
Monitor Level AQS Data Completeness

Howard Schmidt, MS, MBA

US EPA Region 3 Air Protection Division

Air Monitoring Quality Assurance Workshop

June 26, 2014



Overview

- OAQPS/Agency role
- Background on Regional level DV development
- Challenges in Regional DV calculations
- Potential modifications



DV Process Overview

- Monitor level info provided to regional offices
 - Closure, type (SLAMS, SPM), nonattainment, network
 - Two weeks to complete review & provide feedback
- Draft DVs provided to regional offices
 - Compare to in-house data calculations
 - Four weeks to complete review & provide feedback
- Final DVs provided to regional offices
 - Last opportunity to provide feedback to OAQPS
 - Two weeks for last chance to review data
- Final DVs posted to EPA Air Trends website
 - Three month process start to finish



Regional Level DV involvement

- Draft DVs provided to regional offices
 - *Compare to in-house data calculations*
- R3 APD maintains individual spreadsheets for each monitor for each NAAQS pollutant
 - Howard = SO₂, Pb, NO₂
 - Lori = PM, O₃, CO
 - DVs for each pollutant for each monitor calculated based on the form of the standard
 - Pb is easiest; PM most difficult
 - In-house DVs used as supplemental information for reference between annual final DVs from OAQPS



Example Process – SO₂

- AMP 501 – Extract Raw Data (one year)
 - Looking for a complete data set of 8,760 hourly values

```
RD I 51 650 0008 42401 1 1 008 060 20131231 00:00 1.0
RD I 51 650 0008 42401 1 1 008 060 20131231 01:00 1.1
RD I 51 650 0008 42401 1 1 008 060 20131231 02:00 1.3
RD I 51 650 0008 42401 1 1 008 060 20131231 03:00 1.1
RD I 51 650 0008 42401 1 1 008 060 20131231 04:00 1.2
RD I 51 650 0008 42401 1 1 008 060 20131231 05:00 1.4
RD I 51 650 0008 42401 1 1 008 060 20131231 06:00 1.4
RD I 51 650 0008 42401 1 1 008 060 20131231 07:00 1.4
RD I 51 650 0008 42401 1 1 008 060 20131231 08:00 1.5
RD I 51 650 0008 42401 1 1 008 060 20131231 09:00 1.6
RD I 51 650 0008 42401 1 1 008 060 20131231 10:00 4.1
RD I 51 650 0008 42401 1 1 008 060 20131231 11:00 2.9
RD I 51 650 0008 42401 1 1 008 060 20131231 12:00 1.5
RD I 51 650 0008 42401 1 1 008 060 20131231 13:00 0.7
RD I 51 650 0008 42401 1 1 008 060 20131231 14:00 0.6
RD I 51 650 0008 42401 1 1 008 060 20131231 15:00 0.6
RD I 51 650 0008 42401 1 1 008 060 20131231 16:00 0.6
RD I 51 650 0008 42401 1 1 008 060 20131231 17:00 0.7
RD I 51 650 0008 42401 1 1 008 060 20131231 18:00 0.6
RD I 51 650 0008 42401 1 1 008 060 20131231 19:00 0.7
RD I 51 650 0008 42401 1 1 008 060 20131231 20:00 0.6
RD I 51 650 0008 42401 1 1 008 060 20131231 21:00 0.6
RD I 51 650 0008 42401 1 1 008 060 20131231 22:00 0.6
RD I 51 650 0008 42401 1 1 008 060 20131231 23:00 0.6
# 8760 records were written
```



Why is 51-650-0008 Complete?

RD	I	51	650	0008	42401	1	1	008	060	20130130	00:00	3.6
RD	I	51	650	0008	42401	1	1	008	060	20130130	01:00	4.4
RD	I	51	650	0008	42401	1	1	008	060	20130130	02:00	5.1
RD	I	51	650	0008	42401	1	1	008	060	20130130	03:00	2.4
RD	I	51	650	0008	42401	1	1	008	060	20130130	04:00	0.4
RD	I	51	650	0008	42401	1	1	008	060	20130130	05:00	2.7
RD	I	51	650	0008	42401	1	1	008	060	20130130	06:00	3.7
RD	I	51	650	0008	42401	1	1	008	060	20130130	07:00	2.3
RD	I	51	650	0008	42401	1	1	008	060	20130130	08:00	0.6
RD	I	51	650	0008	42401	1	1	008	060	20130130	09:00	0.4
RD	I	51	650	0008	42401	1	1	008	060	20130130	10:00	BC
RD	I	51	650	0008	42401	1	1	008	060	20130130	11:00	BC
RD	I	51	650	0008	42401	1	1	008	060	20130130	12:00	BC
RD	I	51	650	0008	42401	1	1	008	060	20130130	13:00	BC
RD	I	51	650	0008	42401	1	1	008	060	20130130	14:00	0.6
RD	I	51	650	0008	42401	1	1	008	060	20130130	15:00	0.2
RD	I	51	650	0008	42401	1	1	008	060	20130130	16:00	0.2
RD	I	51	650	0008	42401	1	1	008	060	20130130	17:00	0.2
RD	I	51	650	0008	42401	1	1	008	060	20130130	18:00	0.1
RD	I	51	650	0008	42401	1	1	008	060	20130130	19:00	0.2
RD	I	51	650	0008	42401	1	1	008	060	20130130	20:00	0.1
RD	I	51	650	0008	42401	1	1	008	060	20130130	21:00	0.0
RD	I	51	650	0008	42401	1	1	008	060	20130130	22:00	0.0
RD	I	51	650	0008	42401	1	1	008	060	20130130	23:00	0.1

BC	Multi-point Calibration	Null Data Qualifier
----	-------------------------	---------------------



Why is 23-011-2005 Incomplete?

(Example taken outside the region so the guilty remain anonymous)

RD	I	23	011	2005	42401	1	1	008	560	20130131	00:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	01:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	02:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	03:00	1.2
RD	I	23	011	2005	42401	1	1	008	560	20130131	04:00	1.2
RD	I	23	011	2005	42401	1	1	008	560	20130131	05:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	06:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	07:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	08:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	09:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	11:00	AI
RD	I	23	011	2005	42401	1	1	008	560	20130131	12:00	1.2
RD	I	23	011	2005	42401	1	1	008	560	20130131	13:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	14:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	15:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	16:00	1.1
RD	I	23	011	2005	42401	1	1	008	560	20130131	17:00	1.2
RD	I	23	011	2005	42401	1	1	008	560	20130131	18:00	1.2
RD	I	23	011	2005	42401	1	1	008	560	20130131	19:00	1.3
RD	I	23	011	2005	42401	1	1	008	560	20130131	20:00	1.3
RD	I	23	011	2005	42401	1	1	008	560	20130131	21:00	1.3
RD	I	23	011	2005	42401	1	1	008	560	20130131	22:00	1.2
RD	I	23	011	2005	42401	1	1	008	560	20130131	23:00	1.2

8750 records were written

AI	Insufficient Data (cannot calculate)	Null Data Qualifier
----	--------------------------------------	---------------------



Resolution for Incomplete Dataset

- STEP 1 - Convert raw data file date string from yyymmdd to yyyy/mm/dd

```
bof()  
  
while(!At_eof){  
    c(4)  
    ins_text("/")  
    c(2)  
    ins_text("/")  
    l(1)  
}
```



Resolution for Incomplete Dataset

- STEP 2 – Open file with converted date string in Excel – then convert dates in yyyy/mm/dd format to 5-digit serial date
 - For example today is 41816 based on January 1, 1900 being Day 1
 - Why do this? Much easier to program knowing a consistent sequential pattern



Resolution for Incomplete Dataset

STEP 3 – Read last two digits of Excel serial date, determine missing hours, insert blank lines where necessary

```
bof()
```

```
//Loop 4 times to cover 1 year (365 days) for year 2013 (Excel dates 41275 thru 41639), finds missing date, inserts blank line.
```

```
repeat (4) {
```

```
//Julian Day 1
```

```
#1 = 0
```

```
while (#1 <= 23) {
```

```
  c(3)
```

```
  if(cc=='7'){
```

```
    c(1)
```

```
  }
```

```
  if(cc=='5') {
```

```
    l(1)
```

```
  }else{
```

```
    bof()
```

```
    ins_newline()
```

```
  }
```

```
  #1++
```

```
}
```



Finished Product 1

Date	Hour	ppb	daily % daily %	daily count	quarter %	daily high	complete?	DV decrease
20070101	0:00	0.003					2007Q1 100%	2007H4 22
20070101	1:00	0.002					2007Q2 68%	2008H4 12
20070101	2:00	0.001					2007Q3 100%	2009H4 10
20070101	3:00	0.001					2007Q4 91%	2010H4 10
20070101	4:00	0.001					2008Q1 98%	2011H4 9
20070101	5:00	0.002					2008Q2 98%	2012H4 5
20070101	6:00	0.002					2008Q3 100%	2013H4 6
20070101	7:00	0.001					2008Q4 100%	
20070101	8:00	0.001					2009Q1 100%	2009 DV 15
20070101	9:00	0.001					2009Q2 100%	2010 DV 11 27%
20070101	10:00	0.001					2009Q3 100%	2011 DV 10 11%
20070101	11:00	0.002					2009Q4 100%	2012 DV 8 19%
20070101	12:00	0.004					2010Q1 100%	2013 DV 6 17%
20070101	13:00	0.002					2010Q2 100%	
20070101	14:00	0.002					2010Q3 100%	
20070101	15:00	0.001					2010Q4 92%	
20070101	16:00	0.002					2011Q1 100%	
20070101	17:00	0.002					2011Q2 100%	
20070101	18:00	0.001					2011Q3 99%	
20070101	19:00	0.001					2011Q4 100%	
20070101	20:00	0.001					2012Q1 100%	
20070101	21:00	0.001					2012Q2 100%	
20070101	22:00	0.001					2012Q3 99%	
20070101	23:00	0.001	100%	1		0.004	2012Q4 100%	
20070102	0:00	0.001					2013Q1 98%	
20070102	1:00	0.001					2013Q2 36%	
20070102	2:00	0.002					2013Q3 100%	
20070102	3:00	0.002					2013Q4 100%	



Finished Product 2

20080409	0:00	4	✓				✓
20080409	1:00	3	✓				✓
20080409	2:00	3	✓				✓
20080409	3:00	3	✓				✓
20080409	4:00	4	✓				✓
20080409	5:00	4	✓				✓
20080409	6:00	5	✓				✓
20080409	7:00	4	✓				✓
20080409	8:00		✓				✓
20080409	9:00		✓				✓
20080409	10:00		✓				✓
20080409	11:00		✓				✓
20080409	12:00	6	✓				✓
20080409	13:00	14	✓				✓
20080409	14:00	6	✓				✓
20080409	15:00		✓				✓
20080409	16:00		✓				✓
20080409	17:00		✓				✓
20080409	18:00		✓				✓
20080409	19:00		✓				✓
20080409	20:00		✓				✓
20080409	21:00		✓				✓
20080409	22:00		✓				✓
20080409	23:00		✓				✓

