sel gas off\0D::\B3sel ozon perm off\0D

|8|0000 0000 0011 1111 0000 0000 0000 0000|Purge started***

***89=

| |0000 0000 0011 0000 0000 0000 0000 0000|Purge Ended***

***134=

| |0000 0000 0000 0000 0000 0000 0000 0000|SO2 Recovery end***

- Glen Colwell, BAAQMD Air Monitoring Manger, provides more details on Wednesday

Automated Data QC

- Tests configured in DMS to check incoming raw data and resulting averages
- Tests include Range, Zero, Sticking, Rate of Change, Absolute Difference, Compare, and Compare % Difference
- Alters QC codes based on test results
- Creates Data Chain-of-Custody records
- Sends Alert emails to staff

QC Check	Test Site	Test Parameter	Test Duration	Target Parameter	Target Duration	Start Hour	End Hour	Value	Data Points	QC Code	Email Notifications
Range (>)	1023 - Livemore - Rincon Ave.	O3 (ppb)	1 Min	O3 (ppb)	1 Min	0	23	200.000		9 - Invalid	Email (5)
Range (<)	1023 - Livemore - Rincon Ave.	O3 (ppb)	1 Min	O3 (ppb)	1 Min	0	4	-4.000		0 - Valid	Email (5)
Range (>)	1023 - Livemore - Rincon Ave.	StaT (degC)	1 Hr	O3 (ppb)	1 Hr	0	23	30.500		9 - Invalid	
Range (<)	1023 - Livemore - Rincon Ave.	StaT (degC)	1 Hr	O3 (ppb)	1 Hr	0	23	19.500		9 - Invalid	
Range (<)	1023 - Livemore - Rincon Ave.	O3 (ppb)	1 Hr	O3 (ppb)	1 Hr	0	23	-5.000		9 - Invalid	
Range (<)	1023 - Livemore - Rincon Ave.	O3_bncht (degC)	1 Min	O3_bncht (degC)	1 Min	0	23	15.000		0 - Valid	Email (5)
Range (>)	1023 - Livemore - Rincon Ave.	O3_bncht (degC)	1 Min	O3_bncht (degC)	1 Min	0	23	40.000		0 - Valid	Email (5)

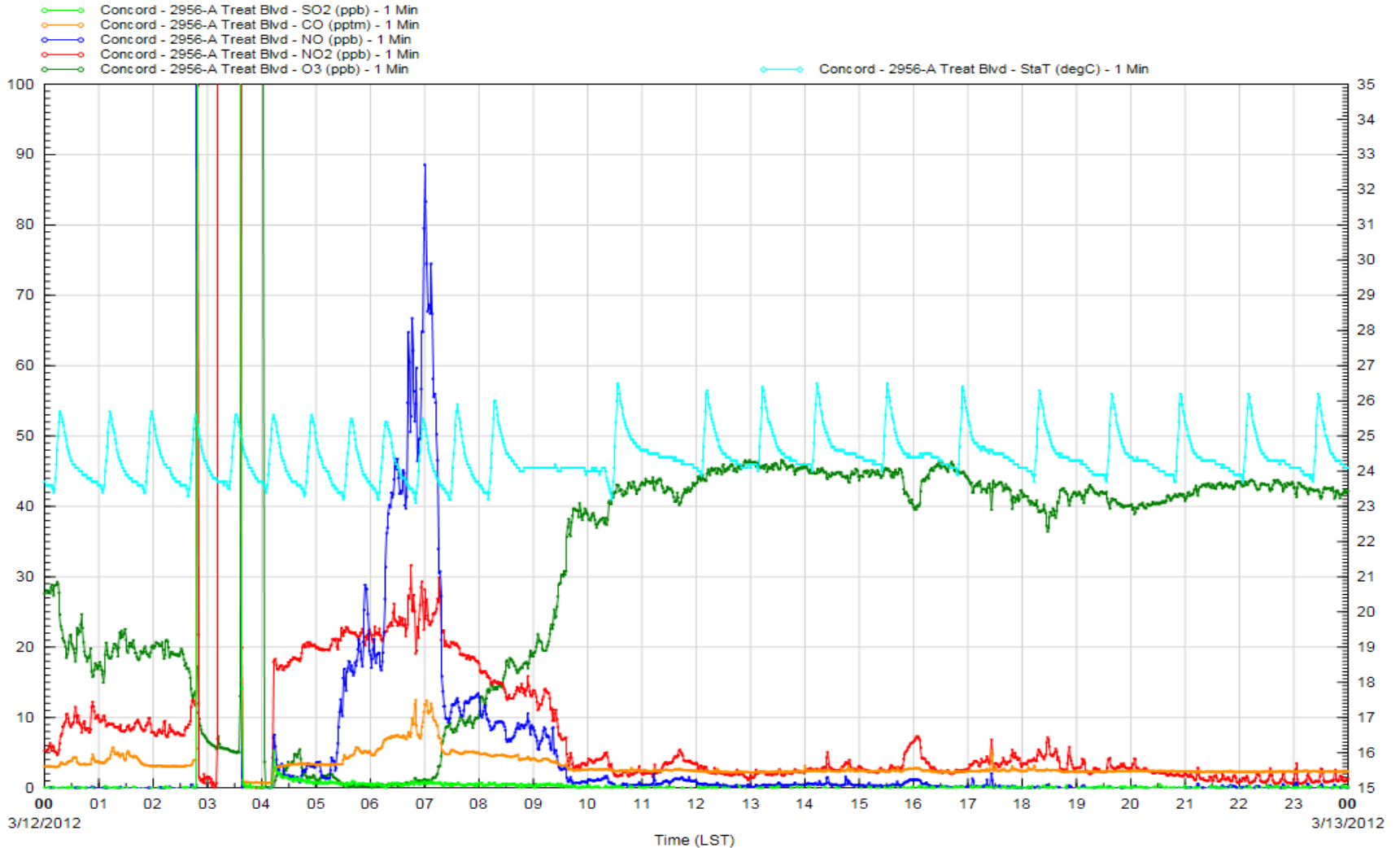
- Checks can apply to 1-min, hourly averages, pollutant concentrations, instrument metadata, etc.
- Glen Colwell provides more details on Wednesday

Graphical Data Review

- DAS/DMS data management tools all focused on efficient data review to improve data quality
- Level 0.5 review is conducted by DMS AutoQC tests every hour before export to web pages, AIRNow, etc.
- **Level 1** review conducted graphically at 1-min resolution by Air Monitoring staff (approximately daily)
- **Level 2** review conducted graphically at the hourly/network level by air quality meteorologists prior to AQS submission

Daily Level 1 Review

2036 - Concord



Analysis

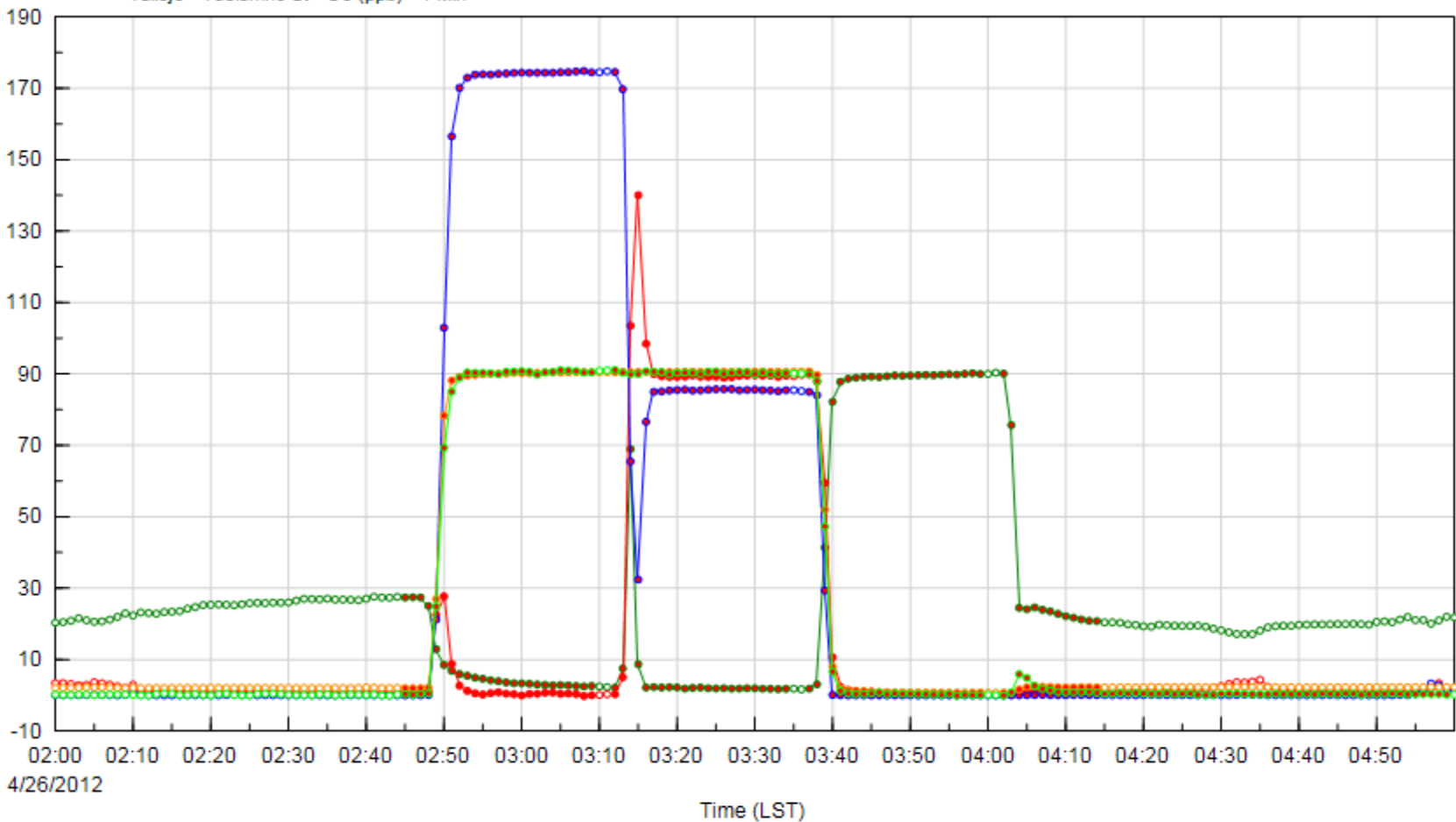
(1-day review panel composed of 8600 data records)

- March 12 was particularly cold in the Bay Area
- Overnight Concord temperatures approached freezing at 2°C and nearby locations were colder
- Pre-dawn winds were nearly calm indicating a strong, stable temperature inversion
- Sunrise occurred at about 6:30 AM and surface heating began during the morning commute
- The surface temperature inversion was consumed about 7:30 AM and aloft ozone was circulated down to the station analyzers and diluted surface pollutant concentrations

Precision AutoQC with OpCodes

8004 - cal

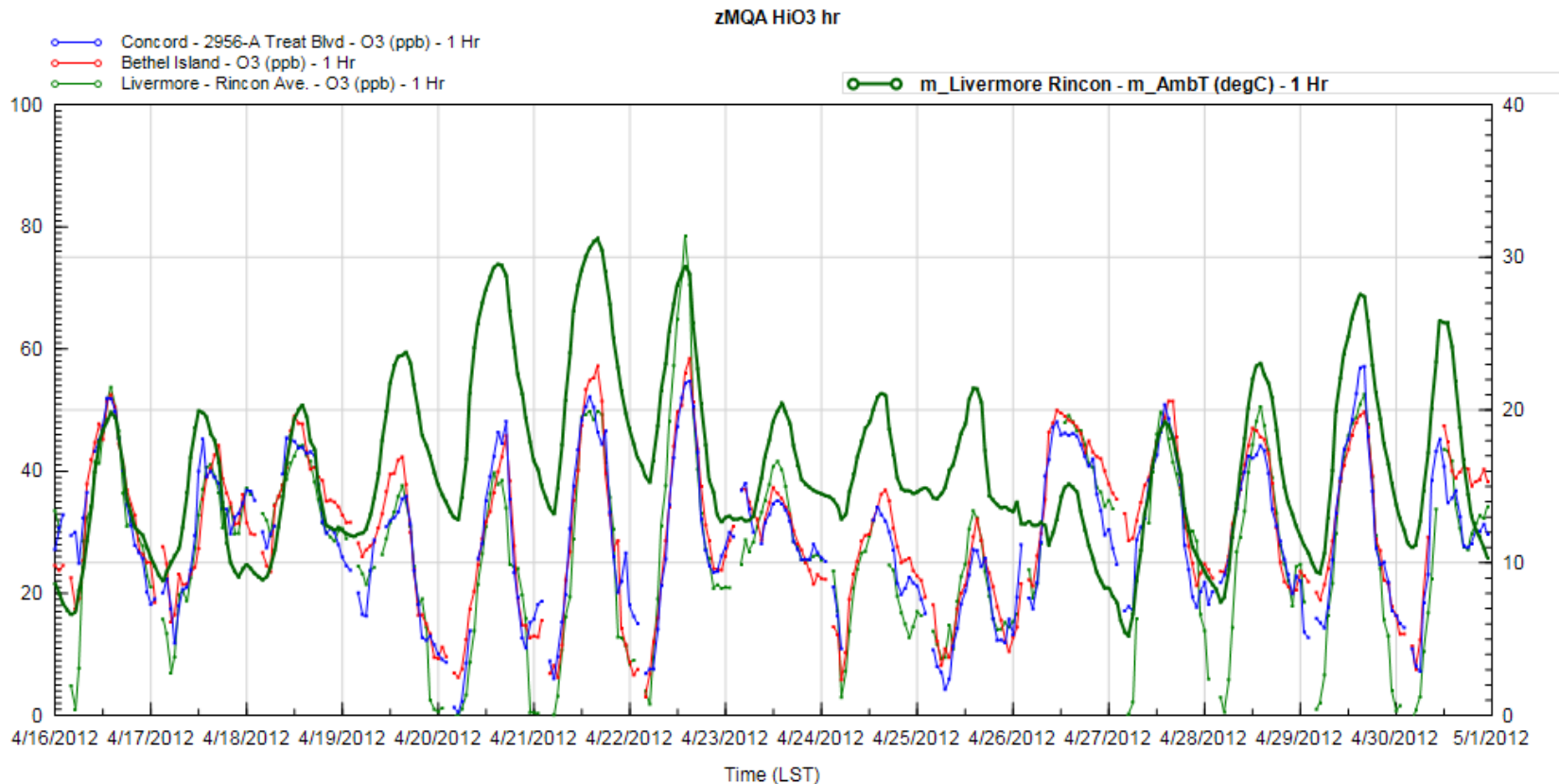
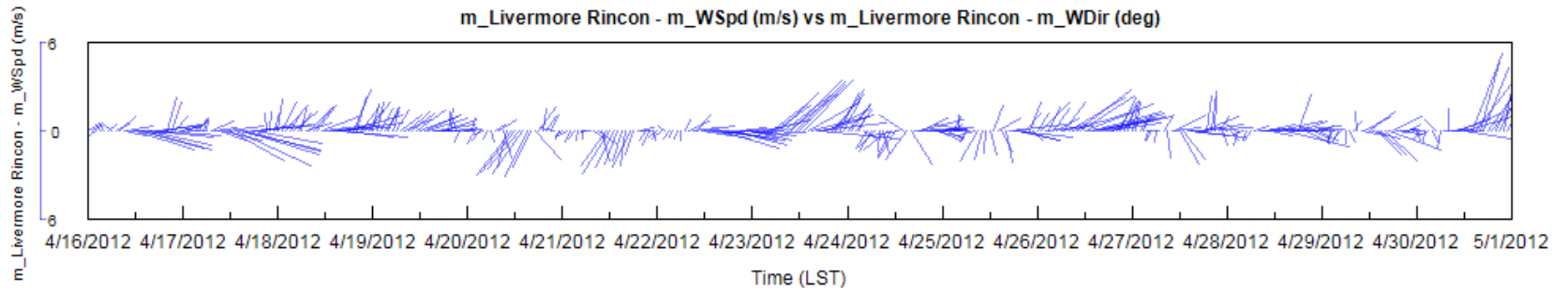
- Vallejo - Tuolumne St - SO2 (ppb) - 1 Min
- Vallejo - Tuolumne St - CO (pptm) - 1 Min
- Vallejo - Tuolumne St - NO (ppb) - 1 Min
- Vallejo - Tuolumne St - NO2 (ppb) - 1 Min
- Vallejo - Tuolumne St - O3 (ppb) - 1 Min



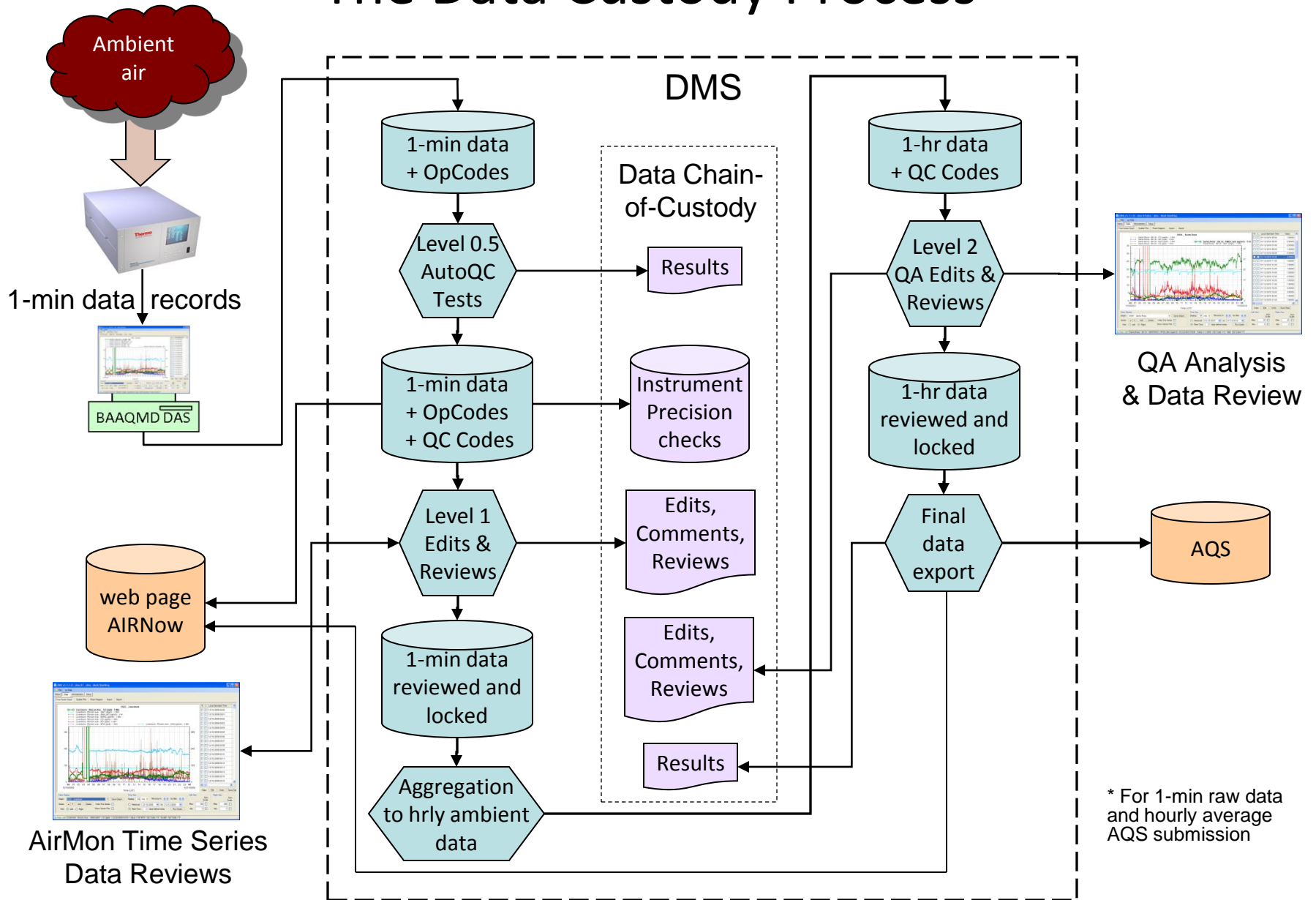
Precision AutoQC

- AutoQC Script starts at 0245 hrs and ends at 0414 hrs (except SO₂), allowing valid ambient monitoring hours at 0200 and 0400
- Calibrator system concentrations True Values were 90 ppb for Ozone, NO₂, and SO₂, 90 pptm for CO, and 177 ppb for NO
- CO, NO, and SO₂ gas precision OpCode = 3
- NO₂ GPT precision OpCode = 13
- Ozone precision OpCode = 23

Monthly Level 2 Review Example



The Data Custody Process*



* For 1-min raw data and hourly average AQS submission

Managing Data Volume

Managing 1-minute data involves very large amounts of data. DMS:

- ingests 1M metadata values/day (instrument temps, flows, voltages, etc.)
- ingests 100K ambient values/day
- creates 2000 hourly averages/day
- purges about 1M metadata values/day after Level 2 data review is complete
- maintains a working volume of about 0.5B records

Review

- All DAS/DMS data management features presented here are available in other systems to one degree or another
- 1-minute raw data, OpCoding, and autoQC tools are critical for efficient, high resolution data management and graphical review
- More complex and difficult monitoring environments in urban settings require more powerful tools for review/analysis

Graphical Review Evaluation Opportunities

- In the 9AM Wednesday morning Technical Session on Ambient Air Monitoring Network Automation, Glen Colwell presents Air Monitoring autoQC and data review procedures
- Sonoma Technology is hosting a DMS demonstration at the conference
- A DMS demonstration is also available using Remote Desktop:

host = dmsterm.sonomatechdata.com

UserName = DMSDemo@sonomatech.com

Alternate UserNames = DMSDemo1 or DMSDemo2

Password = 1Welcome!

To log into DMS, User=Guest, PW=Welcome

Contact Information

Mark Stoelting

Bay Area AQMD

San Francisco, CA

mstoelting@baaqmd.gov

415-749-4619