Developing Data Management Policy and Guidance Documents for your NARSTO Program or Project

---------- A different approach to developing a data management plan in the NARSTO context.

Following is a compilation of data management policy and guidance documents for program and project use in developing data management plans. Documents can be downloaded and implemented individually or as a set, depending upon your data management needs. Please be advised that this guidance and the referenced resources will be periodically updated and that users should visit the QSSC web site (link below) for the latest versions.

Getting started –

- **Select the data management guidance documents needed in your Program or Project from the table of model documents that follows.**
- **Adopt, adapt, or refine these model documents as appropriate for your needs with input from managers, investigators, modelers, data coordinators, etc.**
- **Consult with the NARSTO QSSC for more information and assistance.**
- **Distribute the approved documents to participants to inform them of their data collection and reporting responsibilities.**
- **Ensure that adequate data coordination support is provided to all participants to facilitate implementing the plans.**

Prepared by the NARSTO Quality Systems Science Center (QSSC)
http://cdiac.ornl.gov/programs/NARSTO/

Les A. Hook and Sigurd W. Christensen,
NARSTO Quality Systems Science Center
Environmental Sciences Division
Oak Ridge National Laboratory

Contact: Les Hook, hookla@ornl.gov , 865-241-4846

ORNL research was sponsored by the U.S. Department of Energy and performed at Oak Ridge National Laboratory (ORNL). ORNL is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

Overview of Data Policy and Management Plan Development

Rationale: Providing this information to ***Project*** participants will inform them of their data reporting responsibilities, promote consistency and standardization in data and metadata collection and reporting processes, and greatly facilitate data sharing, integration, synthesis, and analysis. Guidance should be consistent with the needs of the ***Project.***

Target Audience: The audience for these guidance documents is the investigators, experimentalists, modelers, and data coordinators responsible for generating and submitting data to a ***Project*** database, creating other data products, and archiving these data.

Guidance Documents: Each document should be 1-2 pages in length (plus attachments) and contain information that has been reviewed in light of your ***Project*** data management needs. Guidance in the model DM documents incorporates existing NARSTO data management protocols and will often be suitable for use as is. Final guidance should be consistent with the needs of the ***Project*** within the NARSTO context. **Add additional project-specific guidance as needed.**

Document Development Process: Ideally, the ***Project*** data coordinator will take the lead on selecting the needed DM documents, coordinating the project review, and modifying the guidance documents. The provided model DM documents are in MSWord format and may be copied and edited as needed. Please contact the QSSC if you have any problems with the DM documents or have questions about the DM NARSTO guidance.

Authority: Each guidance document should be approved by ***Project*** management to ensure acceptance and implementation.

Distribution: Ideally these will be web documents and would include links to on-line ***Project*** documents (e.g., DM-4, Site ID table) and NARSTO QSSC resources (e.g., variable name reference tables and DES format template) at http://cdiac.ornl.gov/programs/NARSTO. Hardcopies could be provided as needed.

Proposed Project Data Management Policy / Guidance Documents

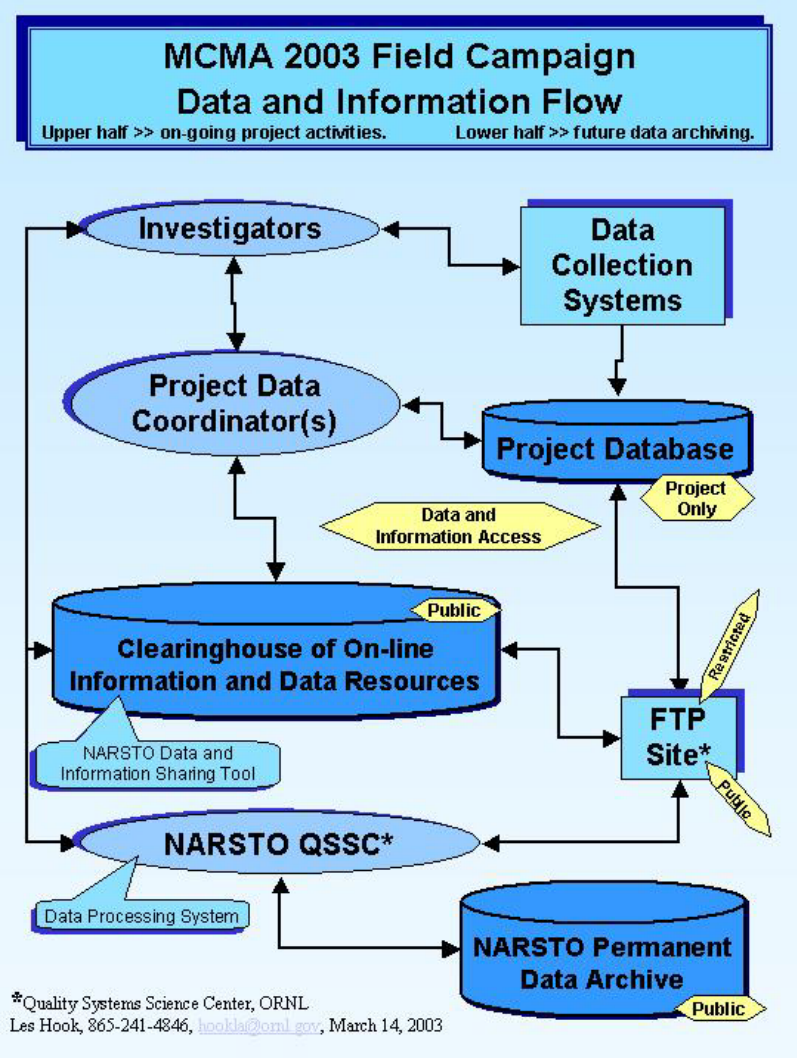
Data Management Policy / Guidance Documents		Status / Contact	Approved by / Date (yyyy/mm/dd)
> Organization			
DM-1	Data Flow Overview		
DM-2	Data Policy Considerations		
DM-3	Project Name Information		
DM-4	Identifying Measurement and Sampling Sites		
> Data and Metadata Reporting			
DM-5	Reporting Sampling and Measurement Dates and Times		
DM-6	Identifying Chemical and Physical Variables and Descriptive Field Information		
DM-7	Reporting Units for Chemical Variables, Particles, and Physical and Descriptive Variables		
DM-8	Assigning Project-Specific and NARSTO Data Quality Flags		
DM-9	Reporting and Flagging Values below Detection Limits		
DM-10	Reporting Missing Data		
DM-11	Reporting Uncertainty Estimates		
DM-12	Reporting Conventions for Mass Measurements, Meteorological Data, and Temperature and Pressure Conditions		
> Data Documentation and Archiving			
DM-13	Planning to Archive Data		
DM-14	Creating Archive Documentation for Your Data Sets		
DM-15	Creating a Searchable Index of Your Data Sets with Links to the Data Files		
DM-16	Capturing Sampling and Analysis Information – Pre- and Post-Measurement		
DM-17	Defining the Quality Level of Data		
> Data Systems Management			
DM-18	Day-to-Day Operation of Data Management Systems		
DM-19	Managing Electronic and Hardcopy Format Project Records		
DM-20	Data Management System and Software Configuration Control Guidelines		

DM-1: Data Flow Overview

BACK TO TABLE

SCOPE: ***Project (MCMA 2003 example)***

PURPOSE: To inform investigators and potential data users of the general flow of data and information before, during, and after the current field campaign. Data collected by investigators will be provided to the MCMA database to meet project data analysis needs. Certain data and metadata reporting standards are necessary (e.g., DM-6, Variable naming) to facilitate efficient data reporting, processing and analysis. Data will ultimately be sent to the NARSTO Permanent Data Archive (PDA). Our reporting standards are consistent with those for the NARSTO PDA.



Discussion:
The information is a general guide to carry out this process. Some larger projects have onsite Data Managers who work with both the Principal Investigators and the NARSTO QSSC. Other smaller projects do not have Data Managers, and the PIs interact directly with the QSSC. While projects may have varying assigned roles and responsibilities for data management, the QSSC is the source for information and assistance with data, metadata, and archiving activities.

DM-2: Data Policy Considerations

BACK TO TABLE

SCOPE: *Project*

PURPOSE: To involve all project managers and participants, as well as potential data users in the formulation of a data policy. A clear statement of the importance of the data collection effort and of the flow of the data and information before, during, and after the current activities in the broadest possible context is needed. It is a shared responsibility of all participants to implement the data policy.

Vision:

Is it safe to assume that data and metadata will be shared among Project investigators, and ultimately made available to the public in a timely manner through an archive facility?

Who do you consider to be the audience for data beyond the Project team?

Will there be a Project data integration or synthesis effort in the future?

Do you see the value of the data as being short-term (3-5 years), mid-term (10 years), or longer (20 years)?

Are these considerations the same for field measurement data, laboratory data, and modeling products (input data, model code, and output results)?

Compliance with (as may be applicable):

- U.S. Government OMB CIRCULAR A 110, (REVISED 11/19/93, As Further Amended 9/30/99) [http://www.whitehouse.gov/OMB/circulars/a110/a110.html#72]

- U.S. Government Agency implementations of "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies," OMB, 2002. (67 FR 8452) [http://www.whitehouse.gov/omb/fedreg/reproducible2.pdf]

*** Example vision statement: The atmospheric sciences community is experiencing an unprecedented increase in the types and amount of data being collected, modeled, and assessed. As projects evolve to more focused, multi-investigator, interdisciplinary efforts in a period of limited resources, the timely availability and sharing of data and documentation among participants becomes increasingly important. The need for the use of this information beyond the project for climate assessments and air quality management decisions has never been greater thus placing the additional responsibility on the project of providing for the timely submission of quality controlled data to national data centers for wider public use. ***

Timeliness of Data Availability:

Considerations for timing of field measurement, laboratory, and modeling activities?

Considerations for timing of laboratory results feeding modeling projects?

Rapid turn around of draft data within the Project? Justification?

Will data that are the subject of student theses or dissertations need special consideration?

Will investigators be expected to maintain or archive raw data for specified periods of time?

History tells us enforcement of data policies requires direct involvement by the Program Manager (i.e., threat of no funding for non-compliance)

Quality Assurance:

Will each investigation develop a QA project plan? Will the Program have an overarching QA plan? A final investigation QA summary report?

What level of QA is desirable for data to be shared within project? With the public?

Flagging data?
Encourage reporting of uncertainty measures with data values?
Detection limits?
Reporting of instrument calibrations and intercomparisons?
Will common data-processing protocols be used (e.g., gap-filling, block averaging, standard software packages to convert voltages to concentrations)?

Data and Metadata Reporting:
Investigators have an obligation to make their data easy to use by others?
The Project will develop or adapt (e.g., from the QSSC) a formal description of preferred conventions?
Consider extending use of uniform metadata reporting conventions beyond date and time to include site names, parameter names, CAS RNs, units, methods, missing values codes, quality flagging, etc.
Consider that searchable, standardized metadata improves synthesis and integration efforts.

Data Archive:
Considerations for archiving: long-term system stability and longevity?
Consider types and amount of documentation for long-term data archiving – "twenty year test".

- Scientists are encouraged to document their data at a level sufficient to satisfy the well-known "20-year test". That is, someone 20 years from now, not familiar with the data or how they were obtained, should be able to find data of interest and then fully understand and use the data solely with the aid of the documentation archived with the data.(National Research Council, Committee

on Geophysical Data, Solving the Global Change Puzzle, A U.S. Strategy for Managing Data and Information, National Academy Press, Washington, D.C., 1991.)

Consider project maintenance and retention of raw/minimally processed instrument data, software codes used for data processing, model code with input data and output products, and hardcopy records.

Data Ownership/Control:

- The issue of data "ownership" is a difficult one.
  - On the one hand a system must allow an instrument operator to reap the rewards of their efforts.
  - On the other hand the common good is served by sharing.
- The metadata should clearly state source of data, whether data are preliminary and for use only among the project or suitable for widespread dissemination and citation requirements.
- At some point there is a legal obligation for data collected with government funds to be freely available.
- A decision is needed as to when the data sets are freely available to the outside community.
- Conflict resolution?

Protection of Intellectual Property Rights:

- How will the Project help to ensure that intellectual property rights are protected and co-authorship, acknowledgement, or credit is given to data originators and principal investigators?

Consider the use of data in synthesis and integration studies that result in derived and value-added products.

Example statement:

- When data are required for modeling or integrating studies, the originator of the data should be consulted before data or derived products are incorporated or published in a review or integrated study. The scientist collecting such data shall be credited appropriately by either co-authorship or citation. (SAFARI 2000 DATA POLICY, February 5, 2001, http://mercury.ornl.gov/safari2k/s2kpolicy.pdf])

Example statement: AmeriFlux Data Fair-Use Policy

- The AmeriFlux data provided on this site are freely available and were furnished by individual AmeriFlux scientists who encourage their use. Please kindly inform the appropriate AmeriFlux scientist(s) of how you are using the data and of any publication plans. Please acknowledge the data source as a citation or in the acknowledgments if the data are not yet published. If the AmeriFlux Principal Investigators (PIs) feel that they should be acknowledged or offered participation as authors, they will let you know and we assume that an agreement on such matters will be reached before publishing and/or use of the data for publication. If your work directly competes with the PI's analysis they may ask that they have the opportunity to submit a manuscript before you submit one that uses unpublished data. In addition, when publishing, please acknowledge the agency that supported the research. Lastly, we kindly request that those

publishing papers using AmeriFlux data provide preprints to the PIs providing the data and to the data archive at the Carbon Dioxide Information Analysis Center (CDIAC).
[http://public.ornl.gov/ameriflux/data-fair-use.shtml]

DM-3: Project Name Information

BACK TO TABLE

SCOPE: ***Project (MCMA 2003 example)***

PURPOSE: Provide standard names to identify the project, sampling sites, data files, data sets, and FTP site area. Resources, examples, and use in the NARSTO Data Exchange Standard (DES) template are shown.

MCMA Names

Study or Network Short Acronym (Starts with a letter. Use in site names, columns 1 - 4)
MCM3
Resource: DM-4 : Identifying fixed measurement sites and mobile measurement platforms

***STUDY OR NETWORK ACRONYM** (Use in data file and data set names, chars 1-15)	***STUDY OR NETWORK NAME**
MCMA_2003	Mexico City Metropolitan Area 2003 Field Campaign
Resource: Data Exchange Standard Template	

***ORGANIZATION ACRONYM**	***ORGANIZATION NAME:**
MIT_IPURGAP	Massachusetts Institute of Technology Integrated Program on Urban, Regional, and Global Air Pollution
Others?	
Resource: Data Exchange Standard Template	

Shared-Access FTP Site Information	
Item	**Project Info**
UID	**mcma** (lower case)
Password	**xxxxxxxx** (case sensitive)
Internal/ directory name	**mcma2003** (lower case)
Resource: [http://cdiac.ornl.gov/programs/NARSTO/sharedaccess.htm]	

Data File and Data Set Naming	Limits
Data File: [*STUDY OR NETWORK ACRONYM]_[unique data file descriptors]_V1.csv Example: MCMA_2003_SMPS_WHERE_WHEN_V1.csv Projects should define a standard syntax for the [unique data file descriptors] portion of the data file name.	57 chars max, uppercase (except .csv)
Data Set Title: NARSTO [*STUDY_OR_NETWORK_ACRONYM] [Data Description] Example: NARSTO MCMA_2003 Scanning Mobility Particle Size Data	80 chars max, title case
Data Set Name: NARSTO_[STUDY_OR_NETWORK_ACRONYM]_[Abbreviated_Data_Description] Example: NARSTO_MCMA_2003_SMPS_DATA	40 char max, uppercase
Resource: http://cdiac.esd.ornl.gov/programs/NARSTO/pdf/archiving.pdf	

Use of Project Name Information in DES Template (See large bolded cells)

*DATA EXCHANGE STANDARD VERSION	NARSTO 2005/04/29 (2.302)
*COMMENT	Further instructions on how to fill in this template are provided at:
*COMMENT	http://cdiac.ornl.gov/programs/NARSTO/narsto.html
*QUALITY CONTROL LEVEL	
*DATE THIS FILE GENERATED/ARCHIVE VERSION NUMBER	2004/06/30
***ORGANIZATION ACRONYM**	
***ORGANIZATION NAME**	
***STUDY OR NETWORK ACRONYM**	
***STUDY OR NETWORK NAME**	
*FILE CONTENTS DESCRIPTION--SHORT/LONG	
*PRINCIPAL INVESTIGATOR NAME--LAST/FIRST	
*PRINCIPAL INVESTIGATOR AFFILIATION	
*CO-INVESTIGATOR NAME--LAST/FIRST	
*CO-INVESTIGATOR AFFILIATION	
*COUNTRY CODE	
*STATE OR PROVINCE CODE	
*SAMPLING INTERVAL AS REPORTED IN MAIN TABLE	
*SAMPLING FREQUENCY OF DATA IN MAIN TABLE	
*PRINCIPAL INVESTIGATOR CONTACT INFORMATION	
*DATA USAGE ACKNOWLEDGEMENT	
*NAME AND AFFILIATION OF PERSON WHO GENERATED THIS FILE	
*DATE OF LAST MODIFICATION TO DATA IN MAIN TABLE	
*FILE CHANGE HISTORY--VERSION NUMBER/DESCRIPTION	
*NAME AND VERSION OF SOFTWARE USED TO CREATE THIS FILE	
*STANDARD CHARACTERS	!#$%&'()*,- +=[]|./0123456789:;<>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ\^_`abcdefghijklmnopqrstuvwxyz{}~
*COMPANION FILE NAME/FORMAT AND VERSION	

DM-4: Identifying Measurement and Sampling Sites BACK TO TABLE

SCOPE: *Project*

PURPOSE: Provides a standard for identifying and characterizing fixed measurement locations and mobile measurement platforms used by the project as measurement and sampling sites. Specifications, resources, examples, and use in the DES template are shown.

NARSTO Standard Site Identifier

The NARSTO Standard Site Identifier is constructed as follows for both fixed and mobile sites.

Columns	Contents
1 - 4	Study or network acronym (see DM-3), beginning with a letter
5 - 6	Country code (following the ISO3166 Standard)
7 - 8	State or Province
9 - 12	Site abbreviation (site mnemonic, 1 – 4 chars), beginning with a letter

Limits: The full 12 columns must be used, and no blanks are permitted. The last character of the site identifier can be repeated to avoid blanks, or underscore (_) character(s) can be used instead of a blank.

Resource: Site Identifier Consensus Metadata Standard [http://cdiac.esd.ornl.gov/programs/NARSTO/metadatastandards/consensus_site_id_standard.txt]
Data Exchange Standard Template for country and state codes

Examples:

Fixed Site:
,SS99USTNBNA_,BNA_,BNA - NASHVILLE INTL AIRPORT,US (UNITED STATES)
,TN,36.1244767,-86.6781822, …

Mobile Site: (mobile platform is based at fixed site)
,SS99USTNG1PN,G1PN,Grumman G-1,US (UNITED STATES),TN,-999.99999,-999.99999, …

Fixed Site:
,ES2HUSTXEFD_,EFD_,EFD - ELLINGTON FIELD AIRPORT,US (UNITED STATES)
,TX,29.607333,-95.158750, …

Mobile Site: (same mobile platform is based at different fixed site)
,ES2HUSTXG1PN,G1PN,Grumman G-1,US (UNITED STATES),TX,-999.999999,-999.999999,
…

Project Master List of Site Information

The project should maintain a master list of site identifiers, characteristics, and other available information. Some items have picklists in the NARSTO Data Exchange Standard template which also serve to explain the meaning of the item.

Key Information Needed to Adequately Characterize Measurement and Sampling Sites

*TABLE COLUMN NAME	FIELD REQUIRED OR OPTIONAL	*TABLE COLUMN UNITS	*TABLE COLUMN FORMAT TYPE	*TABLE COLUMN FORMAT FOR DISPLAY	*TABLE COLUMN MISSING CODE	PICKLIST AVAILABLE IN THE DES TEMPLATE
Site ID: standard	REQUIRED	None	Char	12	None	
Site abbreviation: standard	OPTIONAL	None	Char	4	None	
Description	OPTIONAL	None	Char	50	None	
Country code	REQUIRED	None	Char	50	None	Yes
State or province code	REQUIRED	None	Char	20	None	Yes
Latitude: decimal degrees	REQUIRED	Decimal degrees	Decimal	10.5	-999.99999	
Longitude: decimal degrees	REQUIRED	Decimal degrees	Decimal	10.5	-999.99999	
Lat/lon reference datum	REQUIRED	None	Char	120	None	Yes
Sampling height above ground	OPTIONAL	m (meter)	Decimal	6.1	-99999.9	
Ground elevation: above mean sea level	REQUIRED	m (meter)	Decimal	6.1	-99.9	
Pressure: site ground level	OPTIONAL	hPa (hectopascal)	Decimal	7.2	-99.99	
Site land use	REQUIRED	None	Char	30	None	Yes
Site location setting	REQUIRED	None	Char	40	None	Yes
Measurement start date at site	REQUIRED	yyyy/mm/dd	Date	10	9999/12/31	
Measurement end date at site	REQUIRED	yyyy/mm/dd	Date	10	9999/12/31	
Co-incident measurements at site	OPTIONAL	None	Char	200	None	
Site ID: study	OPTIONAL	None	Char	12	None	
Lat/lon accuracy	OPTIONAL	m (meter)	Decimal	7.1	-999.9	
Lat/lon method	OPTIONAL	None	Char	50	None	Yes
AIRS ID	REQUIRED (conditional)	None	Char	9	None	
City/town	OPTIONAL	None	Char	25	None	
County	OPTIONAL	None	Char	25	None	
WMO region	OPTIONAL	None	Char	1	None	
WDCA/GAW station type	OPTIONAL	None	Char	1	None	
GAWID	OPTIONAL	None	Char	6	None	
Comment	OPTIONAL	None	Char	120	None	
Site nature (fixed or mobile site)	REQUIRED (conditional)	None	Char	6	None	Yes
Site location type	OPTIONAL	None	Char	20	None	Yes
Site: start date	OPTIONAL	yyyy/mm/dd	Date	10	9999/12/31	
Site: end date	OPTIONAL	yyyy/mm/dd	Date	10	9999/12/31	
Site: study start date	OPTIONAL	yyyy/mm/dd	Date	10	9999/12/31	
Site: study end date	OPTIONAL	yyyy/mm/dd	Date	10	9999/12/31	
Site population class	OPTIONAL	None	Char	30	None	Yes
Site type	OPTIONAL	None	Char	20	None	Yes
Site monitoring support	OPTIONAL	None	Char	30	None	Yes

Site topography	OPTIONAL	None	Char	15	None	Yes
Site monitoring duration	OPTIONAL	None	Char	50	None	

Project Site Information Template

There is a Project Site Information Template available to facilitate gathering the required and optional site information for the Master List. *Table Column Name row cells have embedded comments that describe the sought after information. The Excel template has several picklists to help ensure consistency of entered values. Additional values can be added to the lists. [http://cdiac.esd.ornl.gov/programs/NARSTO/NARSTO_template_atmospheric_measurements.xls]

The information collected in this site information template will be used when submitting measurement data with the DES Template to the NARSTO archive.

Use of Project Site Identification Information in DES Template

***TABLE NAME**	**Site information**							
*TABLE FOCUS	Metadata							
*TABLE COLUMN NAME	**Site ID: standard**	**Site abbreviation: standard**	**Description**	**Country code**	**State or province code**	**Latitude: decimal degrees**	**Longitude: decimal degrees**	**Additional site info in adjacent columns**
*TABLE COLUMN UNITS	None	None	None	None	None	Decimal degrees	Decimal degrees	
*TABLE COLUMN FORMAT TYPE	Char	Char	Char	Char	Char	Decimal	Decimal	
*TABLE COLUMN FORMAT FOR DISPLAY	12	4	50	50	20	10.5	10.5	**See Key Info Table above**
*TABLE COLUMN MISSING CODE	None	None	None	None	None	-999.99999	-999.99999	
*TABLE BEGINS								
	SS99USTNBNA_	BNA_	BNA - NASHVILLE INTL AIRPORT	US (UNITED STATES)	TN	36.124476 7	-86.6781822	**More…**
	SS99USTNG1PN	G1PN	Grumman G-1	US (UNITED STATES)	TN	-999.999999	-999.999999	
*TABLE ENDS								

Key Information Needed to Adequately Characterize Measurement and Sampling Sites

DES template provides guidance on identifying mobile measurement platforms (e.g., airplanes, vans, and ships). The Site information table documents the site information for sites with data appearing in this file. The table name must be as shown ("Site information"). Variables may be presented in a different order than shown, but we urge that the Site ID: standard variable appear first. Other variables besides those shown may be added to this table. We suggest you consult with your local data manager before adding other variables; standard names and picklists exist for some other variables.

DM-5: Reporting Sampling and Measurement Dates and Times

BACK TO TABLE

SCOPE: (Example Mexico City Metropolitan Area 2003 Field Campaign)

PURPOSE: Provides a standard for reporting sampling and measurement dates and times. **Resources, examples, and use in the DES template are shown**.

Because reporting dates and time is so important to the success of a project, we have designed redundancy into the reporting fields for date and time to prevent many of the reporting problems encountered by similar intensive monitoring projects.

Time Basis:

Investigators will report data on a **Central Standard Time (CST)** basis.

Dates and Times to Report:

- Start date and time must be reported as time at the beginning of the sampling/measurement/averaging period.
- End date and time must be reported as the time at the end of the sampling/measurement/averaging period.
- For continuous processes, the end date and time of the preceding period may be the start date and time of the next period.
- There is no 24:00 time. 23:59, then 00:00 the next day. Contact the QSSC for guidance if fractions of seconds are needed.

Reporting Dates and Times:

Local Time Zone is specified on every data record. Specify **CST**.

Sample dates and times must be reported in both **CST** and Coordinated Universal Time (UTC).

(1) **CST**. Formats: 2003/02/28 and 07:00. (Note leading zero. See footnote.)
(2) UTC. Formats: 2003/02/28 and 13:00.

CST lags UTC time by 6 hours. If the Universal Time is 14:30 UTC, Central Standard Time would be 08:30 CST.

Time Resources:

Discussion of Coordinated Universal Time (UTC) [http://aa.usno.navy.mil/faq/docs/UT.html].

U.S. Naval Observatory [http://www.usno.navy.mil/]

To set your PC to the correct U.S. time [http://nist.time.gov/]

Important note: A formula is provided in the main data table of the DES template for converting local dates and times to UTC.

Footnote: (Exact steps may vary slightly depending upon your operating system.)
For MS-Windows users, the default date and time format should be changed to the ISO format on every computer used to create Data Exchange Standard files, as follows:

a) On the Windows desktop, click on Start, Settings, Control Panel
b) Click on Regional Settings
c) Click on Date
d) In the "Short date style" field, enter yyyy/mm/dd
e) Click OK, and under Regional Settings, click on Time
f) In the Time style field, enter hh:mm:ss tt (this causes the hour to display leading zeros. e.g., 08:00)
g) Click OK

Reporting Sampling and Measurement Dates and Times in DES Template

*TABLE NAME	**Main Data Table**										
*TABLE FOCUS	Surface—fixed										
*TABLE USER NOTE											
*TABLE KEY FIELD NAMES	Site ID: standard	Instrument co-location ID	Date start: local time	Time start: local time							
*TABLE COLUMN NAME	**Site ID: standard**	Instrument co-location ID	**Date start: local time**	**Time start: local time**	**Date end: local time**	**Time end: local time**	**Time zone: local**	**Date start: UTC**	**Time start: UTC**	**Date end: UTC**	**Time end: UTC**
*TABLE COLUMN NAME TYPE	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable
*TABLE COLUMN CAS IDENTIFIER	None	None	None	None	None	None	None	None	None	None	None
*TABLE COLUMN USER NOTE	None	None	None	None	None	None	None	None	None	None	None
*TABLE COLUMN UNITS	None	None	**yyyy/mm/dd**	**hh:mm**	**yyyy/mm/dd**	**hh:mm**	**None**	**yyyy/mm/dd**	**hh:mm**	**yyyy/mm/dd**	**hh:mm**
*TABLE COLUMN FORMAT TYPE	Char	Char	**Date**	**Time**	**Date**	**Time**	**Char**	**Date**	**Time**	**Date**	**Time**
*TABLE COLUMN FORMAT FOR DISPLAY	12	2	10	5	10	5	6	10	5	10	5
*TABLE COLUMN MISSING CODE	None	None	None	None	None	None	None	None	None	None	None
*TABLE COLUMN LOOKUP TABLE NAME	Site information	Instrument co-location	None	None	None	None	None	None	None	None	None
*TABLE COLUMN OBSERVATION TYPE	Supplementary data	Supplementary data	Supplementary data	Supplementary data	Supplementary data	Supplementary data	Supplementary data	Supplementary data	Supplementary data	Supplementary data	Supplementary data
*TABLE COLUMN FIELD SAMPLING OR MEASUREMENT PRINCIPLE	Not applicable	Not applicable	Pending assignment	Pending assignment	Pending assignment	Pending assignment	Pending assignment	Not applicable	Not applicable	Not applicable	Not applicable
*TABLE BEGINS											
	SS99USTNBNA_	**P**	**2000/01/01**	**08:00**	**2000/01/02**	**07:00**	**CST**	**2000/01/01**	**14:00**	**2000/01/02**	**13:00**
*TABLE ENDS											

DM-6: Identifying Chemical and Physical Variables and Descriptive Field Information

BACK TO TABLE

SCOPE: ***Project***

PURPOSE: Provides the approach for identifying (i.e., naming) chemical and non-chemical/physical measured variables and various descriptive metadata elements. Resources are identified and examples are shown.

This document points to references tables of CAS Registry Numbers, names for chemical and nonchemical/physical variables and various metadata elements to use in the Data Exchange Standard files. Data providers are expected to use these tables to determine the appropriate identifiers for the chemical substances, physical properties, and metadata elements (e.g., date, time, locations) they are reporting.

Identifying Chemical Substances with a CAS[1] Registry Number:

Chemical Substances with a CAS RN	Limits
CAS Identifier: (CAS Registry Number with "C" prefix) Examples: C112-12-9 C6175-49-1 C15869-80-4 **Chemical Name:** (Prefered is CAS-9CI nomemclature. IUPAC for polycyclics. Other common name might be acceptable.) Exmples: (source 9CI) 2-Undecanone 2-Dodecanone Heptane, 3-ethyl-	Valid CAS number. The "C" prefix prevents spreadsheet programs from converting some CAS numbers to dates. Please request CAS RNs and 9CI names from the QSSC as needed.
Resource: http://cdiac.esd.ornl.gov/programs/NARSTO/SS_Chem_Ref_Tables.xls	

[1]The CAS Registry Number and the CAS-9CI name (Chemical Abstracts Service, 9th Collective Index Nomenclature) are the copyrighted property of the American Chemical Society. The NARSTO QSSC has the permission of CAS to use this information in NARSTO archive data sets. By extension, EPA Supersites Projects and NARSTO affiliated projects may incorporate CAS numbers and CAS-9CI names into data being processed for NARSTO archiving. Furthermore, the use of CAS numbers and CAS-9CI names is permitted as required in supporting regulatory requirements and/or for reports to Government Agencies and in copyrighted scientific publications when the CAS information are incidental to the publication. Any use or redistribution other than that described

here is not permitted without the prior, written permission of the American Chemical Society. Please contact the QSSC if you have any questions about the use of CAS information.

Identifying Chemical Substances/Measurements/Calculated Quantities that do not have a designated CAS Registry Number:

Chemical Substances without a CAS RN	Limits
Chemical Substances Identifier: Examples: Carbon: elemental (EC) Hydrocarbons: non-methane (NMHC) NOx (nitric oxide + nitrogen dioxide)	Formal syntax with key phrase and detailed modifier if needed, separated by a ":" Please request new names from the QSSC as needed.
Resource: http://cdiac.esd.ornl.gov/programs/NARSTO/Chems_without_CAS.xls	

Identifying Physical/Non-chemical Measurements:

Physical/Non-chemical Measurements	Limits
Physical/Non-chemical Identifier: Examples: PM2.5: mass Aircraft: heading Humidity: relative	Formal syntax with key phrase and detailed modifier if needed, separated by a ":" Please request new names from the QSSC as needed.
Resource: http://cdiac.esd.ornl.gov/programs/NARSTO/non-Chem_variable_names.xls	

Identifying Descriptive Metadata Elements:

Descriptive Metadata Elements	Limits
Descriptive Field Information Examples: MOUDI stage Flight ID Detection limit: sample-level Flag: NARSTO sample-level Date start: local time Time start: local time	Formal syntax with key phrase and detailed modifier if needed, separated by a ":" Please request new names from the QSSC as needed.
Resource: http://cdiac.esd.ornl.gov/programs/NARSTO/non-Chem_variable_names.xls	

DM-7: Reporting Units for Chemical Variables, Particles, and Physical and Descriptive Variables

BACK TO TABLE

SCOPE: *Project*

PURPOSE: Provides guidance for reporting units for chemical variables, particles, and physical and descriptive variables. **Resources, examples, and use in the DES template are shown.**

Use the following guidance about units:

- Typically particles are reported in mass/volume units, varying depending on the magnitude: ug/m3, ng/m3, pg/m3, etc;
- Gas phase species are typically reported in units of ppm, ppb and ppt by volume;
- Organic compounds in the gas phase are to be reported as ppmC, ppbC, or pptC by volume;
- Fog or cloud measurements are typically reported in units of microequivalents per liter or micromoles per liter. Please avoid using mass/volume for these kinds of measurements;
- It is difficult to come up with a single unit for all chemical species, and we allow variation due to convention;
- Conversion factors may be used to achieve these conventional units. (Ref. 1)
- We provide a table of standard non-chemical variable names with recommended units.

These are SI units in most cases. Please use these units (with appropriate prefixes), or consult with the QSSC if there is a compelling reason to use different units.

Reference:

1. Finlayson-Pitts, B.J. and Pitts, Jr., J.N., 1986. Atmospheric Chemistry. John Wiley & Sons, New York.

Values for *TABLE COLUMN UNITS Key Phrase that must be specified for units. (Additional values may be requested from the QSSC.)

KEY PHRASE	VALUES FOR DESCRIBING UNITS
***TABLE COLUMN UNITS**	---- MIXING RATIO (MOLE FRACTION) ---- g/kg (gram per kilogram) ppbv (part per billion by volume) ppmv (part per million by volume) ppmvC (part per million Carbon by volume) pptv (part per trillion by volume) ppbvC (part per billion Carbon by volume) mol/mol (mole per mole) ---- CONCENTRATION ---- mol/L (mole/liter) mg/L (milligram per liter) ng/m3 (nanogram per cubic meter)

	pg/ml (picogram per milliliter) pg/m3 (picogram per cubic meter) ug/m3 (microgram per cubic meter) ---- OTHER PARTICLE MEASUREMENT ---- g/cm3 (gram per cubic centimeter) number/cm3 (number per cubic centimeter) number/m3 (number per cubic meter) um3/cm3 (cubic micrometer per cubic centimeter) molecule/cm3 (molecule per cubic centimeter) ---- AREA ---- um2/cm3 (square micrometer per cubic centimeter) ---- TEMPERATURE ---- deg C (degree Celsius) ---- DIRECTION ---- degree from true north ---- LENGTH/ALTITUDE/HEIGHT---- cm-1 (inverse centimeter) km (kilometer) m (meter) mm (millimeter) nm (nanometer) m-1 (inverse meter) Mm-1 (inverse megameter) um (micrometer) ---- VOLUME ---- m3 (cubic meter) ---- TIME ---- yyyy/mm/dd hh:mm hh:mm:ss d (day) h (hour) min (minute) s (second) ---- PRESSURE ---- hPa (hectopascal) kPa (kilopascal) Pa (pascal) mb (millibar) ---- SPEED ---- km/h (kilometer per hour) m/s (meter per second) ---- FREQUENCY ---- Hz (hertz) ---- FORCE ---- N (newton) ---- ENERGY ----

	joule ---- POWER ---- W (watt) W/m2 (watt per square meter) mW/m3 (milliwat per cubic meter) ---- ELECTRIC CHARGE ---- C (coulomb) ---- LOCATION ---- decimal degree ---- OTHER ---- l/min (liter per minute) % (percent) DV (Deciview) None Not applicable Pending assignment pH units volt

DM-8: Assigning Project-Specific and NARSTO Data Quality Flags

BACK TO TABLE

SCOPE: *Project*

PURPOSE: Provides a resource document for Projects to use as they determine the most appropriate data flagging approach. Reported data values must be assigned at least one data quality flag by the data originator that indicates to a data user whether the data are valid without qualification, valid but qualified/suspect, or invalid due to serious sampling or analysis problems. **Resources, examples, and use in the DES template are shown.**

Flag Guidance and Resources

The data originator set flags may be the NARSTO data qualification flags (Table 1) or other flags as defined by a Project. Project-defined flags (e.g., Table 2) may be carried to the archive, but they **must** also be mapped to NARSTO flags (i.e., NARSTO flags must be added) before sending the data for archiving by NARSTO.

A more comprehensive list of detailed data quality flags that ***may*** be used by a project to flag individual atmospheric monitoring and analytical laboratory measurement results is shown in Table 2.

A set of exceptional event flags (U.S. EPA source) that a project ***may*** use to flag a sample or set of samples is shown in Table 3.

Although either the NARSTO data qualification flags or the detailed data quality flags might adequately qualify measurement values, the detailed quality flags provide much more information about the nature of the qualification and would be the preferred flag to be assigned by the data originator. Exceptional event flags may be assigned to a sample as needed to indicate the possible influence of local or larger scale events that may impact the representativeness of the sample.

Table 1. NARSTO Data Qualification Flags

NARSTO Flag Code	NARSTO Flag Definition	NARSTO Flag Description	Applicability
V0	Valid value	Flag valid data values.	Applies to all measurement data types.
V1	Valid value but comprised wholly or partially of below-MDL data	Applies to both single and averaged data values. Measured value reported, even when below MDL. Define MDL (minimum detectable limit).	Applies to all measurement data types. See related V7 flag.
V2	Valid estimated value	Can apply to calculated values, approximate/out-of-range values, and measured values with, for example, EPA "J" flag.	Applies to all measurement data types, model input and output data products, and gridded data products.
V3	Valid interpolated value	Valid interpolated value. Provide interpolation method in documentation.	Applies to all measurement data types, model input and output data products, and gridded data products.
V4	Valid value despite failing to meet some QC or statistical criteria	Apply this flag based on evaluation of field and laboratory QC sample data and subsequent statistical outlier tests on the entire data set.	Applies to all measurement data types.
V5	Valid value but qualified because of possible contamination (e.g., pollution source, laboratory contamination source)	Apply this flag for possible contamination of blanks and regular samples.	Applies to all measurement data types.
V6	Valid value but qualified due to non-standard sampling conditions (e.g., instrument malfunction, sample handling)	Provide description of sampling conditions or variance from SOP in documentation.	Applies to all measurement data types.
V7	Valid value but set equal to the detection limit (DL) because the measured value was below the DL	Applies to both single and averaged data values. The measurement was below DL and the Principal Investigator lacks confidence in it and the DL was substituted in its place. Define MDL (minimum detectable limit).	Applies to all measurement data types. See related V1 flag.
M1	Missing value because no value is available	Use this flag when no result was reported. Identify the missing value code that is used in the result field.	Applies to all measurement data types.
M2	Missing value because invalidated by Data Originator	Use this flag when the result reported to a project database was invalid. Invalid results are not sent to the NARSTO archive. Identify the missing value code that is used in the result field.	Applies to all measurement data types.
H1	Historical data that have not been assessed or validated	Use this flag when, for example, historical data may have been used for preliminary characterization or range finding purposes. It will not be used in subsequent analyses but is part of the project record.	Applies to all measurement data types, model input and output data products, and gridded data products.

Table 2. Comprehensive List of Detailed Data Quality Flags That *May* Be Used by a Project

NARSTO Data Flag	Flag Category	Detailed Flag	Acronym Derivation	Detailed Flag Description
V0	All Categories	**NIE**	**N**o **I**ssues **E**ncountered	No Problems or Issues Encountered
V0	Analysis	**NLC**	**N**on-**L**isted **C**ompound	Compound Not Listed but Compound Found
V0	Analysis	**VCD**	**V**alue **C**onfirme**D**	Value Confirmed
V0	Calculated	**DAV**	**D**ata **AV**eraged	Data Averaged
V0	Calibration	**CFC**	**C**orrection **F**actor **C**alibration	Correction Factor, Calibration
V0	Calibration	**CTP**	**C**orrection **T**emp **P**ressure	Correction Factor, Std. Temp. and Pressure
V0	Procedure	**WUS**	**W**rong **U**nit**S**	Wrong Units
V0	QC	**BDQ**	**B**etween **D**etection/**Q**uantification	Between Instrument Detection and Quantification Limits
V0	Calculated	**ROR**	**R**ecalculated **O**r **R**eprocessed	Recalculated or Reprocessed
V1	QC	**AND**	**A**nalyte **N**ot **D**etected	Analyte Not Detected
V1	QC	**BDD**	**B**elow **D**aily **D**etection	Daily Detection Limit, Less than
V1	QC	**BID**	**B**elow **I**nstrument **D**etection	Instrument Detection Limit, Less than
V1	QC	**BMD**	**B**elow **M**ethod **D**etection	Method Detection Limit, Less than
V1	QC	**BSA**	**B**elow **SA**mple Detection	Sample-Specific Detection Limit, Less than
V1	QC	**BSY**	**B**elow **SY**stem Detection	System Detection Limit, Less than
V1	QC	**OBS**	**O**peration **B**eyond **S**cale	Value Not Confirmed, Operation Beyond Scale Setting
V2	Calculated	**NCS**	**N**o **C**alibration **S**tandard	Estimated Value, No Calibration Standard
V2	Calculated	**OLP**	**O**utside **L**imit of **P**recision	Estimated Value, Outside Limit of Precision
V3	Calculated	**ITV**	**I**n**T**erpolated **V**alue	Interpolated Value
V2	QC	**AOR**	**A**bove **O**perating **R**ange	Operating Range, Greater than
V2	QC	**BOR**	**B**elow **O**perating **R**ange	Operating Range, Less than
V2	QC	**MOL**	**M**ass **O**utside **L**imits	Fraction of Total Mass, Out of Acceptable Limits
V2	Analysis	**ALM**	**AL**ternate **M**ethod	Alternate Method
V2	QC	**OOR**	**O**utside **O**perating **R**ange	Operating Range, Not within
V2	Analysis	**RSL**	**R**e**SL**oped	Resloped
V2	Analysis	**VNC**	**V**alue **N**ot **C**onfirmed	Value Not Confirmed
V4	Calculated	**UHA**	**U**nacceptable **H**ourly **A**verage	Time Period Average with less than 75% of possible data points
V5	Procedure	**WRC**	**W**eather **R**elated **C**ontamination	Weather Related Contamination
V6	Procedure	**WTO**	**W**rong **T**imes/**O**versampled	Wrong Times, Oversampled
V6	Procedure	**WTU**	**W**rong **T**imes/**U**ndersampled	Wrong Times, Undersampled
V4	QC	**FDF**	**F**ield **D**uplicated **F**ailed	Field Duplicate, Failed
V4	QC	**FRS**	**F**low **R**ate **S**uspect	Flow Rate, Problem or Suspect
V4	QC	**VSF**	**V**erification **S**olution **F**ailed	Lab Calibration Verification Solution, Failed
V5	Procedure	**SCN**	**S**uspected **C**ontaminatio**N**	Suspected Contamination, Lab Analysis or Field
V4	QC	**BSF**	**B**lank **S**ample **F**ailed	Blank Sample, Failed

NARSTO Data Flag	Flag Category	Detailed Flag	Acronym Derivation	Detailed Flag Description
V4	QC	**CSF**	**C**alibration **S**ample **F**ailed	Calibration Sample, Failed
V4	QC	**DCF**	**D**rift **C**heck **F**ailed	Drift Check, Failed
V4	QC	**ISF**	**I**nternal **S**tandard **F**ailed	Internal Standard, Failed
V4	QC	**LBH**	**L**ikely **B**iased **H**igh	Likely Biased High
V4	QC	**LBL**	**L**ikely **B**iased **L**ow	Likely Biased Low
V4	QC	**LCF**	**L**inearity **C**heck **F**ailed	Linearity Check, Failed
V4	QC	**PCF**	**P**erformance **C**heck **F**ailed	Performance Check, Failed
V4	QC	**RMD**	**R**eplicate **M**easurements **D**isagree	Replicate Measurements Disagree
V4	QC	**RMF**	**R**eference **M**aterial **F**ailed	Reference Material, Failed
V4	QC	**RMI**	**R**eference **M**easurement **I**nconsistent	Reference Method Measurement, Inconsistent
V4	QC	**SRF**	**S**tandard **R**eference **F**ailed	Standard Reference Material, Failed
V6	Sampling	**PSO**	**P**artial **S**ample **O**verloaded	Partial Sample, Due to Overload
V6	Procedure	**COC**	**C**hain **O**f **C**ustody	Chain-of-Custody Evidence Trail Broken or Lost
V6	Procedure	**DNL**	**D**amaged/**N**ot **L**ost	Sample Container Damaged, No Sample Lost
V6	Procedure	**DSL**	**D**amaged/**S**ample **L**ost	Sample Container Damaged, Sample Lost
V6	Procedure	**EHT**	**E**xceeded **H**olding **T**ime	Exceeded Holding Time
V6	Procedure	**ISP**	**I**mproper **S**ample **P**reservation	Improper Sample Preservation
V6	Procedure	**ILG**	**Il**legible/**G**uess	Illegible Paperwork or Mislabeled, Made Guess
V4	QC	**LDF**	**L**ab **D**uplicate **F**ailed	Lab Duplicate, Failed
M1	Analysis	**ANC**	**AN**alysis **C**ancelled	No Result Reported, Analysis Canceled
M1	Analysis	**NAL**	**N**ot **A**nalyzed/**L**isted	Compound Not Analyzed but Compound Listed
M1	QC	**NRI**	**N**othing **R**eported - **I**nterference	No Result Reported, Interference
M1	Instrument	**OEL**	**O**perator **E**rror **L**oading	Sample Loading, Operator Error
M1	Instrument	**SAM**	**SA**mpler **M**alfunction	Sampler Malfunction
M1	Instrument	**SAU**	**S**ite **A**ccess **U**navailable	Site Access, Unavailable
M1	Instrument	**SHP**	**SH**ipment **P**roblems	Shipment Problems
M1	Instrument	**SIS**	**S**afety **I**ssue**S**	Safety Issues
M1	Instrument	**SNA**	**S**ampler **N**ot **A**ctivated	Sampler Not Activated
M1	Procedure	**EER**	**E**ntry **E**rro**R**	No Result Reported, Entry Error
M1	Procedure	**FAC**	**F**ield **AC**cident	No Result Reported, Field Accident
M1	Procedure	**WDY**	**W**rong **D**a**Y**	Wrong Day
M1	Instrument	**INF**	**IN**strument **F**ailure	Instrument Failure
M1	Instrument	**NSQ**	**N**ot **S**ufficient **Q**uantity	No Result Reported, Insufficient Quantity of Sample
M1	Instrument	**VAN**	**VAN**dalism	Vandalism
M1	Procedure	**ILP**	**I**llegible **L**abel **P**aperwork	Illegible Label or Paperwork
M1	Procedure	**ZME**	**Z**ero **M**od**E**	Zero Mode
M2	QC	**OSR**	**O**ff-**S**cale **R**eading	Off-Scale Reading
M2	QC	**VCH**	**V**alue **C**hange **H**igh	Value Change Too High, Above Physical Limit
M2	QC	**VTH**	**V**alue **T**oo **H**igh	Value Too High, Above Physical Limit
M2	Instrument	**EMM**	**E**lectrical or **M**echnical	Problem, Electrical or Mechanical

NARSTO Data Flag	Flag Category	Detailed Flag	Acronym Derivation	Detailed Flag Description
			Malfunction	Malfunction
M2	Instrument	**IAS**	**I**nvalid **A**ir **S**ample	Invalid, Air Sample
M2	Procedure	**CSU**	**C**ontamination **S**ource **U**nknown	Known Contamination, Source Unknown
M2	Procedure	**WFR**	**W**rong **F**ilte**R**	Wrong Filter
M2	Sampling	**NRE**	**N**ot **R**epresentativ**E**	Not Representative
M2	Calibration	**CFB**	**C**orrection **F**actor **B**lank	Correction Factor, Blank
M2	QC	**QSF**	**Q**C **S**amples **F**ailed	Multiple QC Samples, Failed

Table 3. Exceptional Event Flags

The exceptional event flag codes and definitions are U.S. EPA AIRS flag codes and definitions for compatibility.
Source:

U.S. EPA AIRS Database, Air Quality Subsystem (AQS)
Exceptional Event Flags (Re-Engineered AQS, http://www.epa.gov/ttn/airs/)

Exceptional_ Event_Flag_ Code	AIRS_ Flag_Code	Code Definition: The reason for an abnormal observation
EXA	A	HIGH WINDS
EXB	B	STRATOSPHERIC OZONE INTRUSION
EXC	C	VOLCANIC ERUPTIONS
EXD	D	SANDBLASTING
EXE	E	FOREST FIRE
EXF	F	STRUCTURAL FIRE
EXG	G	HIGH POLLEN COUNT
EXH	H	CHEMICAL SPILLS & INDUST. ACCIDENTS
EXI	I	UNUSUAL TRAFFIC CONGESTION
EXJ	J	CONSTRUCTION/DEMOLITION
EXK	K	AGRICULTURAL TILLING
EXL	L	HIGHWAY CONSTRUCTION
EXM	M	REROUTING OF TRAFFIC
EXN	N	SANDING/SALTING OF STREETS
EXO	O	INFREQUENT LARGE GATHERINGS
EXP	P	ROOFING OPERATIONS
EXQ	Q	PRESCRIBED BURNING
EXR	R	CLEAN UP AFTER A MAJOR DISASTER
EXS	S	SEISMIC ACTIVITY
EXT	T	SAHARA OR SEVERE DUST

Reporting NARSTO flags in the DES Template

NARSTO Standard Flag Look-up Table in DES Template

*TABLE NAME	NARSTO standard flags	
*TABLE FOCUS	Metadata	
*TABLE COLUMN NAME	Flag: NARSTO	Description
*TABLE COLUMN UNITS	None	None
*TABLE COLUMN FORMAT TYPE	Char	Char
*TABLE COLUMN FORMAT FOR DISPLAY	2	120
*TABLE BEGINS		
	V0	**Valid value**
	V1	**Valid value but comprised wholly or partially of below detection limit data**
	V2	**Valid estimated value**
	V3	**Valid interpolated value**
	V4	**Valid value despite failing to meet some QC or statistical criteria**
	V5	**Valid value but qualified because of possible contamination (e.g., pollution source, laboratory contamination source)**
	V6	**Valid value but qualified due to non-standard sampling conditions (e.g., instrument malfunction, sample handling)**
	V7	**Valid value but set equal to the detection limit (DL) because the measured value was below the DL**
	M1	**Missing value because no value is available**
	M2	**Missing value because invalidated by data originator**
	H1	**Historical data that have not been assessed or validated**
*TABLE ENDS		

Use of Data Quality Flags in the DES file.

This guidance applies to NARSTO standard flags (not necessarily to sample-level or subsample-level flags, see Frequently Asked Questions worksheet in DES template).

Values for Key Phrases that must be specified for Flag data columns

KEY PHRASE	VALUES FOR FLAG COLUMNS
***TABLE COLUMN NAME**	**Value must be the same for the Variable and Flag column(s) (**e.g., Ozone, Carbon: total, or Peroxide: total)**.**
***TABLE COLUMN NAME TYPE**	**Select value from picklist as appropriate for flag** (i.e., Flag--NARSTO, Flag--Supersites detailed, Flag--study, or Flag--exceptional event).
***TABLE COLUMN FORMAT TYPE**	**Select "Char" value from picklist.**
***TABLE COLUMN FORMAT FOR DISPLAY**	**Enter the maximum number of characters allowed for the flag.**
***TABLE COLUMN MISSING CODE**	**Enter "None" as missing code.**
***TABLE COLUMN LOOKUP TABLE NAME**	**Enter "Lookup table name" as entered in FLAG description metadata table (e.g., NARSTO standard flags)**
*TABLE COLUMN BASIS	Optional Key Phrase. If used, value should be the same for the Variable and associated column(s). See DES template "Frequently Asked Questions".
*TABLE COLUMN GROUP CODE	Optional Key Phrase. If used, value should be the same for the Variable and associated column(s). See DES template "Frequently Asked Questions".

This excerpt from the NARSTO DES template shows the proper selection of values for essential flag key phrases shown in the previous table. These cells are highlighted in yellow.

All other key phrase values may be identical for the Variable and the Flag columns as shown in the first two Ozone columns. This is the preferred approach so as to minimize the chance for confusion across variables and their flags.

However, this is not a requirement; the fields may be set to "None" or "Not applicable" values as appropriate. See third Ozone column. Often these are the default values in the DES template.

Portion of completed DES format file showing two approaches to completing Key Phrase values for quality flag columns.

*TABLE COLUMN NAME	Ozone	Ozone	Ozone
*TABLE COLUMN NAME TYPE	Variable	Flag--NARSTO	Flag—study
*TABLE COLUMN CAS IDENTIFIER	C10028-15-6	C10028-15-6	None
*TABLE COLUMN EXPLANATION OF DERIVED VARIABLE	Not applicable	Not applicable	Not applicable
*TABLE COLUMN USER NOTE	None	None	None
*TABLE COLUMN USER NOTE2	None	None	None
*TABLE COLUMN UNITS	ppbv (part per billion by volume)	ppbv (part per billion by volume)	None
*TABLE COLUMN FORMAT TYPE	Decimal	Char	Char
*TABLE COLUMN FORMAT FOR DISPLAY	8.3	2	3
*TABLE COLUMN MISSING CODE	-999.999	None	None
*TABLE COLUMN LOOKUP TABLE NAME	None	NARSTO standard flags	Study flags lookup
*TABLE COLUMN OBSERVATION TYPE	Gas	Gas	Not applicable
*TABLE COLUMN FIELD SAMPLING OR MEASUREMENT PRINCIPLE	Chemiluminescence	Chemiluminescence	Not applicable
*TABLE COLUMN SAMPLING HEIGHT ABOVE GROUND (M)	10	10	Not applicable
*TABLE COLUMN INLET TYPE	Open sampling line	Open sampling line	Not applicable
*TABLE COLUMN SAMPLING HUMIDITY OR TEMPERATURE CONTROL	None	None	Not applicable
*TABLE COLUMN BLANK CORRECTION	Blank corrected	Blank corrected	Not applicable
*TABLE COLUMN VOLUME STANDARDIZATION	Ambient temperature and pressure	Ambient temperature and pressure	Not applicable
*TABLE COLUMN INSTRUMENT NAME AND MODEL NUMBER	New model 1	New model 1	Not applicable
*TABLE COLUMN	John Doe	John Doe	Not applicable

MEASUREMENT PRINCIPAL INVESTIGATOR			
*TABLE COLUMN DETECTION LIMIT	0.01	0.01	Not applicable
*TABLE BEGINS			
	6.333	V0	ABC
*TABLE ENDS			

DM-9: Reporting and Flagging Values below Detection Limits

BACK TO TABLE

SCOPE: *Project*

PURPOSE: Provides guidance for reporting detection limits and data below the limit of detection, and how to flag these values with NARSTO standard flags. **Resources, examples, and use in the DES template are shown.**

As general guidance:

Report detection limits (when known) in the data file (See *TABLE EXPLANATION OF REPORTED DETECTION LIMIT VALUES and *TABLE COLUMN DETECTION LIMIT) and

(Preferred) Report the **actual measured value** even if the value is below the detection limit (including zero and negative values) with the value **flagged "V1"** (Valid value but comprised wholly or partially of below detection limit data) or

Alternatvely, report that the **measured value has been substituted with the detection limit** and the value **flagged "V7"** (Valid value but set equal to the detection limit (DL) because the measured value was below the DL). We do not, however, recommend substituting zero or the detection limit for below detection limit data.

- If a measured value is below what is considered to be the normal “Detection Limit” (DL) but is nonetheless considered meaningful, suggesting where between zero and the DL the value lies, the measured value should be provided, and the NARSTO V2 flag ("Valid estimated value”), should be applied. If an investigator's detailed flags are being used one or more of these may also be applied.

DES Template guidance:

***TABLE EXPLANATION OF REPORTED DETECTION LIMIT VALUES** should contain a brief narrative of how below-detection-limit data are reported and flagged. A detailed discussion should be given in an associated metadata document if needed.

Two examples of the narratives are as follows:
1) The laboratory reported all measured values. Those values below detection limit are flagged as V1 (Valid value but comprised wholly or partially of below detection limit data).

2) Values below detection limit are replaced by the detection limit value for that analytical run and flagged V7 (Valid value but set equal to the detection limit (DL) because the measured value was below the DL).

If there are no below-detection-limit values reported in the data table, use a statement such as "No below-detection-limit values are reported this table."

***TABLE COLUMN DETECTION LIMIT** is available to provide detection limits for every measurement column.

For variables having detection limits that vary from one measurement row to another, use the key phrase ***TABLE COLUMN NAME TYPE** (select value **"Detection limit"**) to add a column for detection limits.

Portion of DES format file showing two possible approaches for reporting detection limits.

*TABLE NAME	**Detection Limit Example**	**(omitted columns)**			
*TABLE FOCUS	Surface--fixed				
*TABLE EXPLANATION OF ZERO OR NEGATIVE VALUES					
***TABLE EXPLANATION OF REPORTED DETECTION LIMIT VALUES**	**To be completed by PI.**				
*TABLE EXPLANATION OF REPORTED UNCERTAINTY					
*TABLE KEY FIELD NAMES	Site ID: standard				
***TABLE COLUMN NAME**	**Site ID: standard**		**Variable '1'**	**Variable '1'**	**Variable '1'**
***TABLE COLUMN NAME TYPE**	Variable		Variable	Flag--NARSTO	**Detection limit**
*TABLE COLUMN CAS IDENTIFIER	None		None	None	None
*TABLE COLUMN UNITS	None		m (meter)	None	m (meter)
*TABLE COLUMN FORMAT TYPE	Char		Decimal	Char	Decimal
*TABLE COLUMN FORMAT FOR DISPLAY	12		8.3	2	8.3
*TABLE COLUMN MISSING CODE	None		-999.999	None	-999.999
*TABLE COLUMN LOOKUP TABLE NAME	Site information		None	NARSTO standard flags	None
***TABLE COLUMN DETECTION LIMIT**	Not applicable		Not applicable or repeat DL value for column	Not applicable or repeat DL value for column	**Detection limit value for column**
(omitted Key Phrases)					
*TABLE BEGINS					
	Site		**Value**	**Flag**	**Detection limit for each measured value**
*TABLE ENDS					

Examples of Applying Flags to Values below Detection Limit

Condition	Flag
Value is below the DL but the value is considered meaningful	V2
Value is in an average and part or all of the data in the average are below the DL however the value is considered a meaningful estimate	V1 or V7 (see below)
a) actual measured value is provided	V1
b) substituted value equal to the DL value is provided	V7
c) substitued value equal to a constant (e.g., 50%DL or zero)	This practice is discouraged for research data

Example 1: Reporting below-DL values without substitution

Value	Flag	Notes (these are NOT the flag definitions)
5.32	V0	Valid
3.49	V0	Valid
0.05	V0	Valid (this measured value is at the Detection Limit)
0.04	V2	Valid, below DL but considered quantifiable, meaningful, or usable
0.03	V2	Valid, below DL but considered quantifiable or meaningful, or usable
0.02	V2	Valid, below DL but considered quantifiable or meaningful, or usable
0.01	V1	Valid, below DL and Principal Investigator lacks confidence in it, but may be statistically useful
-0.01	V1	Valid, below MDL and Principal Investigator lacks confidence in it, but may be statistically useful
-0.02	V1	Valid, below MDL and Principal Investigator lacks confidence in it, but may be statistically useful

Example 2: Reporting below-DL value with substitution

Value	Flag	Notes (these are NOT the flag definitions)
5.32	V0	Valid
3.49	V0	Valid
0.05	V0	Valid (this measured value is at the Detection Limit)
0.04	V2	Valid, below DL but considered quantifiable or meaningful
0.03	V2	Valid, below DL but considered quantifiable or meaningful
0.02	V2	Valid, below DL but considered quantifiable or meaningful
0.05	V7	Valid, the measurement was below DL and the Principal Investigator lacks confidence in it and the DL was substituted in its place.

DM-10: Reporting Missing Data

BACK TO TABLE

SCOPE: *Project*

PURPOSE: Provide guidance for reporting missing data and selecting the appropriate flag. Resources, examples, and use in the DES template are shown.

Reporting Missing Data

All data fields must have a value present, either the measured value, substituted detection limit, or a missing value representation.

There may not be blank data fields.

The *TABLE COLUMN MISSING CODE key phrase records the code used to represent a missing value. This code will appear in the data whenever the value is missing.

- Use 'None' as the missing code for Char, Date, and Time values.
- Use '.' (a period) as the missing code for project-derived Date and Time values.
- Missing codes for Int, Decimal, and Scientific format variables need to
  - (1) match the format of the column's *TABLE COLUMN FORMAT TYPE and *TABLE COLUMN FORMAT FOR DISPLAY;
  - (2) in general, be negative and large enough to be impossible as actual data, and
  - (3) use repeated 9's (e.g., Decimal: -999.99; Int: -999), except for the exponent in Scientific notation. Use +02 or a similar appropriate value as the exponent in Scientific notation (e.g., -9.99E+02, NOT -9.99E+99).

Flagging Missing Values

Missing values must be flagged with either the "M1" or "M2" flag as appropriate.

Flag: NARSTO	Description
M1	Missing value because no value is available
M2	Missing value because invalidated by data originator

Note that the QSSC automated quality assurance software checks for consistency of the missing value as reported in the data table with the missing value code as specified in the *TABLE COLUMN MISSING CODE key phrase, and that the NARSTO flag value is either "M1" or "M2".

DM-11: Reporting Uncertainty Estimates

BACK TO TABLE

SCOPE: *Project*

PURPOSE: Provide guidance for providing a brief explanation of the meaning/interpretation of the uncertainty values, or for characterizing the uncertainty, associated with the measurements. Resources, examples, and use in the DES template are shown.

Reporting Uncertainty

Provide an entry for the *TABLE EXPLANATION OF REPORTED UNCERTAINTY key phrase.

> Example: Each tabulated uncertainty value is a number which, when added to or subtracted from the variable's value, provides a 95% confidence interval for the true value.
>
> Example: Uncertainty is not reported with these data but will be included in the final QA report.

Uncertainty values that can be expressed as a constant for all of the data in the data file can be reported in the *TABLE EXPLANATION OF REPORTED UNCERTAINTY key phrase.

Sample-level uncertainty can be reported using the variable name "Uncertainty: sample-level".

Uncertainty values can be provided for each measurement value by adding one or more associated columns to the right of the measurement (using the same variable name but selecting the appropriate uncertainty-related *TABLE COLUMN NAME TYPE).

If uncertainty needs to be provided as a constant value for some or all variables, a *TABLE COLUMN USER NOTE key phrase may be used for this purpose; such variables do not need separate uncertainty columns to repeat the uncertainty value.

Resources

See the "Uncertainty and statistics" and the "Reporting uncertainty" worksheets in the DES Template workbook for more detailed advice about uncertainty.

See the NARSTO QSSC web site for a discussion of Uncertainty – “Submittal of Uncertainty Estimates for the Supersite Program”

Portion of DES format file showing three possible approaches for reporting uncertainty.

*TABLE NAME	**Uncertainty Example**	**(omitted columns)**			
*TABLE FOCUS	Surface--fixed				
*TABLE EXPLANATION OF ZERO OR NEGATIVE VALUES					
*TABLE EXPLANATION OF REPORTED DETECTION LIMIT VALUES					
***TABLE EXPLANATION OF REPORTED UNCERTAINTY**	**To be completed by PI.**				
*TABLE KEY FIELD NAMES	Site ID: standard				
***TABLE COLUMN NAME**	Site ID: standard		**Variable '1'**	**Variable '1'**	**Variable '1'**
***TABLE COLUMN NAME TYPE**	Variable		Variable	Flag--NARSTO	**Uncertainty--95% CI**
*TABLE COLUMN CAS IDENTIFIER	None		None	None	None
***TABLE COLUMN USER NOTE**	None		**Specify Uncertainty**	None	None
*TABLE COLUMN UNITS	None		m (meter)	None	m (meter)
*TABLE COLUMN FORMAT TYPE	Char		Decimal	Char	Decimal
*TABLE COLUMN FORMAT FOR DISPLAY	12		8.3	2	8.3
*TABLE COLUMN MISSING CODE	None		-999.999	None	-999.999
*TABLE COLUMN LOOKUP TABLE NAME	Site information		None	NARSTO standard flags	None
(omitted Key Phrases)					
*TABLE BEGINS					
	Site		**Value**	**Flag**	**Uncertainty Value**
*TABLE ENDS					

DM-12: Reporting Conventions for Mass Measurements, Meteorological Data, and Temperature and Pressure Conditions

BACK TO TABLE

SCOPE: *Project*

PURPOSE: Provides guidance for reporting data for several sample media and data types. **Resources, examples, and use in the DES template are shown.**

Particle Mass Measurements

- Mass/volume measurements (e.g., from filters) should be reported as concentrations, rather than separately as mass and volume.
- PM mass data should be referenced to local ambient temperature and pressure conditions to be comparable to federal reference method PM data.

Volumetric and Mass/volume Measurements

- Gas flow, volumetric measurements, and mass/volume conversions, must be reported at known (ambient or standardized) temperature and pressure conditions.

Associated Meteorological Data

- Meteorological data should be reported either in the same file or in a referenced meteorological data file as described below.
- The associated temperature and pressure conditions should be provided in the same file, or these should be available from file(s) of meteorological data from the same site or closest available site.
- If these are provided in the same file the temperature and pressure measurements should be available at the same time basis as the measurements. However, if temperature and pressure measurements are available less frequently than the target measurement, only those values actually measured should be presented in the data file. Missing value codes (not interpolated values) should be used for times when no actual measurements of temperature and pressure were taken.
- If measurements are standardized, U.S. EPA prefers that the reference conditions, i.e., 298.15 deg K (25.00 deg C) and 760 Torr (760 mm Hg); be used as stated in Section 40 chapter 50.3 of the Code of Federal Regulations.

Documenting Temperature and Pressure Conditions

Whatever temperature and pressure conditions are used must be documented in the Data Exchange Standard's *TABLE COLUMN VOLUME STANDARDIZATION Key Phrase.

Values for Key Phrases that must be specified for measurement temperature and pressure conditions. (These are picklist values in the DES Template; additional values may be requested from the QSSC.)

KEY PHRASE	VALUES FOR DESCRIBING TEMPERATURE AND PRESSURE
*TABLE COLUMN SAMPLING HUMIDITY OR TEMPERATURE CONTROL	Diffusion dryer Humidification Humidification with temperature condition at ambient Humidification with temperature conditioning at 50 deg. C Manual control (see metadata) Nafion dryer Nafion dryer with temperature conditioning at 30 deg. C Temperature conditioning at 20 deg. C Temperature conditioning at 25 deg. C Temperature conditioning at 30 deg. C Temperature conditioning at 40 deg. C Temperature conditioning at 50 deg. C Temperature conditioning at ambient Temperature controlled Other (see metadata) None Not applicable Pending assignment
*TABLE COLUMN VOLUME STANDARDIZATION	0 deg. C; 1 atmosphere 0 deg. C; ambient pressure 20 deg. C; 1 atmosphere 20 deg. C; ambient pressure 25 deg. C; 1 atmosphere 25 deg. C; ambient pressure Ambient temperature and pressure Not applicable Pending assignment

DM-13: Planning to Archive Data

BACK TO TABLE

SCOPE: *Project*

PURPOSE: Advanced planning for archiving project data furthers efforts to identify, collect, and report consistent data and metadata and to facilitate timely data analysis, sharing, integration, and synthesis. **Resources, examples, and use of the DES template are shown.**

Data from NARSTO projects are formatted in the NARSTO Data Exchange Standard (a spreadsheet-compatible layout, which uses standardized and consistent metadata values) and are sent to the NARSTO Quality Systems Science Center (QSSC). The QSSC does some format and content checking, interacts with the data providers as needed, develops additional archive documentation, and sends the data to the NARSTO Permanent Data Archive.

Learn about the Data Exchange Standard (DES) Format and Metadata.

Download the latest version of the Data Exchange Standard template [http://cdiac.esd.ornl.gov/programs/NARSTO/NARSTO_template_atmospheric_measurements.xls]

Access guidance for using the DES to report certain types of data:

Report data collected from mobile platforms such as aircraft, boats, and vans.

- Download an annotated template that has been modified to report data and sampling site information from mobile platforms.

- [mobile_site_annotated_template_2_302_developing_20041124_1.xls]

*DATA EXCHANGE STANDARD VERSION	NARSTO 2005/04/29 (2.302)	
*COMMENT	Further instructions on how to fill in this template are provided at:	
*COMMENT	http://cdiac.ornl.gov/programs/NARSTO/narsto.html	
*QUALITY CONTROL LEVEL		
*DATE THIS FILE GENERATED/ARCHIVE VERSION NUMBER	2005/02/21	1
*ORGANIZATION ACRONYM		
*ORGANIZATION NAME		
*STUDY OR NETWORK ACRONYM		
*STUDY OR NETWORK NAME		
*FILE CONTENTS DESCRIPTION--SHORT/LONG	**Mobile**	**mobile platform template**
*PRINCIPAL INVESTIGATOR NAME--LAST/FIRST		
*PRINCIPAL INVESTIGATOR AFFILIATION		
*CO-INVESTIGATOR NAME--LAST/FIRST		

More……

Report basic meteorological data only.

- Guidance that shows how to simplify the DES template to report straightforward meteorological data.

Reporting Basic Meteorological Data

This is guidance for using the DES template to report only basic meteorological data.

- This data type will typically include ground-based or tower measurements of Humidity, Pressure, Temperature, Precipitation, Wind direction, and Wind speed.
- The Key Phrases shown below, from the Main Data Table portion of the DES template, are those essential for clearly identifying and describing meteorological data.
- Highlighted Key Phrases are used to uniquely identify a row of data.
- DES template data provider header and metadata table Key Phrases (i.e., NARSTO flags and Site Information) do not change with data type.
- The non-essential Key Phrases may be removed from the template or their values for all columns may be set to "Not applicable".
- If other data types are also to be reported, some or all of the Key Phrases that may be omitted, will have to be included. See examples below.
- Meteorological variable names for the *TABLE COLUMN NAME key phrase are contained in the "Non-chemical variables name reference table" found at this link: http://cdiac.esd.ornl.gov/programs/NARSTO/non-Chem_variable_names.xls .

DES Key Phrases Needed to Report Meteorological Data

*TABLE NAME
*TABLE FOCUS
*TABLE EXPLANATION OF ZERO OR NEGATIVE VALUES
*TABLE EXPLANATION OF REPORTED DETECTION LIMIT VALUES
*TABLE EXPLANATION OF REPORTED UNCERTAINTY
*TABLE USER NOTE
*TABLE USER NOTE2
*TABLE KEY FIELD NAMES
*TABLE COLUMN NAME
*TABLE COLUMN NAME TYPE
*TABLE COLUMN CAS IDENTIFIER
*TABLE COLUMN EXPLANATION OF DERIVED VARIABLE

*TABLE COLUMN USER NOTE
*TABLE COLUMN USER NOTE2
*TABLE COLUMN UNITS
*TABLE COLUMN FORMAT TYPE
*TABLE COLUMN FORMAT FOR DISPLAY
*TABLE COLUMN MISSING CODE
*TABLE COLUMN LOOKUP TABLE NAME
*TABLE COLUMN OBSERVATION TYPE
*TABLE COLUMN FIELD SAMPLING OR MEASUREMENT PRINCIPLE
***** **Rows have been omitted – see below** *****
*TABLE COLUMN SAMPLING HEIGHT ABOVE GROUND (M)
***** **Rows have been omitted – see below** *****
*TABLE COLUMN INSTRUMENT NAME AND MODEL NUMBER
*TABLE COLUMN MEASUREMENT PRINCIPAL INVESTIGATOR
*TABLE COLUMN DETECTION LIMIT
*TABLE BEGINS
***** **Insert Data Rows** *****
*TABLE ENDS

Key Phrases essential for clearly identifying and describing other data types:

Add one or more of these Key Phrases if you want to report particle data.

*TABLE COLUMN PARTICLE DIAMETER--LOWER BOUND (UM)
*TABLE COLUMN PARTICLE DIAMETER--UPPER BOUND (UM)
*TABLE COLUMN PARTICLE DIAMETER--MEDIAN (UM)

Add one or more of these Key Phrases if you want to report radiation or optical property data.

*TABLE COLUMN WAVELENGTH (NM)
*TABLE COLUMN WAVELENGTH--LOWER BOUND (NM)
*TABLE COLUMN WAVELENGTH--UPPER BOUND (NM)

Add one or more of these Key Phrases if you want to report data that were measured in the field or were measured in a field sample taken to a laboratory.

*TABLE COLUMN MEDIUM
*TABLE COLUMN COATING OR ABSORBING SOLUTION/MEDIA

*TABLE COLUMN INLET TYPE
*TABLE COLUMN SAMPLING HUMIDITY OR TEMPERATURE CONTROL

*TABLE COLUMN BLANK CORRECTION
*TABLE COLUMN VOLUME STANDARDIZATION

Add one or more of these Key Phrases if you want to report data that were measured in a field sample taken to a laboratory.

*TABLE COLUMN LABORATORY ANALYTICAL METHOD
*TABLE COLUMN SAMPLE PREPARATION

Other Data Types:

SPMS (specialized)

Contact QSSC for guidance about data with special issues (e.g., MOUDI, SPMS, etc)

Archiving Data in Other Formats:

This may be a viable option. Please contact the QSSC for more information.

DM-14: Creating Archive Documentation for Your Data Sets

BACK TO TABLE

SCOPE: ***Project***

PURPOSE: To introduce project data providers to the NARSTO Data and Information Sharing Tool (DIST). DIST can be used to enter project and data set metadata and then export formatted archive documentation. Links to the DIST and examples are given.

The NARSTO Data and Information Sharing Tool has a convenient web-based metadata entry/export feature for efficient preparation of archive documentation.
https://daac.ornl.gov/cgi-bin/MDE/NARSTO//access.pl

The final output looks like this.
http://eosweb.larc.nasa.gov/GUIDE/dataset_documents/narsto_epa_ss_fresno_teom_pm_mass.html

Contact the NARSTO QSSC to obtain a user ID and password for the metadata editor.

DM-15: Creating a Searchable Index of Your Data Sets with Links to the Data Files

BACK TO TABLE

SCOPE: ***Project***

PURPOSE: To introduce project data providers to the NARSTO Data and Information Sharing Tool (DIST) DIST can be used to create a searchable metadata index of your project data sets. DIST is used to enter project and data set metadata (see DM-3a).

The NARSTO Data and Information Sharing Tool has a convenient web-based metadata entry feature. https://daac.ornl.gov/cgi-bin/MDE/NARSTO//access.pl

It enables the data provider to develop a searchable, Web-based inventory of project data using the existing ORNL metadata search and data retrieval system called Mercury. The NARSTO implementation is called the Data and Information Sharing Tool.

Mercury is a web-based, distributed system designed to allow the searching of metadata to identify data sets of interest with the capability to deliver those data sets to the user.

Data providers need not run any database software on their machines, and their data can reside in any convenient format (e.g., ASCII, spreadsheets, gif and/or jpg images, Access data sets). Data providers may make their data available by placing them in a "visible" ftp area on their machines, and periodically the Mercury system harvests the metadata and automatically builds a searchable index with the descriptive information, which resides at ORNL.

The NARSTO DIST system, located within CDIAC, facilitates "one-stop shopping" for the distributed data and information comprising the project data collection. Such a system has already been developed for NARSTO by the QSSC (http://mercury.ornl.gov/narsto/) .

DM-16: Capturing Sampling and Analysis Information – Pre- and Post-Measurement

BACK TO TABLE

Intended use: This document outline can be used by an investigator as a checklist for identifying the sampling and analysis metadata that should be collected and recorded before, during, and after field and laboratory investigations. The information will be used to report field and laboratory data. A complete document can be constructed using this template and submitted with the data as a valuable quality assurance tool.

1. Study
1.1. Study Name
1.2. Principal Investigator/Team and Agency:
1.3. Data Availability
1.4. Data completeness (explanation of large data gaps, etc)
2. Site Information
2.1. Pictures, Site layout diagram, description, lat, lon, elev
3. Measurements
3.1. Objective(s):
3.2. Species and Size Range(s) (if PM):
3.3. Units (e.g., mg/l)
3.4. Measurement Details (brief description)
3.5. Field Information
    3.5.1. Instrument Location
    3.5.2. Measurement Platform (surface, tower, airborne)
    3.5.3. Instrumentation Description (type/principle/make/model/media/coatings)
    3.5.4. Inlet Type (cyclone, 10 um SSI, filter, none, etc.)
    3.5.5. Inlet Height above Ground (and length if applicable)
3.5.6. Sampling Times/Frequency/Period
3.5.7. Sample Handling Methods
3.5.8. Nominal flow rate
3.5.9. Flow measurement/control
3.5.10. Flow Temperature/pressure Conditions
3.6. Standard Operating Procedures
4. Laboratory (if applicable)
4.1. Analytical method and instrumentation
4.2. Extraction method
4.3. Method/Instrument Detection Limit
4.4. Uncertainty Estimates
5. Measurement QA/QC Procedures
5.1. Field QA/QC
    5.1.1. Sample Handling, Documentation, Storage, Chain of Custody
    5.1.2. Traceable to (Standard Reference)
    5.1.3. Calibrations
    5.1.4. Zeroes and Spans
    5.1.5. Audits
    5.1.6. Blanks (frequency, type)
    5.1.7. Field QC (for filter methods)
    5.1.8. Precision Determination (collocation, duplication)
    5.1.9. Comparison with other methods
5.2. Laboratory QA/QC
5.2.1. Sample Handling, Documentation, Storage, Chain of Custody
5.2.2. Traceable to (Standard reference)
5.2.3. Calibrations
5.2.4. Blanks (frequency, type)
5.2.5. Other lab QC (check solutions, spikes, blinds, recoveries, etc)
5.2.6. Precision Determination (duplicates)
5.2.7. Comparison with other methods

5.2.8. Audits
6. Data Management QA/QC
6.1. Data recording method
6.2. Frequency of raw data record
6.3. Frequency of reported data record (averaging period)
6.4. Data Chain of Custody
6.5. Data QC Procedures
6.6. Validity Flags used
6.7. Calibrations or other adjustments performed on measurements
7. Data Quality Measures
7.1. Measurement Quality Objectives (Pre-campaign)
7.1.1. Anticipated Detection Limits (method/analytical) and method of determination:
7.1.2. Accuracy
7.1.3. Precision
7.1.4. Comparability
7.1.5. Representativeness
7.1.6. Completeness
7.2. Post-Campaign DQIs (Data Quality Indicators)
7.2.1. Detection Limits (method/analytical) and method of determination:
7.2.2. Accuracy
7.2.3. Precision
7.2.4. Comparability
7.2.5. Representativeness
7.2.6. Completeness
8. Blank correction (describe whether done and method used)
9. Other Quality Information
Give special information about the measurement and its quality. For example:
Filter blanks had high nitrate levels causing the rejection of many active samples
the analytical method for sulphate was found to be biased high during laboratory round-robin study (reference)
Measurements were biased high compared to collocated xxxxxx method (reference) or external audit.
References to publications related to the measurements.

++

The Reason for Including Metadata In Research Data Sets

As an atmospheric researcher, consider the needs of the future researcher who, in 10 years time, must decide whether or not to use your archived data. The expert will have to determine whether the data are suitable for a research activity, which might involve any of the following:
- comparing your data to other researchers' data collected at the same location and time,
- merging your data with data from other sites to investigate spatial patterns and variability,
- selecting the most accurate of several data sets collected at the same location and time,
- comparing your data against more recent data to look at the temporal changes,
- using your data to evaluate model predictions.

Let us assume that the researcher is not able to talk to you directly, and has access only to your data and supporting information. The question is, "What information would this future researcher need to have available, besides the actual data values, to make a proper judgment about the suitability of your data for use in his/her research?"

DM-17: Defining the Quality Level of Data

BACK TO TABLE

SCOPE: Project

PURPOSE: All project data must be quality assured and all data products should include a flag or quality statement designating the level of quality checks (i.e., validation) that have been performed and reflecting the data provider's confidence in the data. The data product quality flags generally will have values ranging from zero (0) to three (3). Initially, data will be at the lowest level of quality. In response to further quality checks, the quality flag will be upgraded. Resources, examples, and use in the DES template are shown.

All data products submitted to the NARSTO data archive must have been validated to quality level 1 and preferably level 2. All data products should include a flag and/or statement designating the level of validation. (Within a data product (i.e., data set) submitted to the data archive there should be no invalid values. Individual invalid values should be replaced by a missing value code and flagged appropriately.)

DATA VALIDATION

Additional validation activity descriptions, beyond these general guidelines, should be in the individual project's QA plan.

Data validation is the process of determining and denoting the quality of a data set (data having either a common method of collection or data collected by various methods in one location). The validation process consists of evaluating the internal, spatial, temporal and physical consistency of each data set for invalid data and for outliers (data that are physically, spatially, or temporally inconsistent). During validation, physically unrealistic data are invalidated, biases and instrumental drift are noted, and gross errors are identified. The objective of this process is to produce products with values that are of known quality.

All data products must be validated. Recognizing the potential for your project to generate research products other than just numeric measurement data is important. Other possible products are simulation models, methods, procedures, and reports. Each of these products must be validated. A validation level and status discussion must be included in the metadata record or information associated with the data or research product.

*QUALITY CONTROL LEVEL as required in the DES

0 (a reasonably complete data set of unspecified quality that consists of research products subjected to minimum processing in the field and/or in the laboratory by project staff.)

1 (a complete data set of specified quality that consists of research products subjected to quality assurance and quality control checks and data management procedures.)

2 (a complete, externally consistent data set of specified quality that consists of research products that have undergone interpretative and diagnostic analyses by the project staff or user community.)

3 (data that have received intense scrutiny through analysis or use in modeling.)

9 (data submitted with no indication of the quality control level.)

Indicating the quality level of the data being reported in the DES (select one)

*DATA EXCHANGE STANDARD VERSION	NARSTO 2005/04/29 (2.302)
*COMMENT	Further instructions on how to fill in this template are provided at:
*COMMENT	http://cdiac.ornl.gov/programs/NARSTO/narsto.html
***QUALITY CONTROL LEVEL**	0 (a reasonably complete data set of unspecified quality that consists of research products subjected to minimum processing in the field and/or in the laboratory by project staff.) 1 (a complete data set of specified quality that consists of research products subjected to quality assurance and quality control checks and data management procedures.) 2 (a complete, externally consistent data set of specified quality that consists of research products that have undergone interpretative and diagnostic analyses by the project staff or user community.) 3 (data that have received intense scrutiny through analysis or use in modeling.) 9 (data submitted with no indication of the quality control level.)
*DATE THIS FILE GENERATED/ARCHIVE VERSION NUMBER	2005/03/01

Assigning Quality Levels from a QC Validation Perspective

Level 0 Validation indicates a reasonably complete data set of unspecified quality that consists of research products subjected to minimum processing in the field and/or in the laboratory by project staff.

- Level 0 designations will be given to raw data and other research products that have not been audited or peer reviewed.
- Level 0 data contain all available measurement data and may also contain data in the form of quality control checks and flags indicating missing or invalid data.
- Level 0 data consist of instrument outputs expressed in engineering units using nominal calibrations. Missing data from on-site backup loggers or strip charts have been filled in.
- Level 0 data may include flags indicating QC check data, power failures, excessive rate-of-change, and insufficient data for the averaging period, or other logger programmed occurrences.

- Level 0 status continues until all QC checks and peer reviews associated with Level 1 validation of the product have been completed and the investigator's response recorded.

Level 1 Validation indicates a complete data set of specified quality that consists of research products subjected to quality assurance and quality control checks and data management procedures.

- As part of the Level 1 process, site documentation is reviewed for completeness and performance compared with other locations.
- Compliance with documented data quality objectives, standard operating procedures (SOPs), and research protocols is evaluated in the Level 1 process.
- Audit and peer review reports have been evaluated (and necessary corrections made) for all research products designated Level 1. The comparison and cross-checking activities done under Level 1 may be conducted by the project staff, and/or the scientific community.
- Level 1 data are generated by project groups. In response to audits, data may have been adjusted.
- The project group, responsible for submitting the data, will adjust the data for "blank bias" (lab analyses) or "zero drift" (continuous ambient measurements), will determine precision and accuracy, and will perform consistency checks with other data within the same data set.
- These internal consistency checks might include diurnal analyses to look for expected patterns or time series analyses to detect outliers, extreme values, or time periods with too little or too much variation.
- Level 1 designation will be assigned after the project group has performed all quality control activities and addressed all quality issues stemming from audits and reviews.

Level 2 Validation indicates a complete, externally consistent data set of specified quality that consists of research products that have undergone interpretative and diagnostic analysis by the project staff or user community. A validation level and status discussion must be included in the metadata record associated with the research product.

- Level 2 data have been closely examined by the data manager and/or data users for external consistency when compared to other related data sets.
- External checks might include correlation by scattergram, comparison of data with other similar data for the same time period, and comparison of a measurement made by two different methods (e.g., hydrocarbons by auto-GC and canister).
- If comparisons are not within the precision of the measurements, then measurement records and other information will be reviewed.
- If a check of measurement records uncovers a process error, the value will be corrected or invalidated. If such errors are not found, then an annotation will be entered.
- If the value is invalidated, it should be flagged appropriately and identified per project procedures. The value could be deleted from the active project database and replaced by a missing value and flagged appropriately.

- A record of changes will be permanently retained.
- Level 2 designation will be assigned after the project data manager and/or data users have performed comparative tests and addressed the quality issues and the project staff have evaluated the test results and supporting QA documents. Authority for Level 2 designation lies at the project level.

Level 3 Validation consists of data that have received intense scrutiny through analysis or use in modeling. As analysis of the data proceeds, analysts may raise questions about portions of the Level 2 data set.

- Additional checks and tests will be performed on such data and the Level 3 code will be affixed to data passing these tests.
- If this scrutiny reveals an inconsistency that appears to be caused by a measurement error, the entire chain of evidence for the measurement will be reviewed. This includes reviewing site logs and quality control test data as well as reviewing performance audit results and any other relevant documents.
- The data users will recommend a Level 3 designation to project staff on the basis of the reevaluations.
- Alterations to the data validation codes, if warranted, will be made by the project data manager.

Data Quality Codes for External Data Sets

Selected sets of external data, such as data from the EPA AIRS and National Weather Service, might be included in with project data. Data quality codes may be assigned to external data sets, or select fragments of external data sets using the same conventions described above. For external data sets not evaluated by the projects, but archived with project data, users should pay close attention to the data validation procedures and designations used by those networks before using the data for a particular research application.

Revisions to Data Quality Codes

Revisions to flags will be made at the project level and records of these changes will be maintained. When upgraded supplemental data are received or when additional validation work is performed on existing data, the project will send the archive a revised data set reflecting the new quality status. Details on validation activities and actions may be included with the metadata. Because not all activities are performed concurrently, at any given time the archive may contain data at different validation levels.

Submitting Quality Assured Data the NARSTO Archive

All data submitted to the NARSTO data archive must have been validated to quality Level 1 and preferably Level 2. All data products should include a flag and/or statement designating the level of validation.

DM-18: Day-to-Day Operation of Data Management Systems

BACK TO TABLE

SCOPE: Project

PURPOSE: Provide guidance to investigators, technicians, instrument operators, and project data managers responsible for the day to day operation of data collection and data management systems including: backups; access and security; data entry, transfer, transformation; and data control.

These routine data management protocols can be facilitated through checklists or worksheets (electronic or hardcopy) that aid completing project documentation. An ounce of prevention…

System Backups

Project data should be protected from loss through preventative data system backup and recovery mechanisms. Data system backups should be performed on a periodic basis at a frequency to be defined by each project. This frequency should be selected to minimize the extent of consequences of data loss and time required for data recovery. Recovery procedures should be developed and documented in preparation for the event of hardware or software failure.

Data System and Database Access

Projects should protect systems and data from unauthorized access by implementation of administrative and procedural controls. Access controls should be managed based upon specific data user roles that are defined by the types of data and functionality required (e.g., a data management specialist needs the capability to create and update data while a program manager may need read-only access to perform on-line queries). The mechanism for implementing access control should be documented in project data management plans. Maintaining up to date computer security, including operating system patches, and as applicable antivirus and antispyware software on project data systems is essential.

Data Entry, Transfer, and Transformation

Data entry, transfer, and transformation activities should be verified to ensure that data integrity is maintained. This includes movement/copying of data from one storage medium to another and transformation from one format to another. All data, including analytical data produced and reported by a laboratory should be verified.

This verification encompasses all data recording media, handwritten or hard copy produced via electronic means, as well as electronically stored, such as in a database.

It also includes all data collection methods (e.g., electronic collection through real-time monitoring instrumentation, bar coding equipment, and handwritten log entries).
If a data transformation or transfer activity has occurred before receipt of the data by project personnel (i.e., between creation and final reporting), the verification may be performed by the reporting party but only if sufficient evidence to support the validity of the process can be provided by the reporting party. For example, if a laboratory technician captures data from a laboratory instrument and records it in a logbook, enters the data from the logbook into an electronic data deliverable format, and then transfers the data to the project, the verification process may be performed by the laboratory. The mechanism for a project's data entry, transfer, and transformation verification processes should be documented.

Database Content Configuration Control

A project should establish configuration control requirements for the contents of the project database. The requirements should ensure traceability of field and laboratory data from the original reported values, through authorized data changes, to current values stored in the database. The configuration control should define the approval process required for making changes to the database and the documentation required for each database change.

The minimum information maintained for each database change should include

- a description of the change;
- the reason for the change;
- the name of the individual making the change;
- the date of the change; and
- a copy of the data before the change took place.

Identification of Data Products

Practices should be established to assure that all data and data products are clearly identifiable and traceable to the project from which they were produced. It is very important that this identification and traceability be maintained (protected) throughout the needed lifetime of the data. A description of practices to be used on your project should be included in project plans.

Control of Erroneous Data

Practices are needed for controlling data that are erroneous, rejected, superseded, or otherwise unsuited for their intended use. These practices should provide for the identification, flagging, and/or segregation of inadequate data to avoid their inadvertent use. Project plans should describe the practices for controlling invalid data in your project data systems.

While maintaining clearly designated invalid data values may be of value to the project, no invalid values should be sent to the NARSTO archive. Before archiving, invalid

values should be replaced with the appropriate missing value codes and flagged accordingly.

DM-19: Managing Electronic and Hardcopy Format Project Records

BACK TO TABLE

SCOPE: Project

PURPOSE: To ensure that your Project maintains raw data, computer codes, models, and hardcopy records of their data collection and generation activities for quality assurance purposes.

Project records, both electronic and hardcopy, should be specified, prepared, reviewed, and maintained to document the quality for the work completed. Records are completed documents that provide objective evidence of the quality of an item or process. Project plans should state requirements and responsibilities for record transmittal, distribution, retention, protection, preservation, traceability, disposition, and retrievability. The project plan should also identify how the disposition of records, in accordance with regulatory requirements, schedules, or directives from senior management, will be accomplished.

Unprocessed/Raw Measurement Data in Electronic Format

Projects should identify the raw or minimally processed (level 0) measurement data that are recorded by their instruments and plan to store this data for a period of time defined by the project – usually a minimum of 5 years.

Computer codes, models, input and output data sets

Projects should identify the specific software codes used to process measurement data and specific versions of model codes plus their input and output data sets that were used to generate a specific product or publication and plan to store this information for a period of time defined by the project – usually a minimum of 5 years. Consider archiving model codes and input and output data sets.

Hardcopy Format Records

Paper records will **not** be sent to the NARSTO PDA for long-term archival. Conversion to electronic media should be considered. It is the responsibility of the projects to establish a project filing and index system, and storage location, for project records. Projects should determine the best filing structure and indexing mechanism to meet project needs. Projects are encouraged to consolidate data records storage, determine an appropriate tool for maintaining an index, and give careful thought to the records storage area (e.g., limited access, environmental conditions suitable for short-term records storage, administrative controls). Development of data forms and logbooks should be a controlled process.

DM-20: Data Management System and Software Configuration Control Guidelines

BACK TO TABLE

SCOPE: Project

PURPOSE: Provide guidance to project data managers responsible for the documentation, quality assurance, and configuration control of project software and data systems.

This software and computer system implementation guidance is applicable to projects using project specific software and an electronic database. The need for project-specific data systems, databases, and software will vary depending on the scale of the project. This section discusses minimum documentation, QA, and configuration control guidelines for project specific implementations.

Project Database Documentation

Project specific databases include spreadsheets, data sets, and databases (e.g., Excel, ORACLE) defined by investigators and the project data management group to manage project data. The project specific databases should be described in the permanent project record. The description should identify the commercial database product used, the database name, structure, and locations.

The minimum database documentation will consist of

- name and version of commercial software used;
- names of project databases created;
- database structure definitions, including field names and descriptions; and
- storage location and media.

Project Software Documentation

Project specific software includes programs written by investigators, technicians, and the project data management group for data management tasks, and applications written for the production of data products. Data management tasks could include instrument data acquisition and processing, data conversions and derivations, and data quality control checks. Data products are defined as any extraction, summary, or analysis of data that results in a data summary or a hard copy product such as tables, graphs, statistics, or maps.

Software documentation should include the software program name, description, special requirements, author, revision, completion date, and documentation of the QA review. Data products documentation should also include all information to uniquely describe how the data product was produced, including the sources used, the manipulations made, and the tools used to produce the data product. Software documentation can be maintained in electronic or hard copy format or it may be included as comment blocks embedded within the project software program.

The minimum software documentation should consist of

- name and version of the commercial software used;
- name and version of the software program written by the project;
- author;
- date;
- revision;
- system requirements; and
- storage location.

Project Software Quality Assurance

The project should define the QA requirements for project specific software. At a minimum, software programs for data acquisition and processing, data conversions and derivations, data loading, data quality control checks, calculating statistics reported in project deliverables, and producing data products should be reviewed to ensure they meet the desired objectives. The reviewer should be someone other than the person who wrote the software program.

Project Specific Software Configuration Control

Project specific software should be protected from unauthorized modification or deletion. This can be accomplished by administrative controls or file security options provided by many computer operating systems. Changes to project software should be documented and a history of revisions which impact the results or data products should be included in the project file. Commercial products are available to maintain a record of software revisions [e.g., Revision Control Software (RCS)]. Another way to do this is to keep the initial or baseline software in a storage area separate from the working software. Then, when the software changes, the new software can be moved to this separate area also, there maintaining copies of all revisions. Project specific configuration and revision control should be documented in project plans.

The project software configuration control documentation should include

- commercial software used;
- program names;
- approvals
- revisions (including dates of revision); and
- storage locations.

Before the development of software applications, a requirements analysis should be conducted. Developed software applications should be tested and validated to ensure compliance with all user requirements and to provide confidence that the software will perform satisfactorily in service. The technical adequacy of results generated by these applications should also be reviewed by another person, tested and validated. Configuration management of the developed software application programs shall be conducted.