DART Training for PAMS and CSN

(Data Analysis and Reporting Tool for Photochemical Assessment Monitoring Station and Chemical Speciation Network Data)

> Jennifer DeWinter, Steve Brown, and Hilary Hafner Sonoma Technology, Inc.

> > Beth Landis and Kevin Cavender

for



National Ambient Air Monitoring Conference

Data Validation Training Session St. Louis, MO

August 8, 2016



Outline for Training Session

- DART overview
 - Accessing and navigating DART
 - Bringing data into DART
- Using DART as part of the new CSN data review process (with live demo)
- Using DART for PAMS (and other) data validation and analysis (with live demo)
- Future enhancements
- Q&A

Data Analysis & Reporting Tool



- DART is a web-based data validation and analysis tool that allows monitoring agencies to manage, explore, validate, and flag data.
- DART includes automated screening, statistical summaries, and AQS-formatted data.

DART Features

- Manage data
 - Upload data files for validation
 - Request data from AQS
 - Convert units
 - Aggregate data

1 Add data to DART
Manage
crosstab
aqs TXO

DART Features

- Explore data
 - Create time-series plots
 - Edit data
 - Create scatter plots
 - Create bar plots



DART Features

- Validate data
 - Screen data
 - "DART Smarts"



- Export data
 - Create AQS-ready files
 - Obtain summary statistics



Accessing DART http://www.airnowtech.org



Requires an

AirNow-Tech account

Navigating DART



Here's how to get started ...

Bringing Data into DART

Three ways to add data to DART

- 1. Upload data files from your computer
 - AQS format
 - CSV format
 - TX0 format
- 2. Request data from AQS
- 3. Receive data automatically from a laboratory transfer
 - Currently available for CSN data from UC Davis (UCD)

Bringing Data into DART



Your Data in DART

When you upload your PAMS (or other) data, or make an AQS Request:

- Currently, it's unique to you
- In future, you will be able to share your data within your agency (and potentially with other AirNow-Tech users)

DART			Mar	nage Explore Va	alidate Export Help
My Data S	ets				add data 🕇
Date Received	Туре	Data Set Name	Date Range (LST)	Status	
10/23/2015	AQS	HRH test2	06/01/2014 - 08/30/2014	Ready	*
05/11/2015	AQS	My Sample Data Set	11/18/2011 - 12/10/2011	Ready	
Show 10 💌 entri	es				Previous 1 Next
		Sł	nowing 1 to 2 of 2 entries		

Your Data in DART

CSN data from UC Davis are available to approved CSN validators in your agency

Manage | Explore | Validate | Export | Help

DART

Agency data table with CSN data sets

Date Received	iype	Data Set Name	Date Range (LST)	Data Status	Download	Approval Status
06/27/2016	Lab - CSN	360010005 CSN Data	11/26/2015 - 02/06/2016	Ready for use	*	
06/27/2016	Lab - CSN	360050110 CSN Data	11/20/2015 - 02/09/2016	Ready for use	*	
06/27/2016	Lab - CSN	360290005 CSN Data	11/26/2015 - 02/06/2016	Ready for use	*	
06/27/2016	Lab - CSN	360310003 CSN Data	11/26/2015 - 02/06/2016	Ready for use	*	
06/27/2016	Lab - CSN	360551007 CSN Data	11/20/2015 - 02/09/2016	Ready for use	*	
06/27/2016	Lab - CSN	360610134 CSN Data	11/20/2015 - 02/09/2016	Ready for use	*	
06/27/2016	Lab - CSN	360810124 CSN Data	11/20/2015 - 02/09/2016	Ready for use	*	
06/27/2016	Lab - CSN	361010003 CSN Data	11/20/2015 - 02/09/2016	Ready for use	*	
Show 10 🔻 en	tries		Showing 1 to 8 of 8 entries			Previous 1 Next

/ly Data S	Sets					add data 🕇
Date Received	Туре	Data Set Name	Date Range (LST)	Data Status	Download	Delete
05/24/2016	AQS	My Sample Data Set	11/18/2011 - 12/10/2011	20%	*	×
how 10 🔻 ent	ries		Showing 1 to 1 of 1 entries			Previous 1 Next

CSN Data Review Process

- 1. UC Davis generates laboratory data; acquires mass data from continuous FEMs (as available) from AirNow-Tech.
- 2. UCD conducts data validation (details on next slide).
- 3. UCD sends the data batch to DART.
- 4. DART ingests the batch data, then automatically alerts CSN data reviewers (details on following slides) that data are available in DART.
- 5. CSN data reviewers then have 30 days to review and approve the data.
- 6. After 30 days, the data are sent back to UCD for final review and submission to AQS.

DART Data Flow for CSN

UCD sends data to DART Data Validation Contacts are informed electronically of data available for review (30-day window)

Data Validation Contacts review data and any suggested flags in DART

Data Validation Contacts approve data, or suggest additional/ different flagging, with option to include comments to UCD

UCD makes any necessary changes to the data and uploads the data to AQS

CSN Data Validators

- One agency is responsible for data validation for each site (e.g., New York DEC reviews Rochester site data).
- For each site, EPA provided a list of data validators and their associated agency.
- All listed data validators for an individual agency have access to the agency data in DART; they do not have access to data that are not for their agency.
- Once data are available in DART for review, the data validators will be contacted automatically via email.

- Once data are in DART, validators have 30 days to review and approve the CSN batch.
- DART provides summary statistics
 - flags and comments that UCD applied
 - other data characteristics (percent complete, percent above detection limit, etc.)
- In DART, validators can flag and comment on data for UCD to review, and use sortable tables to review the CSN batch.
- Time-series and bar plots are linked to the data table so validators can also graph the data.

From: Dart Email Notification [mailto:noreply@airnowtech.org]
Sent: Friday, June 24, 2016 10:33 AM
To: Angela L. Ekstrand
Subject: Current batch of CSN data expires in 14 days!

Automated Notification Via Email

Data sets

Dear DART User,

Please disregard this email if you have already completed your review.

You currently have a batch of CSN data waiting for you to approve in <u>DART</u>. These data are available until 11:59 pm on Saturday, July 23, 2016; upon expiration, the data are returned to the laboratory and submitted to AQS.

Your current batch has 2 sites.

QUEENS COLLEGE 2 (360810124) has 14 samples from 11/20/2015 00:00:00 to 12/29/2015 00:00:00.

PINNACLE STATE PARK (361010003) has 14 samples from 11/20/2015 00:00:00 to 12/29/2015 00:00:00.

"Review by'

Please email <u>CSNsupport@sonomatech.com</u> if you have questions or trouble accessing your data.

Thanks, DART Support Team





Batch Cr	eated: 07	//29/2016 - Review	By: 08/28/2	2016 23:59	9	"Dei	<i>iiouu</i> b	v" Data		
Total Samples: 6	6	Date Range	e: 01/07/2016 ·	02/06/2016		Rev	new D	y Date		
Date	Species	Total Qualifiers	Total Nul	l Codes	Data (Completeness		Data Abo	ove Detection	
01/07/2016	47	32	11		77%			32%		
01/13/2016	47	45	1		98%			34%		
01/19/2016	47	28	1		98%			40%		
01/25/2016	47	26	1		98%			45%		
01/31/2016	47	30	1		98%			36%		
02/06/2016	47	25	1		98%			47%		
Batch Data Filter:	Flag	as reviewed								
Reviewed Date		arameter	POC	Value	MDL	Uncertainty	Unit	Null Code	Qualifier Code	Comments
01/07	7/7016 To	otal Nitrate PM2.5 LC	5	-999.0	0.0	0.0	ug/m3	AM	ľ	
01/07	//2016 Va	anadium PM2.5 LC	5	-1.2E-4	0.00128	7.8E-4	ug/m3		MD	
1/07	7/2016 Zi	nc PM2.5 LC	5	0.00162	0.00326	0.00199	ug/m3		MD	
01/07	7/2016 Zi	rconium PM2.5 LC	5	-0.00198	0.01561	0.0095	ug/m3		MD	
01/13	/2016 Zi	rconium PM2.5 LC	5	-0.01687	0.01563	0.01039	ug/m3	Ø	5, MD	
01/13	/2016 Zi	nc PM2.5 LC	5	0.00189	0.00327	0.002	ug/m3	Ø	5, MD	

Save

al Samples:	6	Date Range	e: 01/07/2016 - 02/06/2016						
ite	Species	Total Qualifiers	Total Null Codes	"Ou	tlier for XRF-IC C	omparison"			
/07/2016	47	32	11	Add	9/2010 10.29 01C				
13/2016	47	45	1	9	comment.				
Edit Bat	ch	20	•						
Recent Co	mment:	Apply	null or	9					
Comment	has not been addeo	^{d yet.} qualifi	er codes	9				Change	es and
Edit Null C	ode:								
No null co	de		•					comme	ents log
B No null co	ide		•					comme	ents log
Edit Qualif	ier Code:		•		lineerteinte		Null Code	comme	ents log
Edit Qualif	ier Code: ilue less than MDL		•	MDL 0.0	Uncertainty 0.0	Unit ug/m3	Null Code	Qualifier Code	Communits
No null co Edit Qualifi XMD - Va Apply to: Selected S	ier Code: alue less than MDL	•	•	MDL 0.0	Uncertainty 0.0	Unit ug/m3	Null Code AM	Qualifier Code	Comments
Edit Qualifi MD - Ve Apply to: Selected S Edit Comm	ier Code: alue less than MDL Species nent:	T	•	MDL 0.0 0.00128	Uncertainty 0.0 7.8E-4	Unit ug/m3 ug/m3	Null Code AM	Qualifier Code	Combents log
B No null cc Edit Qualif MD - Va Apply to: Selected S Edit Comm	ier Code: alue less than MDL species nent:	T	·	MDL 0.0 0.00128 0.00326	Uncertainty 0.0 7.8E-4 0.00199	Unit ug/m3 ug/m3 ug/m3	Null Code AM Ø	Qualifier Code	Comments Comments
Apply to: Edit Comm	ier Code: alue less than MDL Species nent:	¥		MDL 0.0 0.00128 0.00326 0.01561	Uncertainty 0.0 7.8E-4 0.00199 0.0095	Unit ug/m3 ug/m3 ug/m3 ug/m3	Null Code AM C	Qualifier Code	Combents log
No null cc Edit Qualifi MD - Va Apply to: Selected S Edit Comm	ier Code: alue less than MDL species nent:	T	•	MDL 0.0 0.00128 0.00326 0.01561 0.01563	Uncertainty 0.0 7.8E-4 0.00199 0.0095 0.01039	Unit ug/m3 ug/m3 ug/m3 ug/m3	Null Code AM Ø	Comme Qualifier Code MD MD MD 5, MD	Completes log







360010005 CSN Data	ROCHESTER 2			🖨 Settings 🛛 🕍 Gra
360050110 CSN Data	Select All	Select All CSN Species	Sort by CSN	
	Filter by Nylon	Filter by Quartz	Eilter by Teflon	
360290005 CSN Data	CC PM2.5 LC Tor	Sulfate PM2.5 LC	Total Nitrate PM2.5 LC	Ammonium Ion PM2.5 LC
60310003 CSN Data	EC PM2.5 LC Tor	Sulfur PM2.5 LC	Sodium Ion PM2.5 LC	Silicon PM2.5 LC
360551007 CSN Data	Sodium PM2.5 LC	✓ Iron PM2.5 LC	Potassium PM2.5 LC	Potassium Ion PM2.5 LC
	Calcium PM2.5 LC	Chlorine PM2.5 LC	Aluminum PM2.5 LC	Chloride PM2.5 LC
360610134 CSN Data	Zinc PM2.5 LC	Magnesium PM2.5 LC	Antimony PM2.5 LC	Chromium PM2.5 LC
360810124 CSN Data	Copper PM2.5 LC	Bromine PM2.5 LC	✓ Titanium PM2.5 LC	Tin PM2.5 LC
361010003 CSN Data	Manganese PM2.5 LC	Lead PM2.5 LC	☑ Indium PM2.5 LC	Cadmium PM2.5 LC
	Barium PM2.5 LC	Cesium PM2.5 LC	Vickel PM2.5 LC	Silver PM2.5 LC
v Data Sets	Zirconium PM2.5 LC	Arsenic PM2.5 LC	Vanadium PM2.5 LC	Strontium PM2.5 LC
م Vy Sample Data Set	Cobalt PM2.5 LC	Phosphorus PM2.5 LC	Selenium PM2.5 LC	Cerium PM2.5 LC
	Rubidium PM2.5 LC	Ammonium Nitrate PM2.5 LC	C Ammonium Sulfate PM2.5	Elements
Date Range	lons	Organic Carbon Mass PM2.5	PM2.5 Raw Data	Reconstructed Mass PM2.5

Transition To PAMS Training

Time-Series Graphs



Scatter Plots



Bar Plots



Stacked Bar Charts



Auto-Validation: Screening

Check	Fails If	DART Smarts Action If Check Fails
Abundant Species	Any of Benzene , Propane , N-Butane , Isoprene , N-Hexane , Ethylbenzene are missing or 0	If two or more species missing or = 0, flag sample with code "AQ"
TNMOC	-TNMOC is missing or 0; or -Unidentified exceeds 50% of TNMOC; or -Sum of PAMS exceeds TNMOC	-Flag TNMOC and Unidentified with code "AN" -Flag Unidentified with code "DA" -Flag TNMOC and Sum of PAMS with code "DA"
Variability	Species concentration exceeds the mean + 4*standard deviation	None
Sticking	Species has same non-zero value for 3 or more consecutive samples	Flag species with code "DA"
Benzene : Toluene	Benzene exceeds 0.2 and exceeds Toluene	Flag Benzene and Toluene with code "DA"
Ethylene : Ethane	Ethylene exceeds 0.5 and exceeds Ethane	Flag Ethylene and Ethane with code "DA"
Propylene : Propane	Propylene exceeds 0.5 and exceeds Propane	Flag Propylene and Propane with code "DA"

All checks done in ppbC. AQ = collection error; AN = machine malfunction; DA = aberrant data; BH = interference/coelution/misidentification.

Auto-Validation: Screening

Check	Fails If	DART Smarts Action If Check Fails
O-Xylene : M/P Xylene	O-Xylene exceeds 0.5 and exceeds M/P Xylene	Flag Xylenes with code "DA"
Methylpentanes	3-Methylpentane exceeds 0.1 and exceeds 0.6*2-Methylpentane	If 3-Methylpentane exceeds 0.1 and exceeds 0.65*2-Methylpentane , then flag Methylpentanes with code "BH"
Undecane : Decane	N-Undecane exceeds 0.5 and exceeds N-Decane	Flag N-Undecane and N-Decane with code "DA"
Olefins : Paraffins	Sum of Olefins exceeds Sum of Paraffins	Flag Olefins and Paraffins with code "DA"
Carbon Tetrachloride	Carbon Tetrachloride exceeds 0.16 ppb	Flag Carbon Tetrachloride with code "AQ"
Formaldehyde	Formaldehyde is less than 0.3 ppb	None
Nighttime Isoprene	Isoprene increases between 8 pm and 3 am local time	Flag Isoprene with code "DA"

All checks done in ppbC, except carbon tetrachloride and formaldehyde checks, which use ppb.. AQ = collection error; AN = machine malfunction; DA = aberrant data; BH = interference/coelution/misidentification.

	DART		Manage Exp	plore Validate Export Help
	VALIDATION SESSIONS			
	SampleData T My Sa	et: Site: mple Data LIVERMOF	RE - RINCON	Load
	SETUP NEW SESSION			
1	Select Screening Level			
	Basic Level 0-1	Intermediate Level 2 Coming soon		Advanced Level 3 Coming soon
2	Select Data		(Set up custom screening checks
	Select a task group	Please select a task group	<pre>configure</pre>	
	Select data set	User File 1	•	
	Select site from User File 1	SACRAMENTO - DEL PASO MANOR	▼	

1-Pentene 2,2,4-Trimethylpentane

SampleData	■ Data set: ■ My Sample Data	Site: LIVERMORE - RINCON	Load
Create New Tas	k Group		
Task group name Duration	PAMS Basic		
1. Check typ	e t Species 🔻	Check nickname Abundant Species	Apply null code?
Fails when Benzene N-Butane	Ethylbenzene Isoprene	is missing or	Applies to whole sample for a failed hour
2. Check type TNMOC	•	Check nickname TNMOC	Apply null code?
тимос тимос * тимос	= ▼ 0 0.5 <		
3. 🖌 Check type Variability	2 /	Check nickname Variability	Apply null code?
Fails when 1,2,3-Trir 1,2,4-Trir 1,3,5-Trir	nethylbenzene nethylbenzene nethylbenzene	average + 0 * stddev.	

DART					Ν	Man	age E	xplore	Validate	Export	Help			
Summary	test - PAMS B	asic											Summ	nary
1 Abundant Species	Data Set	Site			Total Samples		Passes	Missing	Failures	Null Codes	Date	Range		
	My Sample Data Set	LIVERMORE - RINCON			552 0		376	420	455	11/18/2011 - 12/10/2011				
2 TNMOC	 Summary 													
3 Variability	Date Time (LST) 🔺	1	2	3	4	5	6	7	8	9	10	11	12	
-	11/18/2011 00:00	-	0	0	•	A	0	0	0	0	0	A	A	*
4 Benzene: Toluene	11/18/2011 01:00		-						P			0	0	-
	11/18/2011 02:00		0	0					 			0	0	
5 Ethylene:Ethane	11/18/2011 03:00		0						-			0	•	
	11/18/2011 04:00	×	-	0					-			A	A	
6 Propylene:Propane	11/18/2011 05:00		P						-			•		
7 O. Yvlene: M/P. Yvlene	11/18/2011 06:00		-	0				-				0	0	
	11/18/2011 07:00		P					F	-			0	•	
8 Methylpentanes	11/18/2011 08:00	×	-	0					P			•	A	
	11/18/2011 09:00	×	-						P			•	•	
9 Undecane:Decane	11/18/2011 10:00		-	۵			11		 			۵	۵	



Getting Help

Search

DART

Contents

1. About DART

- 2. The Importance of Data Validation
- 3. Getting Started
- 4. System Overview
- 🛄 4.1 Manage
- 4.1.1 Supported File Formats
- 4.1.2 Agency Data Sets from a Labo
- 4.2 Explore
- 4.3 Validate
- 4.4 Export
- 5. More Information
- Acknowledgments

•

Glossary and Acronyms

111

1×

(Previous Topic: 4.1.1 Supported File Formats)

4.1.2 Agency Data Sets from a Laboratory

In addition to ingesting file uploads and AQS requests, DART can ingest data directly from an air quality laboratory. Data from a laboratory can be provided to DART via File Transfer Protocol (FTP) and are automatically made available in the correct DART user accounts.

Currently, PM_{2.5} speciation data collected as part of the Chemical Speciation Network (CSN) program are transferred to DART from the Crocker Nuclear Laboratory at the University of California, Davis. CSN data are listed in the **Agency** table on the **Manage** screen, in the DART user accounts that are registered to the appropriate agency. Data are received from the laboratory in batches and are available for review and validation using the DART **Approval Mode** screen.

		AirNow				Weitume, Thereas Olifieri I My Account I Cantactilia I Log Out				
	DART				Ma					
		New York	Dept. of E	Environmental (Conservation Data Sets	Factors				
		03/30/20%	Lab - CSN	368018085 CSN Data	11/26/2015 - 11/26/2015	Ready	 A 	Click the icon to	•	
View the date of the most recent batch from the laboratory.	e date of 🔨	03/30/2016	Lab - CSN	368058110 C SN Date	11/20/2015 - 11/29/2015	Ready		enter Approval	roval	
	st recent	03/30/2016	Lab - CSN	368298085 CSN Data	11/26/2015 - 11/26/2015	Ready	▲ ▲	Mode and review		
	ratory.	03/30/20%5	Lab - CSN	368318083 CSN Data	11/26/2015 - 11/26/2016	Ready		laboratory.	/	
\sim		03/30/20%5	Lab - CSN	368551087 CSN Data	11/20/2015 - 11/29/2015	Ready	A			
		03/30/20%6	Lab - CSN	368619134 CSN Data	11/20/2015 - 11/29/2016	Ready	A			
		03/30/2016	Lab - CSN	368919124 CSN Data	11/20/2015 - 11/29/2015	Ready	A			
		03/30/20%5	Lab - CSN	369918083 CSN Data	11/20/2015 - 11/29/2015	Ready	■ ▲ ✓			
		Show 10 • entr		/	Showing 1 to 8 of 8 entries		Previous 1 No.	Note the green	١	
			View name; include mon	the data se each data s s data for o itoring site	et set one			data that do not require review.	/	

Getting Help

Feedback button (Feedback





Important CSN Contacts

CSNsupport@sonomatech.com

Role	Contact	Phone Number	Email
EPA Project Officer	Jeff Yane	919-541-2962	yane.jeff@epa.gov
EPA Project Manager	Elizabeth Landis	919-541-2262	landis.elizabeth@epa.gov
EPA QA Officer	Jenia Tufts	919-541-0371	tufts.jenia@epa.gov
Shipping & Handling (Amec)	Justin Knoll	352-333-6621	justin.knoll@amecfw.com
Gravimetric Analysis (Amec)	Bill Barnard	352-333-6617	bill.barnard@amecfw.com
Laboratory Analysis (UCD)	Nicole Hyslop	530-754-8979	nmhyslop@ucdavis.edu
DART (STI)	Jennifer DeWinter Steve Brown	707-665-9900	jdewinter@sonomatech.com steveb@sonomatech.com

CSN Regional Representatives

- R1 Alan VanArsdale & Catie Taylor
- R2 Mazeeda Khan
- R3 Lori Hyden
- R4 Keith Harris
- R5 Scott Hamilton

- R6 Frances Verhalen
- R7 Leland Grooms
- R8 Joshua Rickard
- R9 Anna Mebust & Dena Vallano
- R10 Chris Hall & Keith Rose

Potential New DART Features

- Interactive map for Data Mart AQS requests
- Suite of automated screening checks for air toxics, CSN data
- Compare site data to national statistics
- New analyses and plot types



Potential New DART Features

- Plot concentrations with MDL values
- Plot concentrations with annual averages
- Support for more import file formats
- Connection to AirNow-Tech features
- Pollution and wind roses



Potential CSN Features in DART

- Plot concentrations with MDL values
- Plot concentrations with annual averages
- Compare site data to national statistics
- Stacked bar and scatter plots in approval mode
- Pollution and wind roses

Acknowledgments

- Joann Rice, CSN Program Support, EPA
- National Association of Clean Air Agencies Steering Committee
- Nick Mangus and Robert Coats, AQS Support, EPA

Contact Us

Jennifer DeWinter

DART Project Manager (STI) jdewinter@sonomatech.com

Steve Brown

Senior Atmospheric Scientist sbrown@sonomatech.com



Kevin Cavender

EPA PAMS Lead cavender.kevin@epa.gov

Beth Landis

EPA CSN Lead landis.elizabeth@epa.gov

