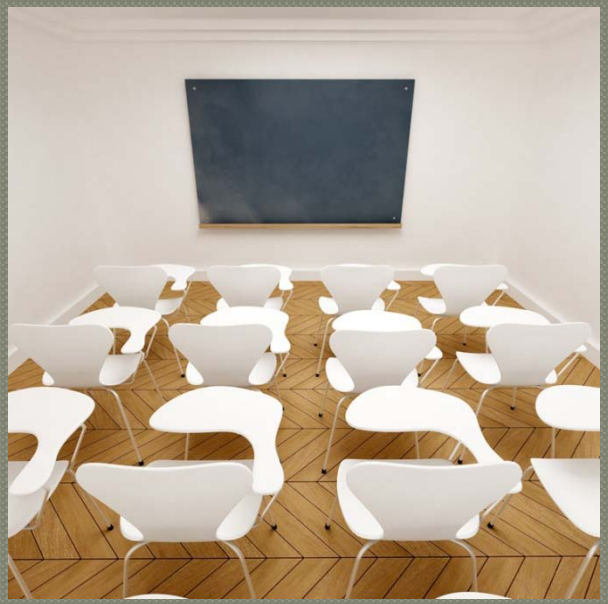
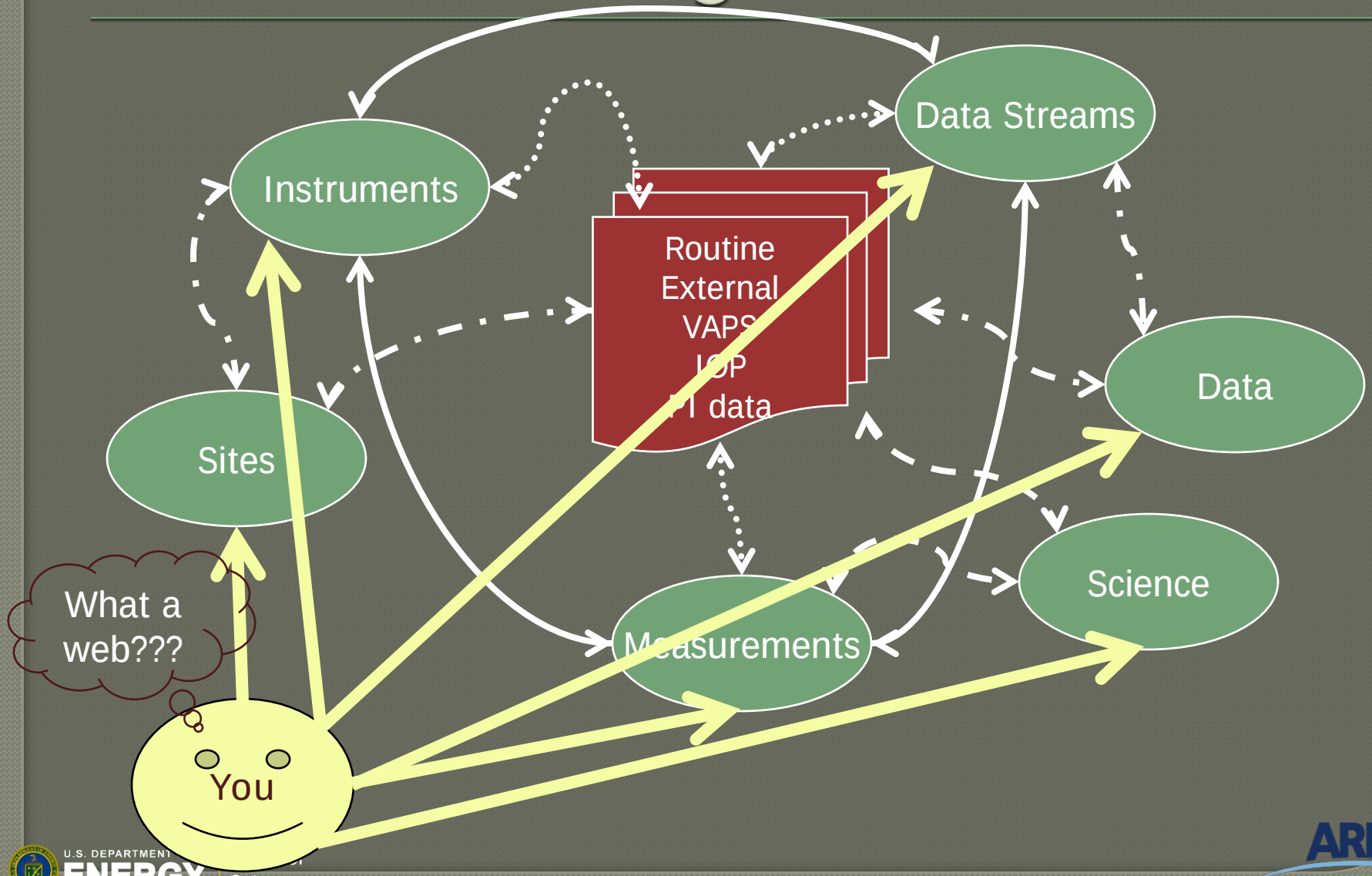
Obtaining Data from the ARM Archive

Orientation for new Science Team Members



Wandering ARM Web "stuff"



One click to the ARM Data Archive

ARM
CLIMATE RESEARCH FACILITY

CART | Home | People | Site Index | Search arm.gov

U.S. DEPARTMENT OF ENERGY

About | Science | Campaigns | Sites | Instruments | Measurements | Data | News | Publications | Education

The ARM Climate Research Facility is a U.S. Department of Energy scientific user facility for the study of global climate change by the national and international research community.

1 2 3 4 5 6 7 8 9 10

SITE: Barrow (NSA C1), Atqasuk (NSA C2)
OPERATION: Jul 1997 - present
INSTRUMENTS: 25
DATASTREAMS: 180
North Slope of Alaska

[Full Kiosk]

FEATURED DATA

11.12.2009
Updated Evaluation Product

FEATURE 01.06.2010

Cirrus Clouds Hold Clues to Climate

On January 4, scientists sponsored by ARM began a five-month aircraft campaign to gather data from cirrus clouds in the skies above Oklahoma. Using an instrumented Learjet 25 research aircraft, their goal is to obtain a new and comprehensive set of in-cloud, or "in situ," measurements about the size and number of ice crystals that make up cirrus clouds.
» Read More

VIEW ALL FEATURES

Recovery Act
Learn about ARM's efforts.

USING OUR FACILITIES

Preproposals for FY2012 for AMF and AAF deployments **due February 16**. Successful candidates will be eligible for up to **$150K** per year for three years of research funding from Atmospheric System Research. To learn more, see announcements.

FIELD CAMPAIGNS

SPARTICUS
Small Particles in Cirrus

RHUBC-II
Chile: Radiative Heating in Underexplored Bands Campaign 2

CAP-MBL
Azores: Clouds, Aerosol and Precipitation in the Marine Boundary Layer

ARM Data Archive Home Page

Get routine ARM data

Data Browser [?]

Select datastreams, view quality information about the data and order data files with the Data Browser. The "Novice Interface" guides new users through the process, while the "Datastream Interface" is designed for users experienced with ARM data.

Data Cart [?]

Browse ARM website pages to find datastreams of interest to place in the Archive data cart. This can be done by clicking "Build an Order" from any instrument, measurement, datastream, or VAP page.

Catalog Browser [?]

The catalog based user interface presents, in an interactive sequence of tab
of available data files organized in a way that will be useful to the inexperie
Archive user.

Thumbnail Browser [?]

View prepared plots of data to quickly find data of interest to you. The thum
measurement type and date range selections to retrieve data plot thumbnai
You can also download high-resolution images of the data plots, or downloa

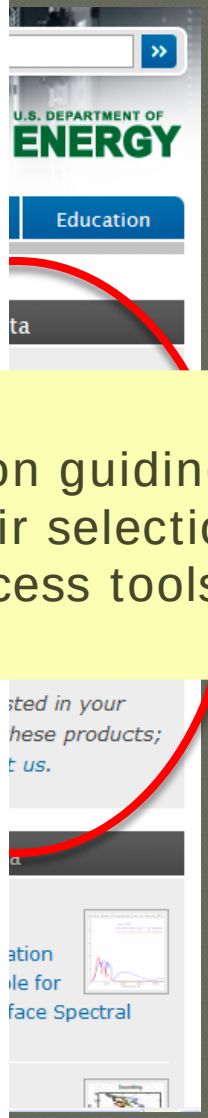
Statistical Browser [?]

Users select a location and measurement and then drill down through time scales ranging from the full period of record to individual months. In addition to viewing graphs displayed by this interface, access to extractions of data behind the statistical graphs, obtain the measurements that were used for the statistics, or place the order for related ARM data files.

Get special data

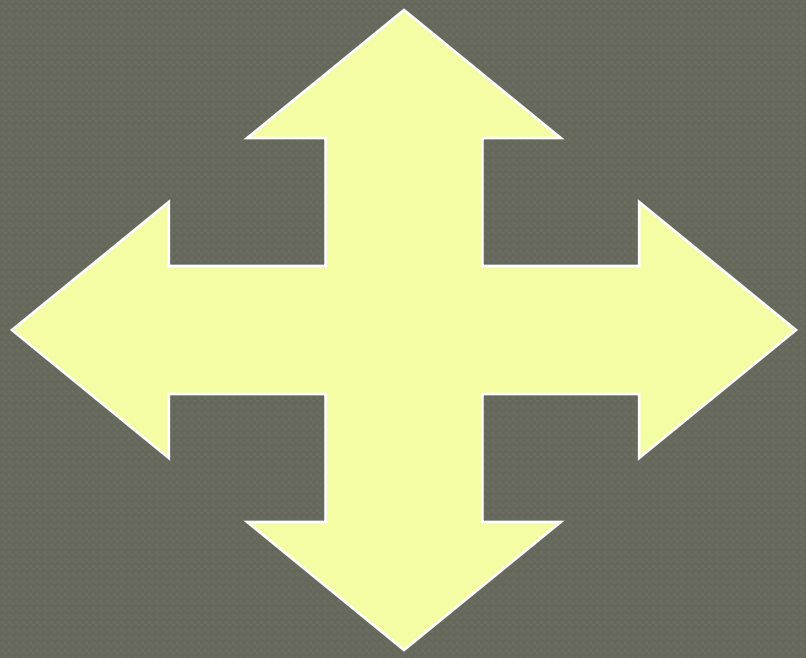
IOP Data **PI Data** **Showcase Data** [?]

Browse and download data generated from ARM Intensive Operation Periods or "IOPs". Data is stored in a directory structure organized by year, site, IOP name and instrument. A README file is included in each directory to provide documentation.



Emphasis on guiding users in their selection of data access tools

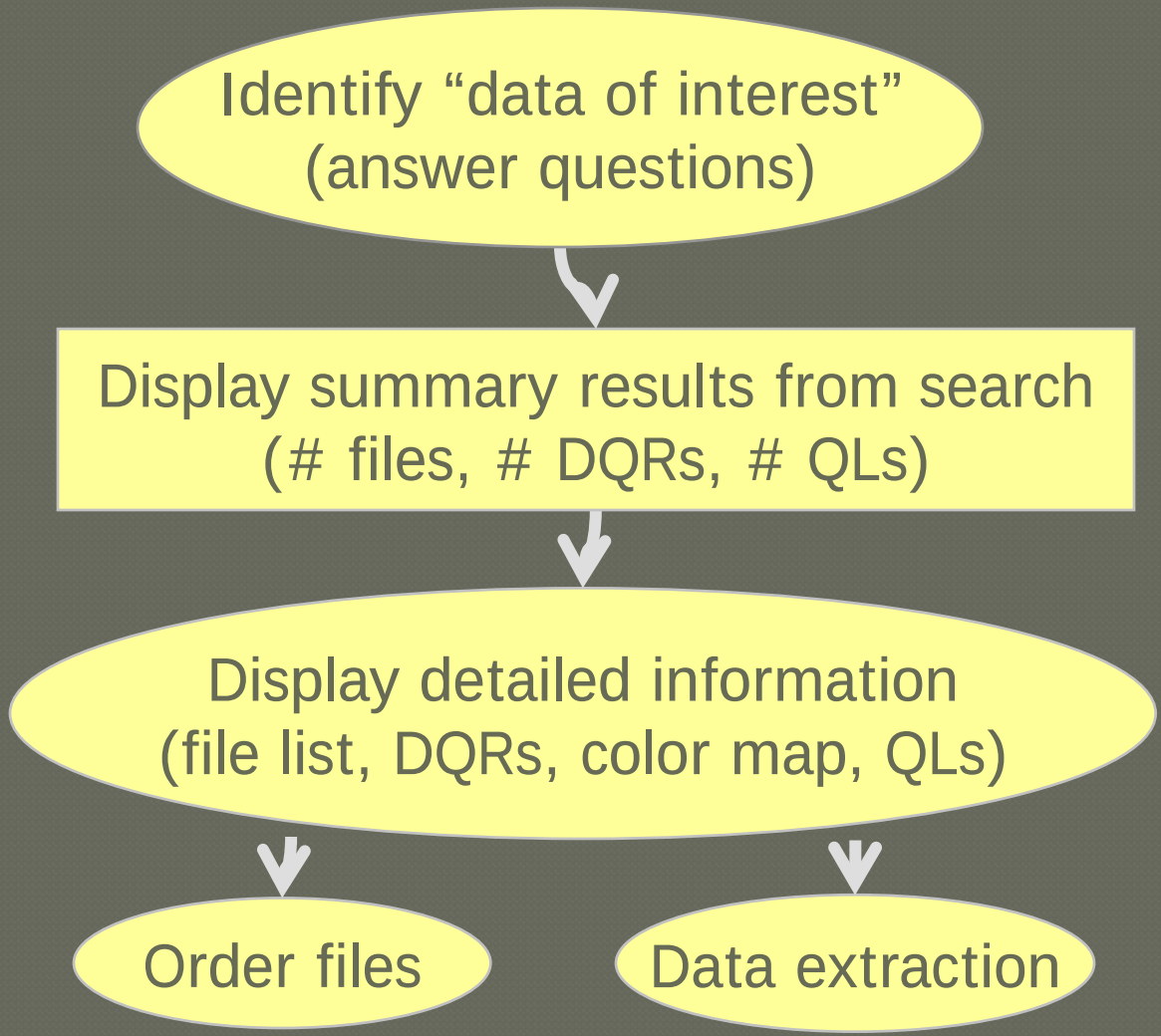
Accessing ARM Data: Options



Comparison of Browser/Interface Options

Interface name	Accessible data	"*Ordering*" approach (armarchive@ornl.gov, 1-888-ARM-DATA)
Data Cart	Routine ARM data and some IOP data	*"I need to read about what you have, then I will decide."* Discover areas of interest by browsing the ARM web documentation and collect items of interest.
Data Browser	Routine ARM data	*"I know what I want. Do you have it?"* Searching with predefined selection criteria.
Catalog Interface	Routine ARM data	*"I am not sure what I want. I need to see what you have available."* Browsing a hierarchy of availability summaries.
Thumbnail Browser	Most routine ARM data	*"I will know what I want when I see it."* Searching with a combination of predefined selection criteria and visual review of data plots
NCVWeb	Routine ARM Data	*"I want to see my own data plot."* Interactive data plotting tool with visualizing, extracting, statistics generation capabilities.
Statistical Browser	Special Data (CMBE, QCRAD, CONSTRVARANA)	*"I' need to see climatological summary of cloud and radiation data at ARM sites, then I'll drill down further."* Gain insight via statistical plots at the main sites for various time periods. Download the statistics, measurements and data files that went into them.
IOP, PI, and Showcase Data Browser	IOP, PI, and Showcase beta data	*"I need to look in the odd parts bin."* Direct access to IOP data. Navigate /year/site/iop directory tree. Also use narrow Google search.

Typical Logic behind Data Access Tools, Browsers, and Interfaces



General flow of User Interface

- Create account (new users only)
- Login to interface
- Formulate data selection
- Accept or reject results - submit request
- Logoff or do another request
- Wait for e-mail notification
- Copy files using anonymous FTP
- Search ARM web for ancillary information

Who are you? Who wants to know?

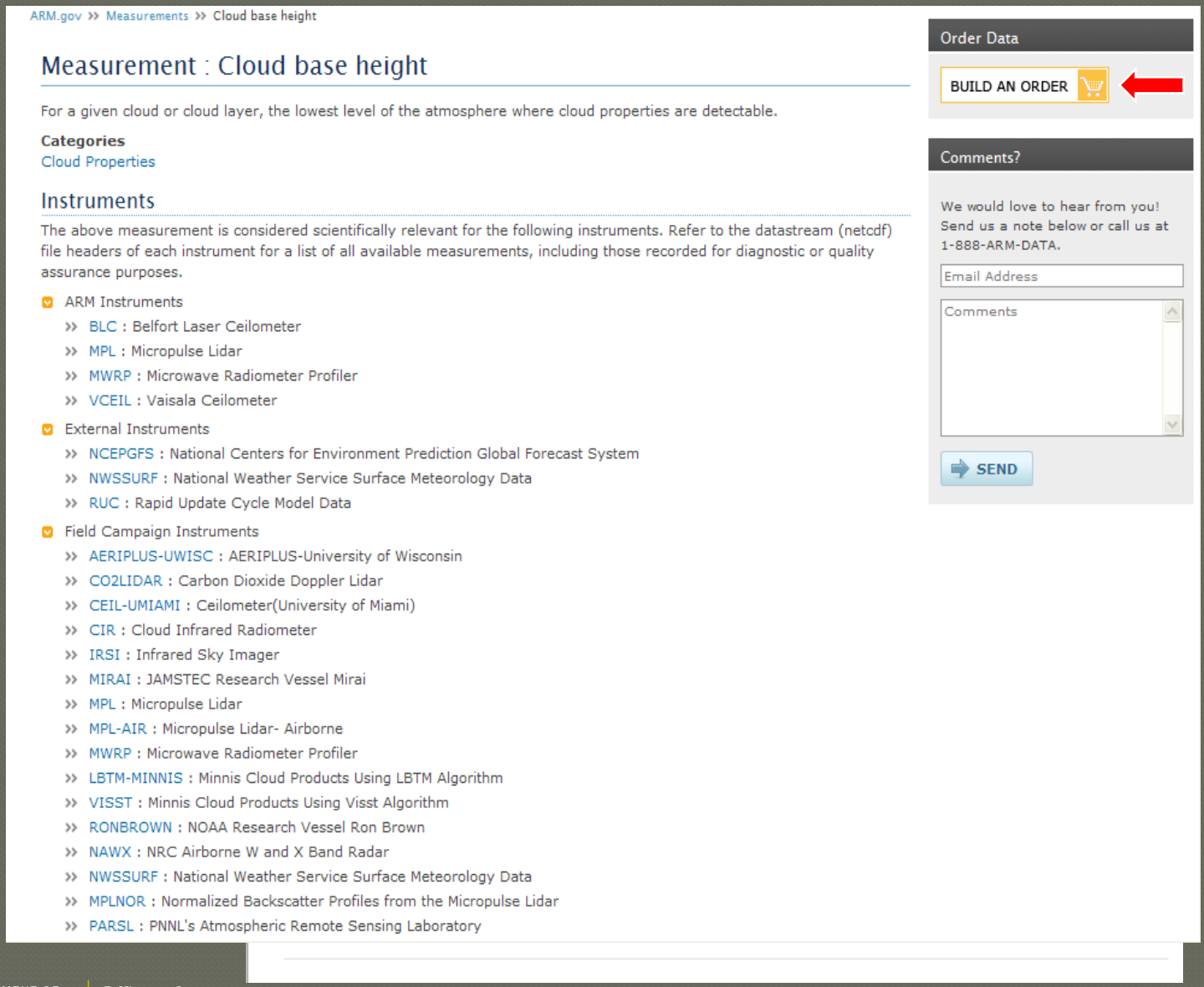
- Archive users must register.
- Notification helps you with data access.
- ARM infrastructure is a "National User Facility"
  - provides access to extra budget!!
  - OMB requests User Facilities to report user statistics for several "demographic categories"
- Some personal information is required*
  - *personal information is not reported individually and is accessible only to Archive staff after entry

"How personal??"

- Archive account creation requires:
  - Name
  - E-mail address
  - Mail address
  - Institution name
  - Institution type
  - Education status (student, graduate student, etc.)
  - Gender*
  - Ethnicity*
  - Citizenship*
  - ARM role and function

*"decline to state" option available for these items

ARM Archive – Data Access Examples



ARM.gov » Measurements » Cloud base height

Measurement : Cloud base height

For a given cloud or cloud layer, the lowest level of the atmosphere where cloud properties are detectable.

Categories
Cloud Properties

Instruments

The above measurement is considered scientifically relevant for the following instruments. Refer to the datastream (netcdf) file headers of each instrument for a list of all available measurements, including those recorded for diagnostic or quality assurance purposes.

- ARM Instruments
  - BLC : Belfort Laser Ceilometer
  - MPL : Micropulse Lidar
  - MWRP : Microwave Radiometer Profiler
  - VCEIL : Vaisala Ceilometer
- External Instruments
  - NCEPGFS : National Centers for Environment Prediction Global Forecast System
  - NWSSURF : National Weather Service Surface Meteorology Data
  - RUC : Rapid Update Cycle Model Data
- Field Campaign Instruments
  - AERIPLUS-UWISC : AERIPLUS-University of Wisconsin
  - CO2LIDAR : Carbon Dioxide Doppler Lidar
  - CEIL-UMIAMI : Ceilometer(University of Miami)
  - CIR : Cloud Infrared Radiometer
  - IRSI : Infrared Sky Imager
  - MIRAI : JAMSTEC Research Vessel Mirai
  - MPL : Micropulse Lidar
  - MPL-AIR : Micropulse Lidar- Airborne
  - MWRP : Microwave Radiometer Profiler
  - LBTM-MINNIS : Minnis Cloud Products Using LBTM Algorithm
  - VISST : Minnis Cloud Products Using Visst Algorithm
  - RONBROWN : NOAA Research Vessel Ron Brown
  - NAWX : NRC Airborne W and X Band Radar
  - NWSSURF : National Weather Service Surface Meteorology Data
  - MPLNOR : Normalized Backscatter Profiles from the Micropulse Lidar
  - PARSL : PNNL's Atmospheric Remote Sensing Laboratory

Order Data

BUILD AN ORDER

Comments?

We would love to hear from you! Send us a note below or call us at 1-888-ARM-DATA.

Email Address

Comments

SEND

ARM Archive – Data Browser Example

ARM CLIMATE RESEARCH FACILITY

ARM Data Browser

My Account | Log out

Home | Site | Date Range | Data Stream | Data Selection Summary

Select ARM Site

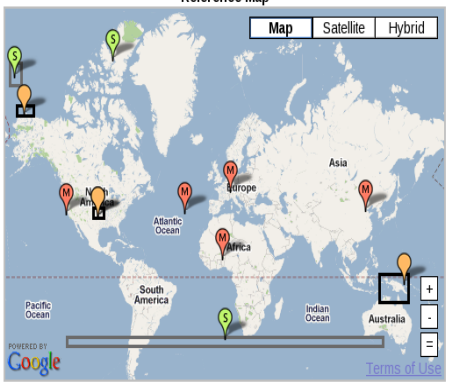
Note: You will be asked to make some selections to establish your search criteria for ordering ARM data or viewing data quick looks and quality information

Please select an ARM Site [?] from the following list

- Black Forest, Germany; Mobile Facility
- Eureka, Ellesmere Is., Canada; for NOAA SEARCH project
- Global Earth Coverage
- Graciosa Island, Azores, Portugal; Mobile Facility
- Niamey, Niger; Mobile Facility
- North Slope Alaska
- Point Reyes CA, USA; Mobile Facility
- SHEBA (Surface HEat Budget of the Arctic)
- Shouxian, Anhui, China; Mobile Facility
- Southern Great Plains
- Southern Great Plains [development]
- Tropical Western Pacific

NEXT | RESET

Reference Map



Navigation

Site
Date Range
Data Stream
Summary Page
Place Order

Questions/Comments?

We would love to hear from you! Send us a note below or call 1-888-ARM-DATA.

Email Address

Submit

Interface Help

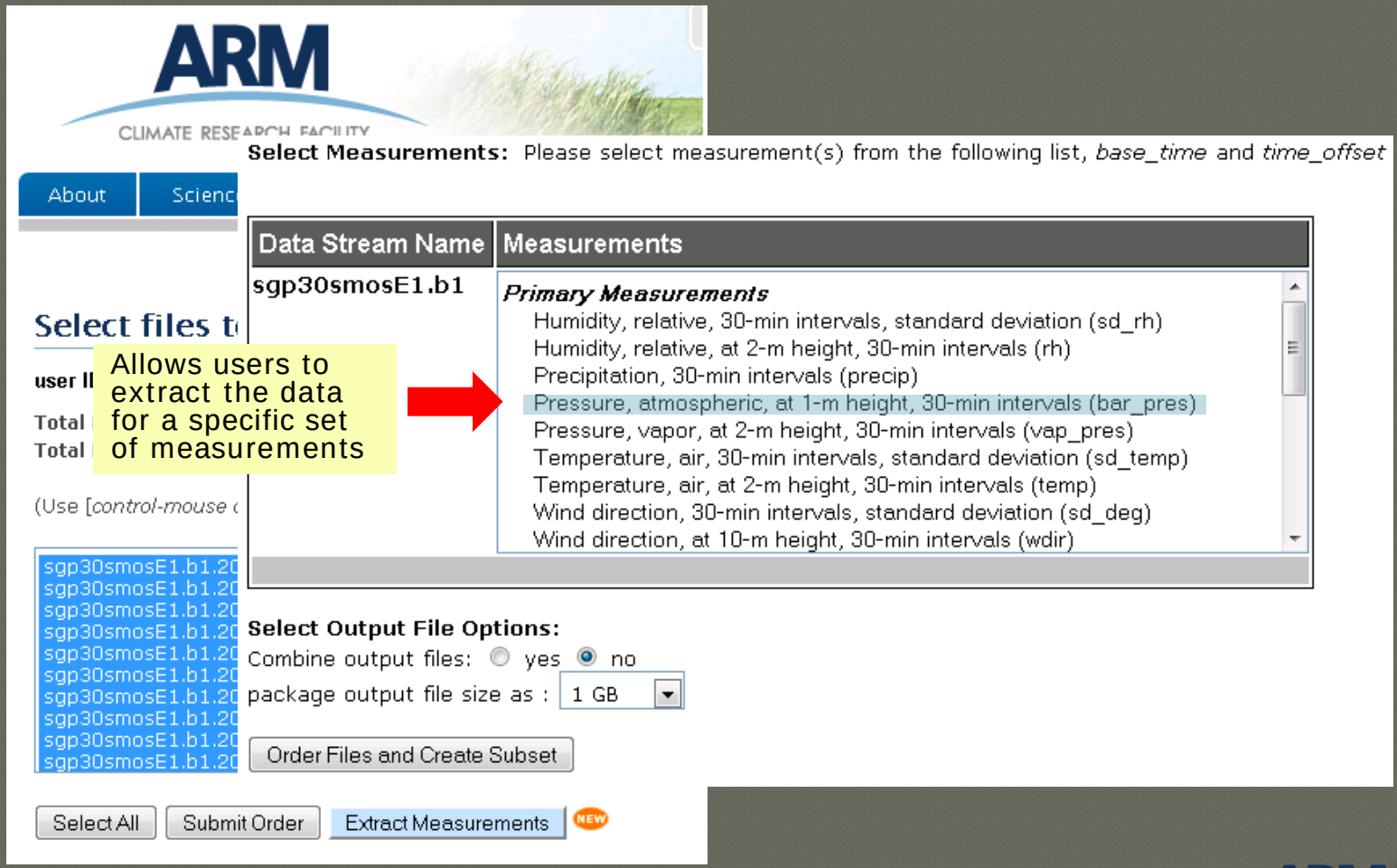
- View interface help documentation

ARM Documentation

- Global Earth Coverage (GEC)
- Niamey, Niger: Mobile Facility (NIM)
- North Slope Alaska (NSA)
- Point Reyes CA, USA: Mobile Facility (PYE)
- Richland WA, USA: Mobile Facility test (RLD)
- SHEBA (Surface HEat Budget of the Arctic) (SHB)

Data Extraction Capabilities

New! Available in all interfaces



Field Campaign Data: ARM provides two options

PRELIMINARY DATA SHARING SYSTEM

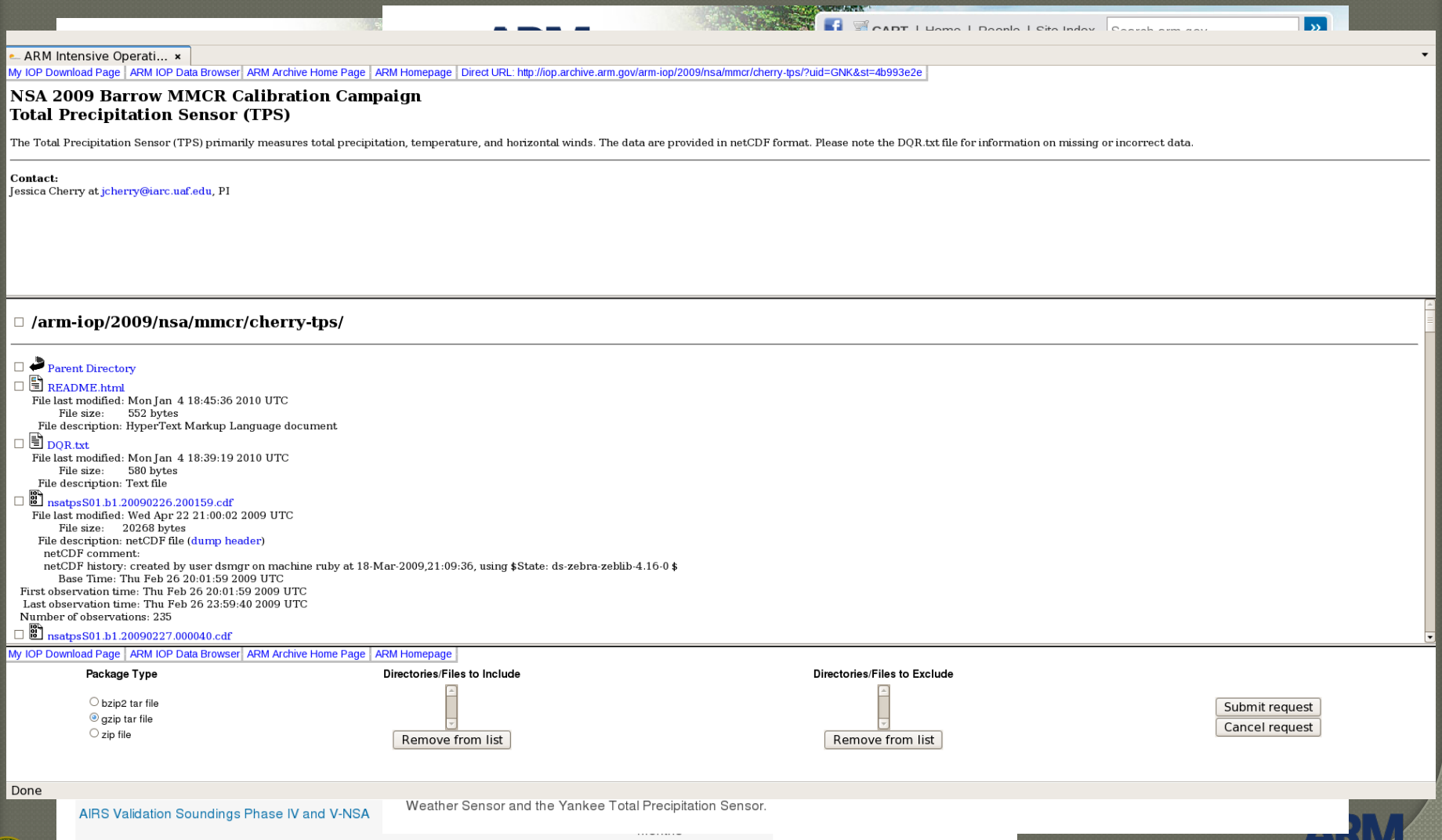
- Provides restricted access distribution during operations and initial data review
  - FTP only access via accounts with passwords
  - No usage tracking
- Provides an option for data review and delivery to ARM in a shared working space
  - Create, write, read, NO delete
- Allows very flexible data structures and documentation
  - Start as "user defined"
  - Evolve to "ARM defined"
- NOT for permanent storage

FINAL DATA SYSTEM

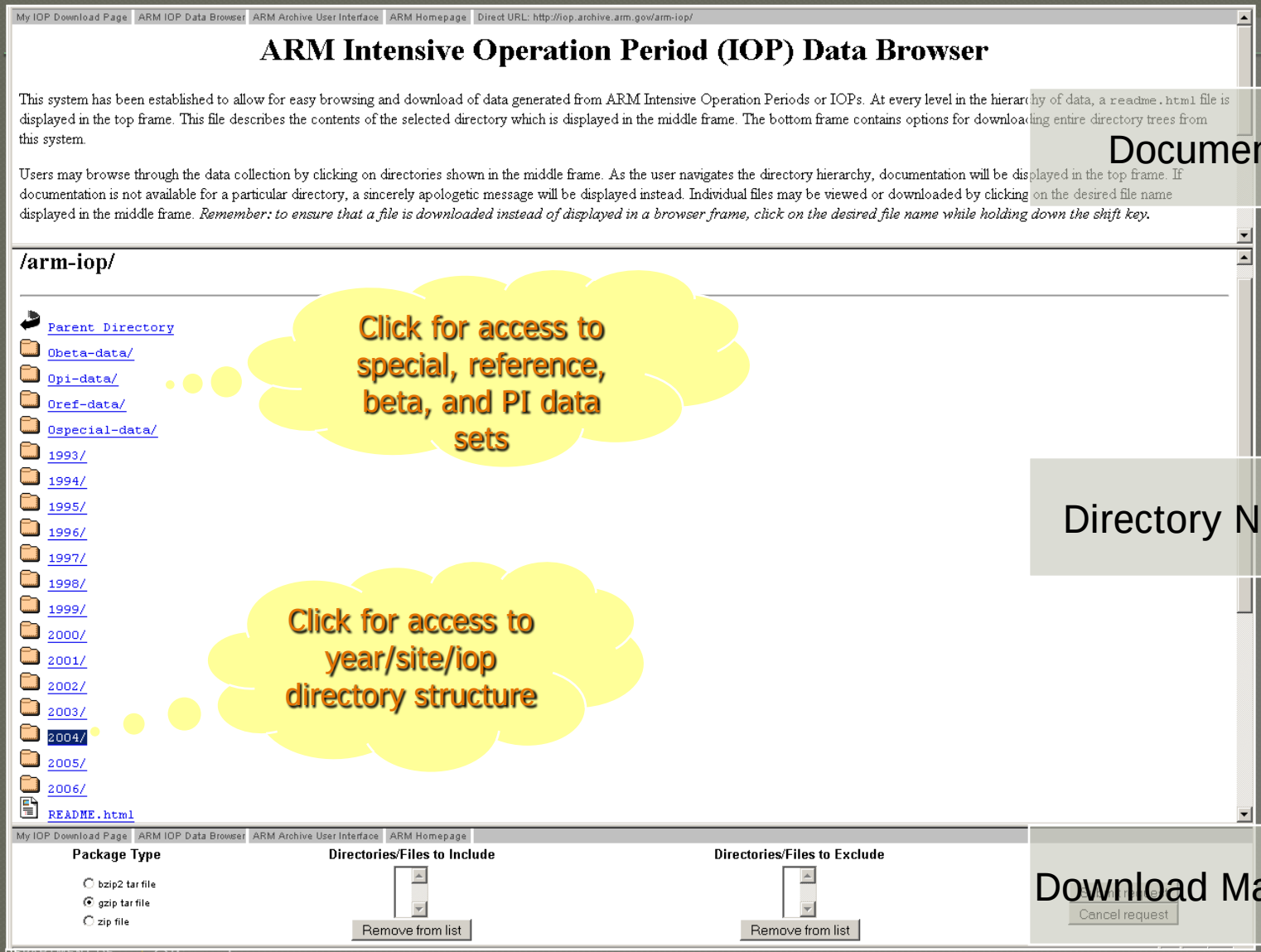
- Provides open access to all Archive users
  - Discovery is also linked to full web descriptions for Campaigns
  - Requires introductory documentation about campaigns, instruments, subdirectory and filename logic
    - README.html file required and displayed for each subdirectory
  - Usage is tracked and can be reported
  - For permanent storage

Data submission and distribution requirements arranged during Field Campaign Proposal process

Field Campaign Data



IOP Data Browser – "home page"

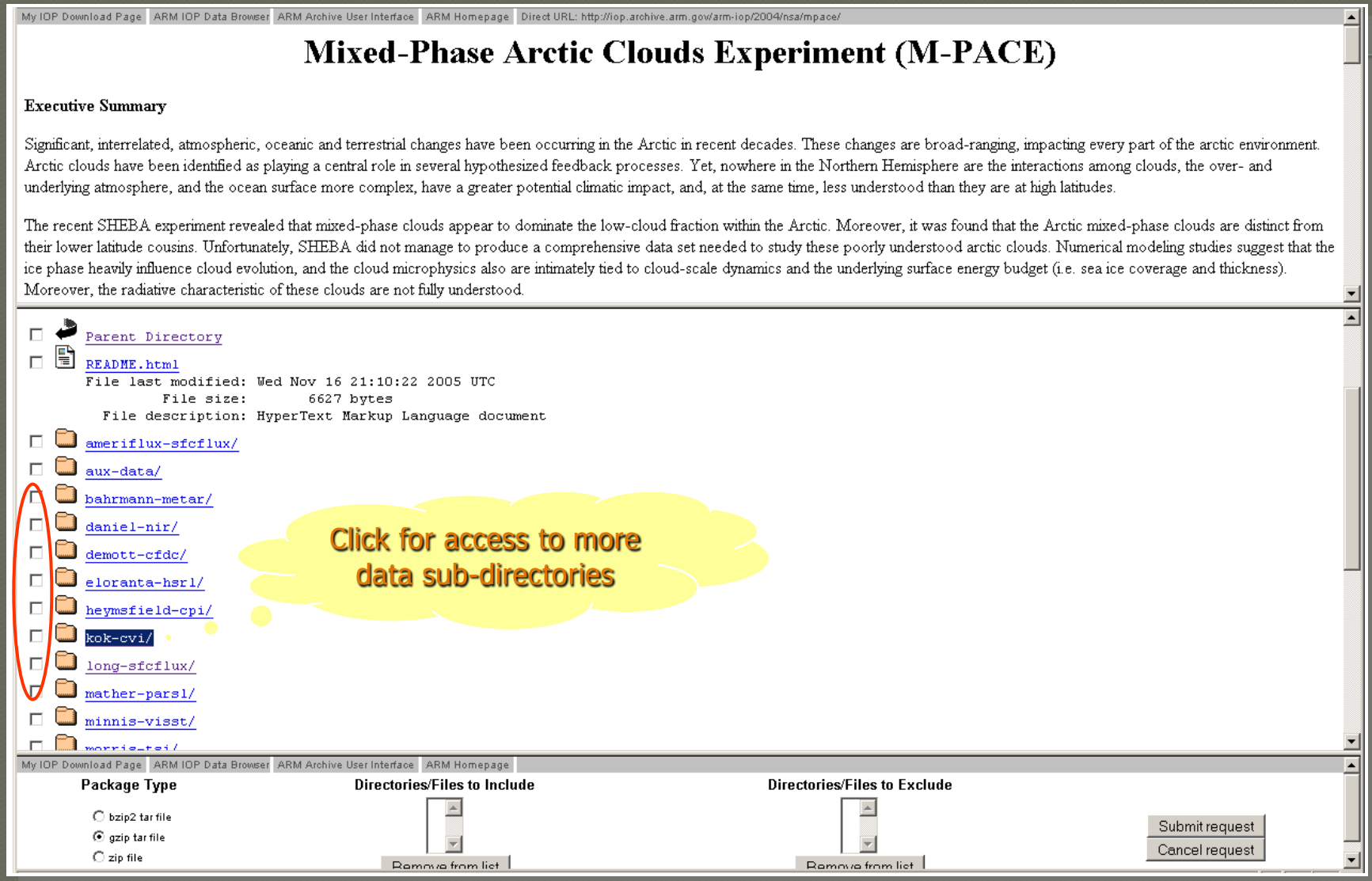


Documentation

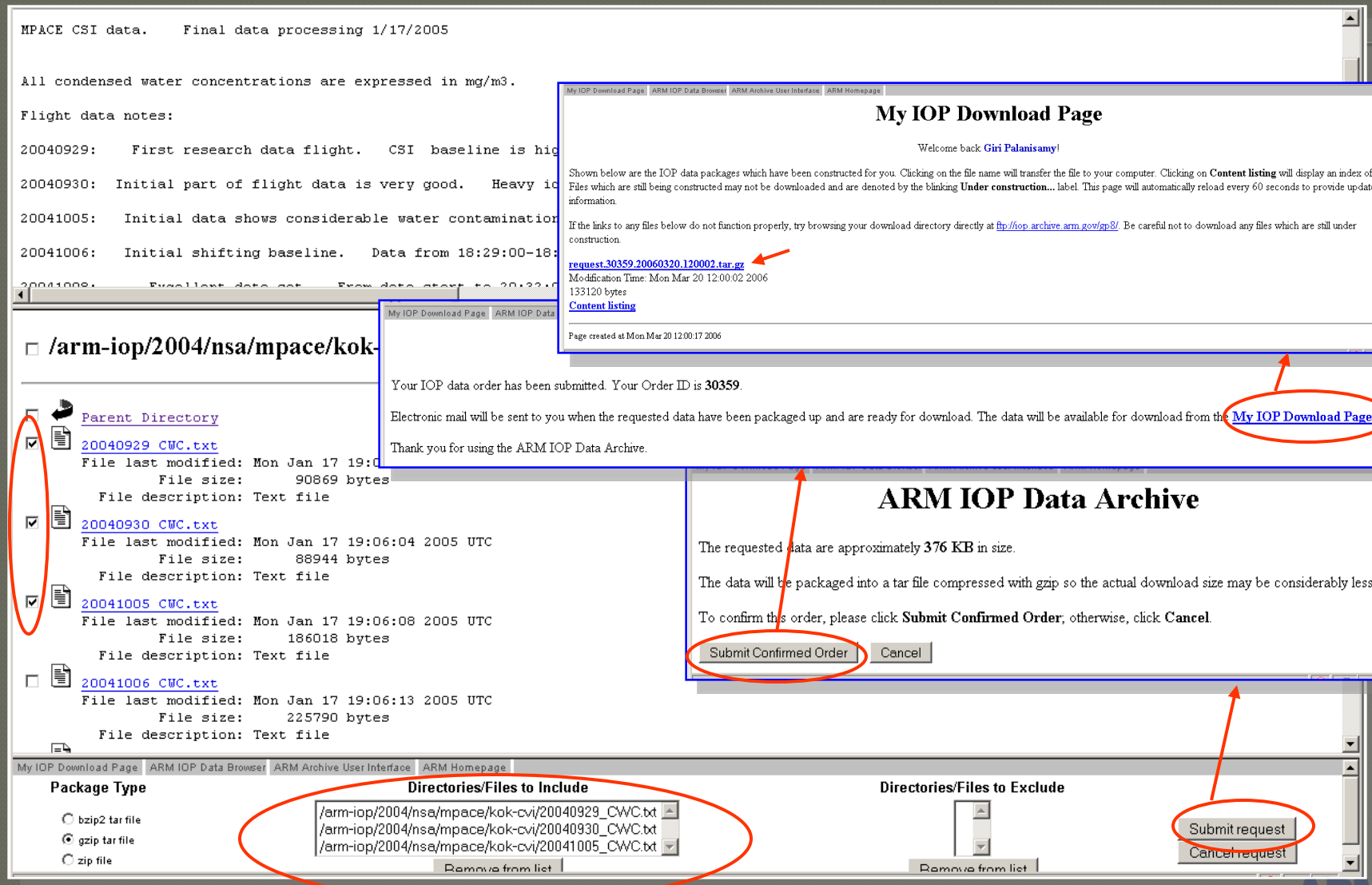
Directory Navigation

Download Management

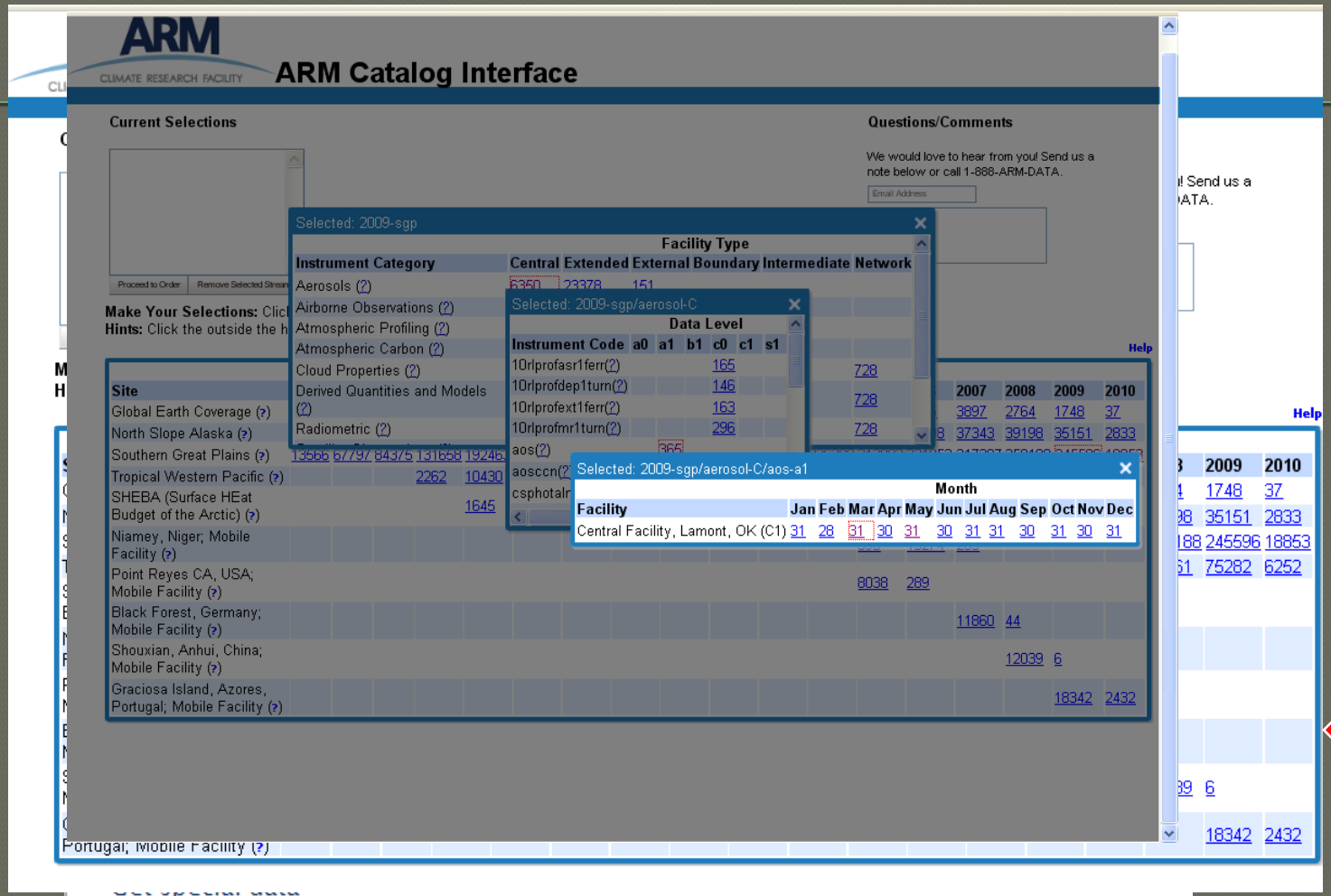
IOP Data Browser – IOP View



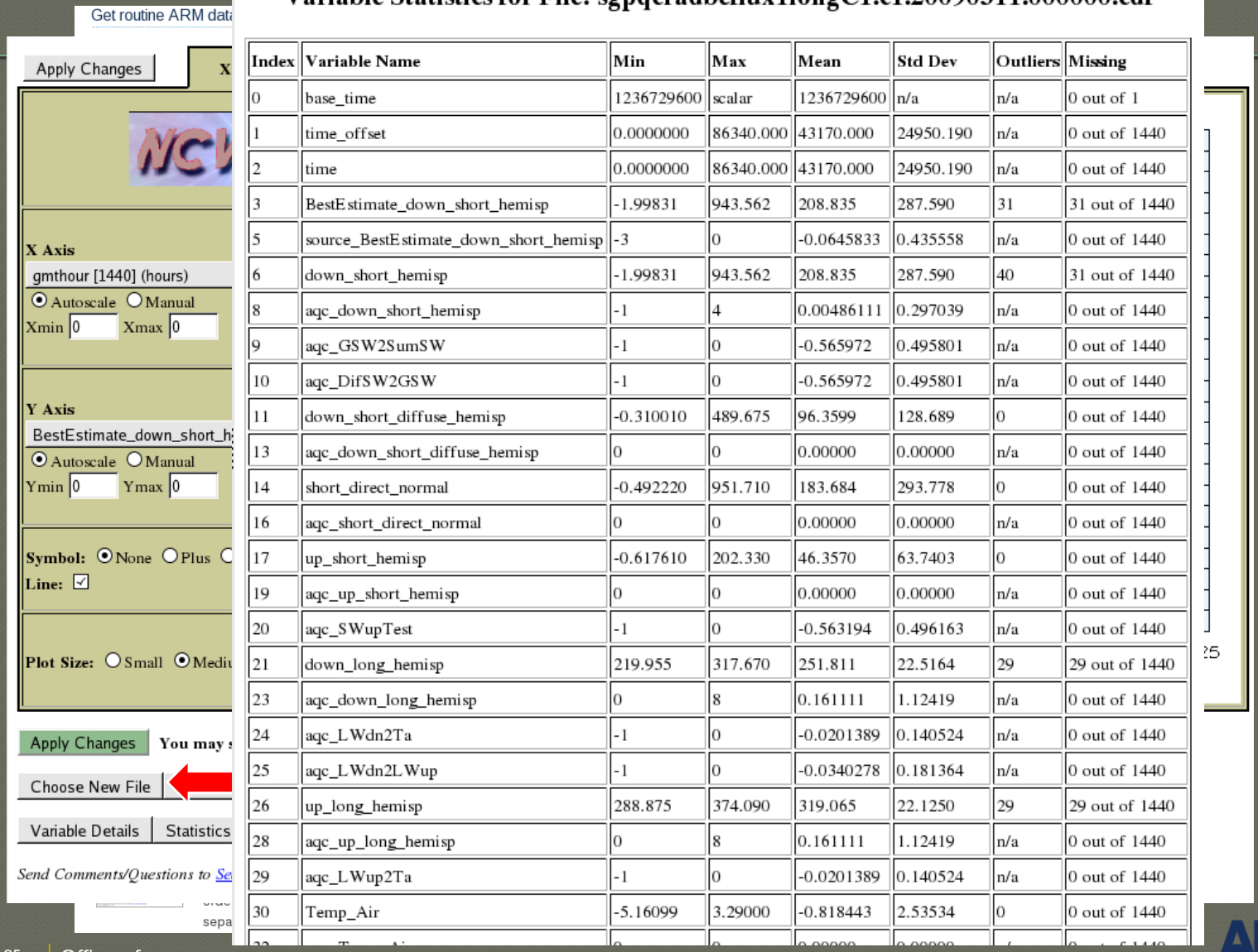
IOP Data Browser – Download Bulk Data



ARM Archive – Catalog Browser Example (new version)



ARM Archive – NCVWeb Example



Variable Statistics for File: sgpqcradbeflux1longC1.c1.20090311.000000.cdf

Index	Variable Name	Min	Max	Mean	Std Dev	Outliers	Missing
0	base_time	1236729600	scalar	1236729600	n/a	n/a	0 out of 1
1	time_offset	0.0000000	86340.000	43170.000	24950.190	n/a	0 out of 1440
2	time	0.0000000	86340.000	43170.000	24950.190	n/a	0 out of 1440
3	BestEstimate_down_short_hemisp	-1.99831	943.562	208.835	287.590	31	31 out of 1440
5	source_BestEstimate_down_short_hemisp	-3	0	-0.0645833	0.435558	n/a	0 out of 1440
6	down_short_hemisp	-1.99831	943.562	208.835	287.590	40	31 out of 1440
8	aqc_down_short_hemisp	-1	4	0.00486111	0.297039	n/a	0 out of 1440
9	aqc_GSW2SumSW	-1	0	-0.565972	0.495801	n/a	0 out of 1440
10	aqc_DifSW2GSW	-1	0	-0.565972	0.495801	n/a	0 out of 1440
11	down_short_diffuse_hemisp	-0.310010	489.675	96.3599	128.689	0	0 out of 1440
13	aqc_down_short_diffuse_hemisp	0	0	0.00000	0.00000	n/a	0 out of 1440
14	short_direct_normal	-0.492220	951.710	183.684	293.778	0	0 out of 1440
16	aqc_short_direct_normal	0	0	0.00000	0.00000	n/a	0 out of 1440
17	up_short_hemisp	-0.617610	202.330	46.3570	63.7403	0	0 out of 1440
19	aqc_up_short_hemisp	0	0	0.00000	0.00000	n/a	0 out of 1440
20	aqc_SWupTest	-1	0	-0.563194	0.496163	n/a	0 out of 1440
21	down_long_hemisp	219.955	317.670	251.811	22.5164	29	29 out of 1440
23	aqc_down_long_hemisp	0	8	0.161111	1.12419	n/a	0 out of 1440
24	aqc_LWdn2Ta	-1	0	-0.0201389	0.140524	n/a	0 out of 1440
25	aqc_LWdn2LWup	-1	0	-0.0340278	0.181364	n/a	0 out of 1440
26	up_long_hemisp	288.875	374.090	319.065	22.1250	29	29 out of 1440
28	aqc_up_long_hemisp	0	8	0.161111	1.12419	n/a	0 out of 1440
29	aqc_LWup2Ta	-1	0	-0.0201389	0.140524	n/a	0 out of 1440
30	Temp_Air	-5.16099	3.29000	-0.818443	2.53534	0	0 out of 1440

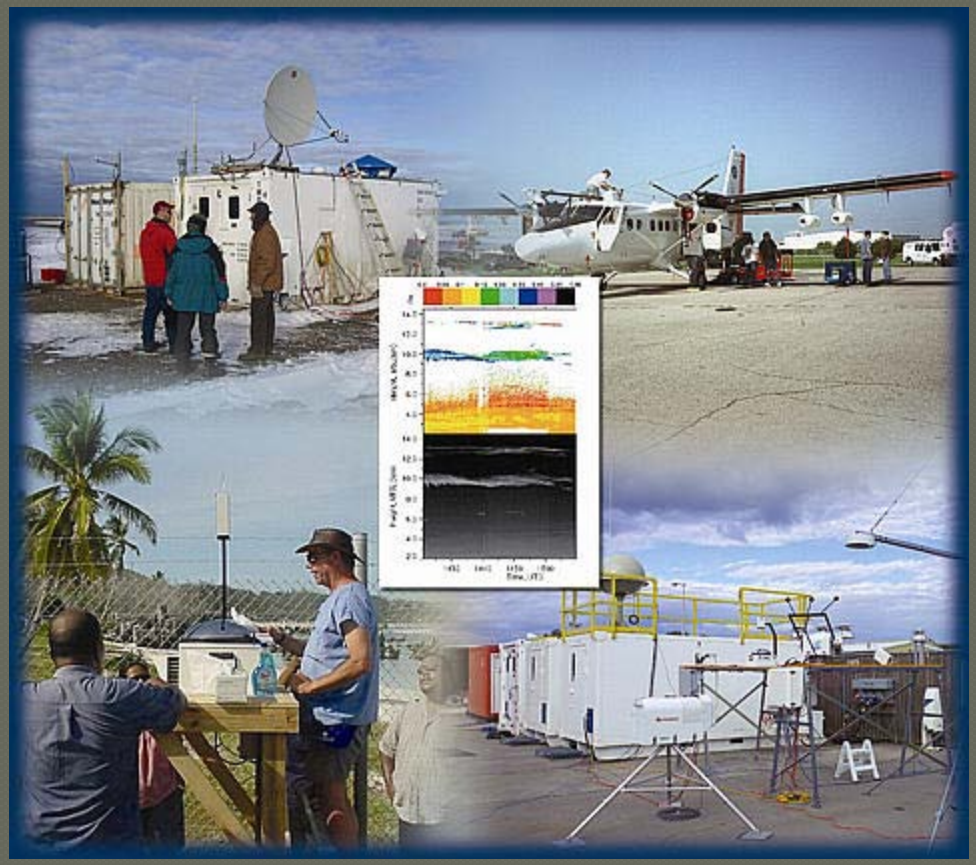
Contact Us at . . .

- ARM Information
  - http://www.arm.gov
  - info@arm.gov
  - 1-888-ARM-DATA (1-888-276-3282)
- Archive Assistance
  - armarchive@ornl.gov
  - Phone: 1-888-ARM-DATA
  - Fax : 1-865-574-4665

Backup Materials

- IOP Data Browser
- Other information
  - Other Data types
  - Sources
  - Filename syntax
- Quality information structure
- Standing Orders
- Archive details
- Web diagram
- Interface details
  - Catalog Browser
  - Thumbnail Browser

ARM IOP* Data Browser



*IOP == Field Campaign

Field Campaigns: Options for data sharing at the Archive

- Optional and available upon request by Campaign PI
- FTP-only accounts create for each Co-PI
- PI and coworkers are responsible for data transfers
- Storage and access valid for only 6-12 months after the end of Campaign
- ARM Archive / External Data Center staff will transfer data from preliminary area to fully open access to final data
  - Campaign data in preliminary area can be finalized there and do not needed to be resubmitted to IOP upload system

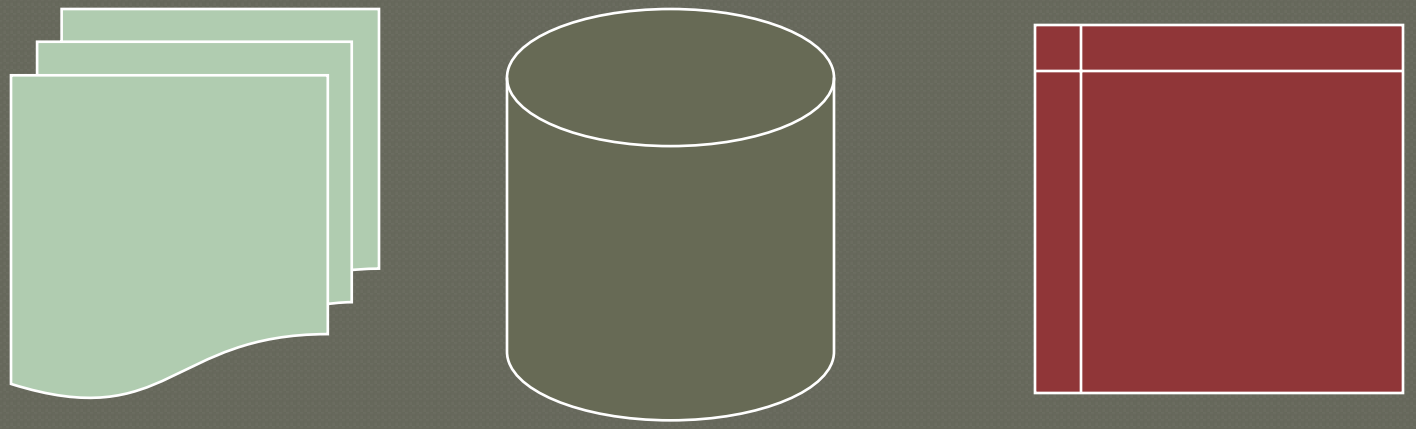
- OR

- Procedure for upload of final data (http://www.xdc.arm.gov/docs/iopsteps.html)
- Contact External Data Center Staff to coordinate data transfer
  - Currently: Alice Cialella (cialella@bnl.gov, 1-631-344-3286

Other ARM Archive Documentation Topics

- Comparison of data types
  - *(routine data, IOP data, etc.)*
- Type of available quality information
  - *(flags, Data Quality Reports)*
- Data access beyond the user interfaces
  - *("contact us", Standing Orders)*
- Archive details
  - *Performance statistics*
  - *Logical configuration*

Data Types



ARM Data Types - overview

- Continuous data (stored offline, accessible by requests from user interface)
  - ARM collected data
  - Value added products
  - External data
- Special data (stored online, accessible from web interface)
  - Field Campaign (IOP) data
  - Beta data
  - PI generated data products

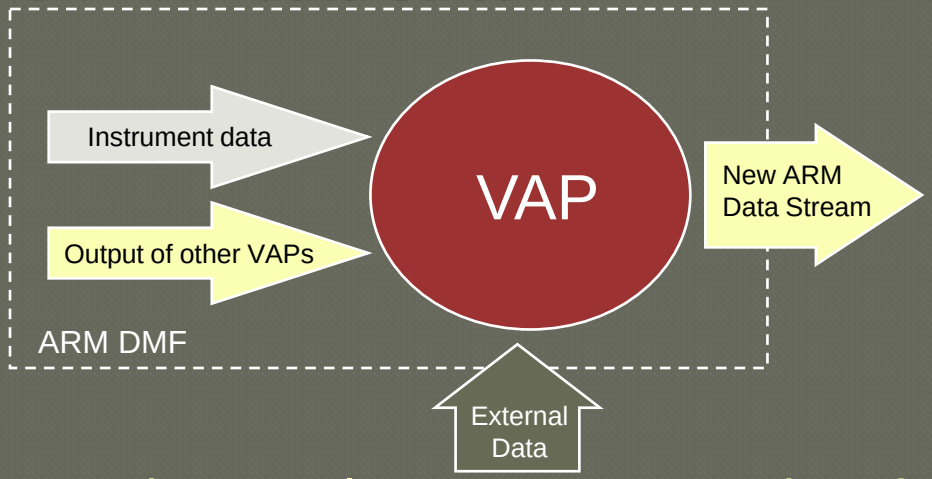
ARM Data Types – more detail

armarchive@ornl.gov
1-888-ARM-DATA

- ARM collected data
  - RAW data files
    - Available upon request, but not accessible from User Interface
    - Minimal documentation; user beware
    - Wide variety of formats; many are binary
  - Processed data files
    - Accessible from user interfaces
    - Common formats include NetCDF and HDF
- Value added products (VAPs)
  - Include one or more of the following
    - Advanced algorithms
    - Multiple data inputs
    - Input from long-time periods
    - Best estimate products
  - ARM produces some VAPs to improve the quality of existing measurements. In addition, when more than one measurement is available, ARM also produces "best estimate" VAPs.

More on VAPS…

- VAPs are products from automated analytical procedures (models, retrievals, etc.) that are run in the ARM data system
- Inputs come from instruments, other VAPs, and/or external data
- Output is a new ARM data stream

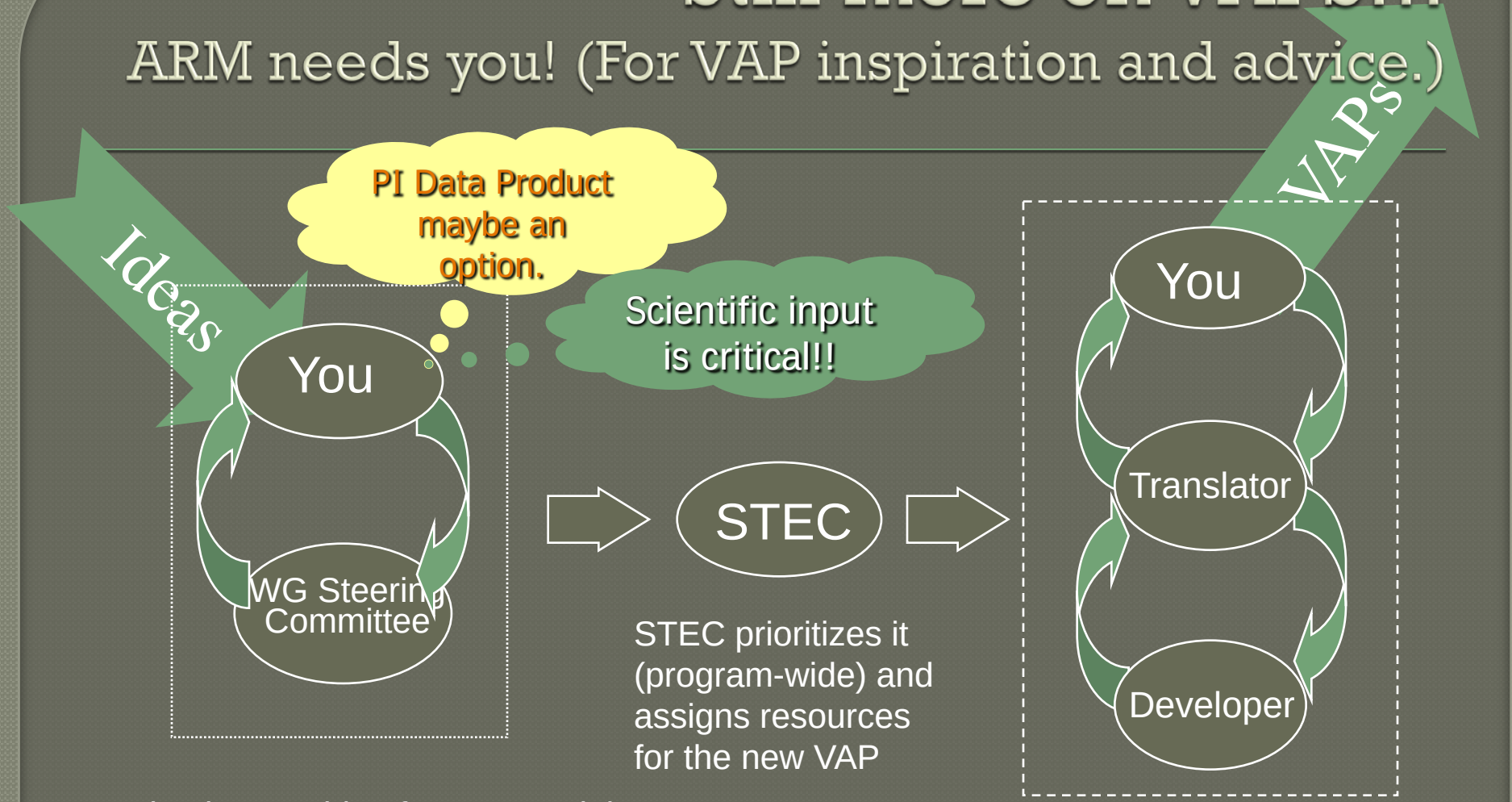


- ARM wants your input. Please note *"Procedure for Submitting Science and Research Products to the Data Archive"* at:

http://www.arm.gov/data/pi_procedure.stm

Still more on VAPS…

ARM needs you! (For VAP inspiration and advice.)



Scientist gets idea for new model or algorithm and presents it to the WG. The WGSC prioritizes the idea and contacts the STEC

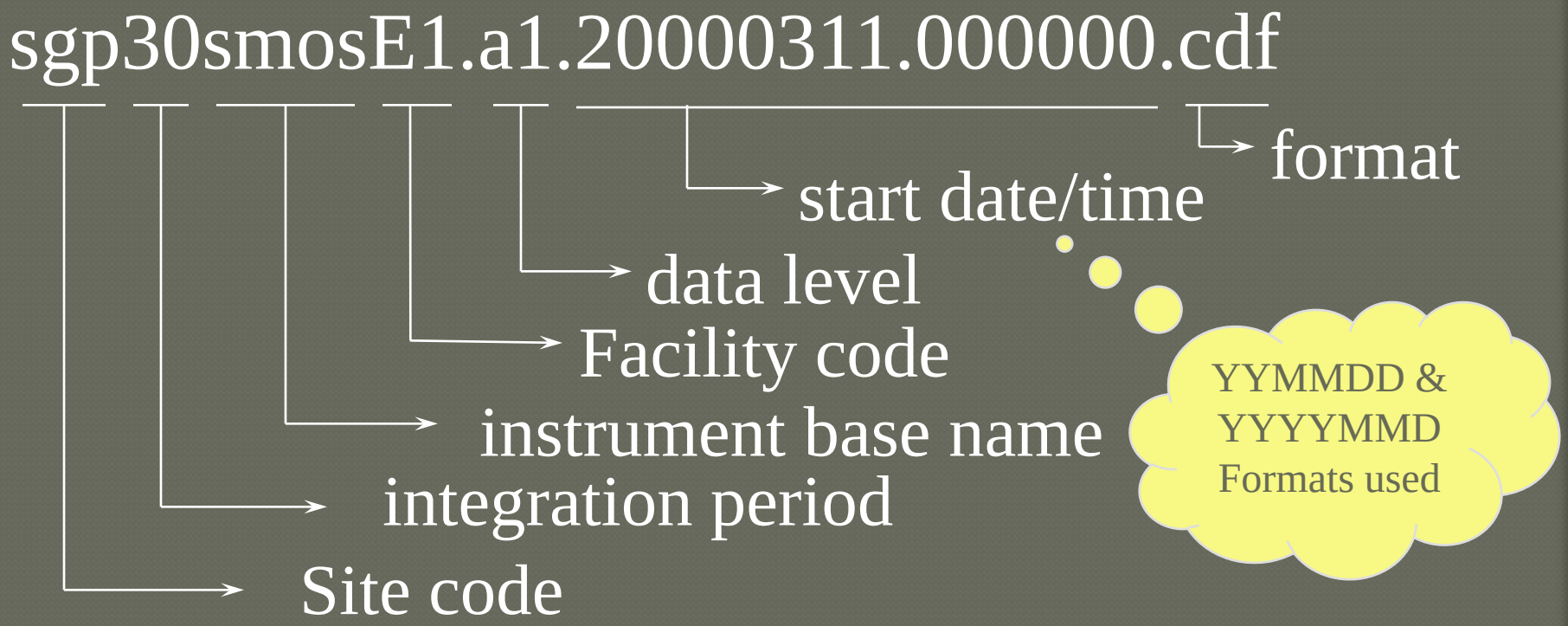
STEC prioritizes it (program-wide) and assigns resources for the new VAP

Translator works with the Scientist to further define the algorithm, and then interacts with the Developer to implement the VAP. Translator and the Scientist then evaluate and document.

ARM Data Types – still more detail

- External data
  - Generated by other programs (e.g., NOAA weather models, NASA satellites, etc.)
  - Many formatted into NetCDF consistent with ARM style
  - Specialized subsets specific to ARM sites
    - Geographic clips of global data
- Field Campaign Data
  - Special experiments (e.g., M-PACE, 2003 Aerosols, etc.)
  - Stored online in separate data structure
- PI generated data products
  - Considered useful to ARM users
  - Provided "at will" by a researcher
  - Supported by the researcher
- Showcase data sets
  - Condensed and integrated subsets of selected ARM datastreams
  - Targeted for a particular research community and contain only a few measurements
  - Usually "best estimates" derived from instruments and/or VAPS

sgp30@%$#&!!! (or, making sense of the ARM file naming convention)



Quality Information

Types of Quality Information

- Automated products
  - QC flags
    - inserted in data files during processing
  - Summaries of flags (data color)
- Manual products
  - Data Quality Reports (DQRs)
    - web accessible reports; delivered as html files after data requests (more later); event driven and problem-based
  - Instrument Mentor Monthly Summary Reports
    - web accessible; linked to instrument web pages.
  - Data Quality Assessment Reports

Beyond User Interfaces

"Standing Orders": Data Distribution Upon Arrival

- A "Standing Order" is an open request for access to a copy of all new files arriving at the Archive
  - Matching a user-specified set of data streams
  - Arriving during a delivery period
- Designed for users wanting to:
  - Access data shortly after Archival
  - Build their own complete collection of selected data streams

Standing Order Processing

Standing Order Information

- Online documentation
  - http://www.archive.arm.gov/docs/standing-orders.html
- Send request:
  - What data streams?
  - What delivery frequency?
  - To: armarchive@ornl.gov
- More details in reference section of handout

Archive Details

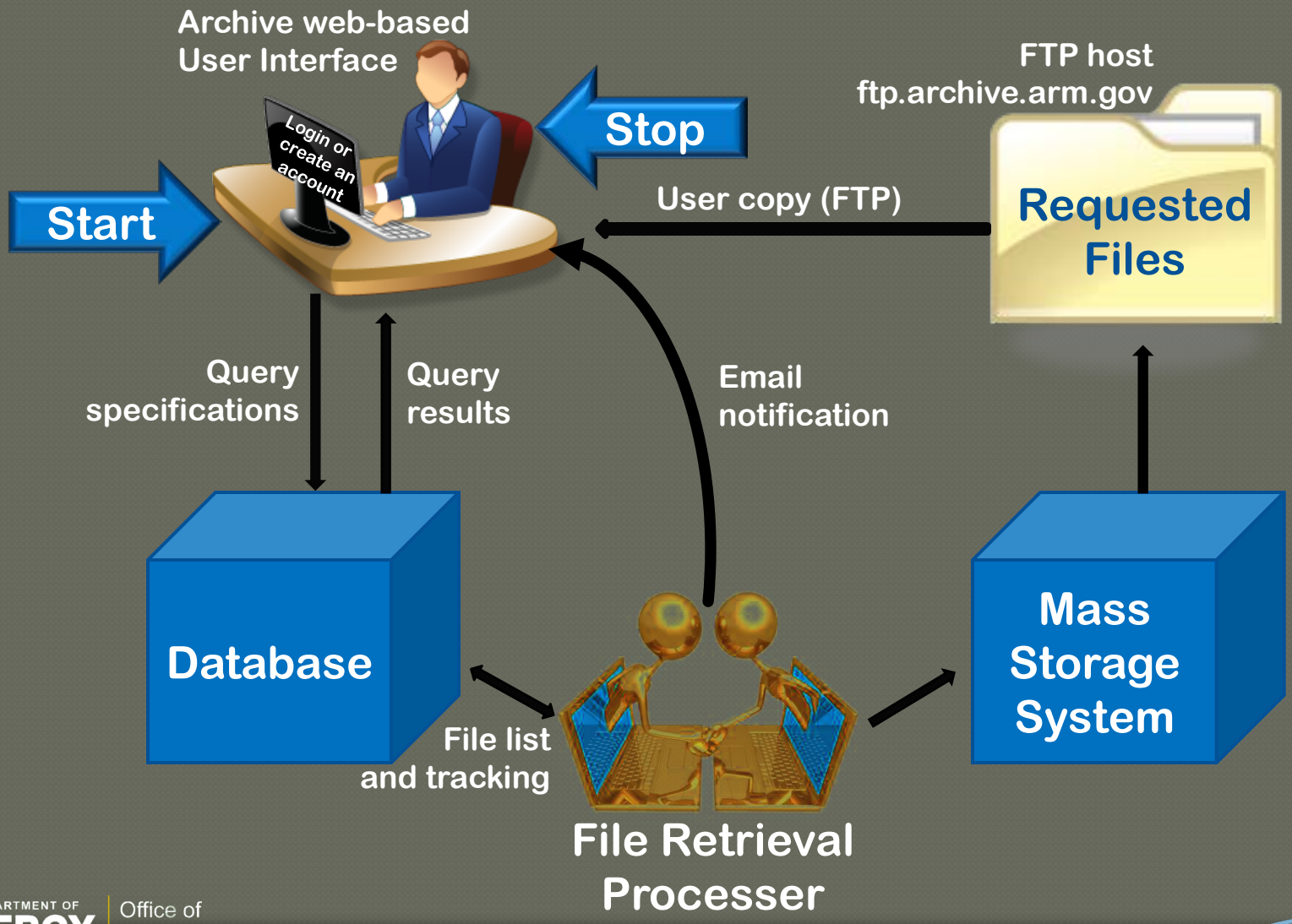
More than you want to know...?

You are NOT alone…

- 3 sites
- 10's facilities
- 100's data sources
- 100's data users
- 1000's measurement types
- 1,000,000's data files
- 1,000,000,000's measurements
- 10,000,000,000,000's bytes

- Storage
  - New data: 50-100,000 files, 2-3 TB per month
  - Total storage: 9.9 million files, 182 TB of data
- Usage
  - Yearly requests – 2 million files, 20 TB of data
  - 800-1000 different users active each year.

You and the Archive 'Guts' (Sanitized)

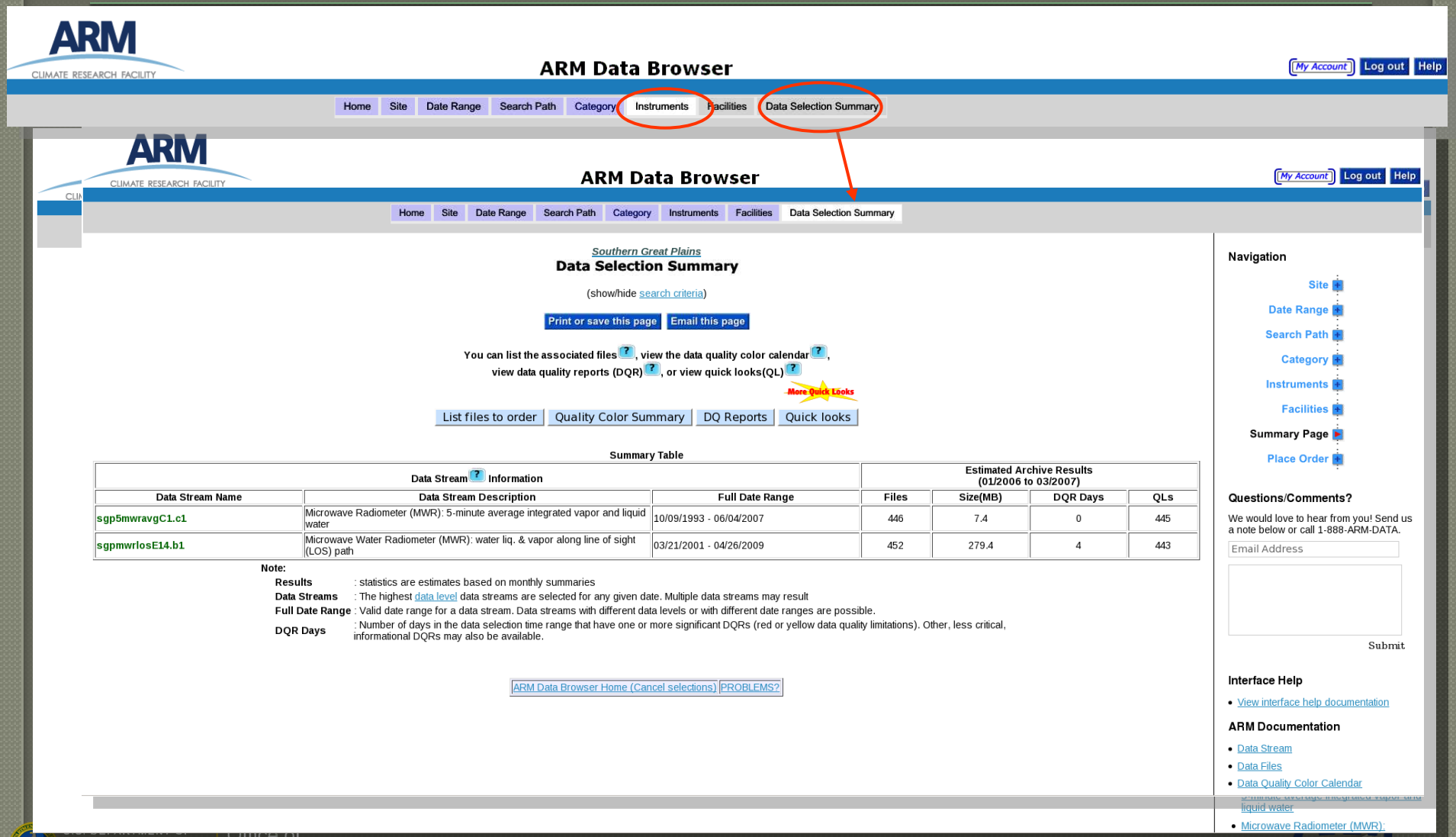


Accessing Data from the Archive

- *Contact Us…..*
  - *1-888-ARM-DATA, armarchive@ornl.gov*
- Continuous data distribution
  - "Standing Orders"

Remaining slides are backups; some taken/moved out of STM2008 talk

ARM Data Browser



ARM Catalog Browser

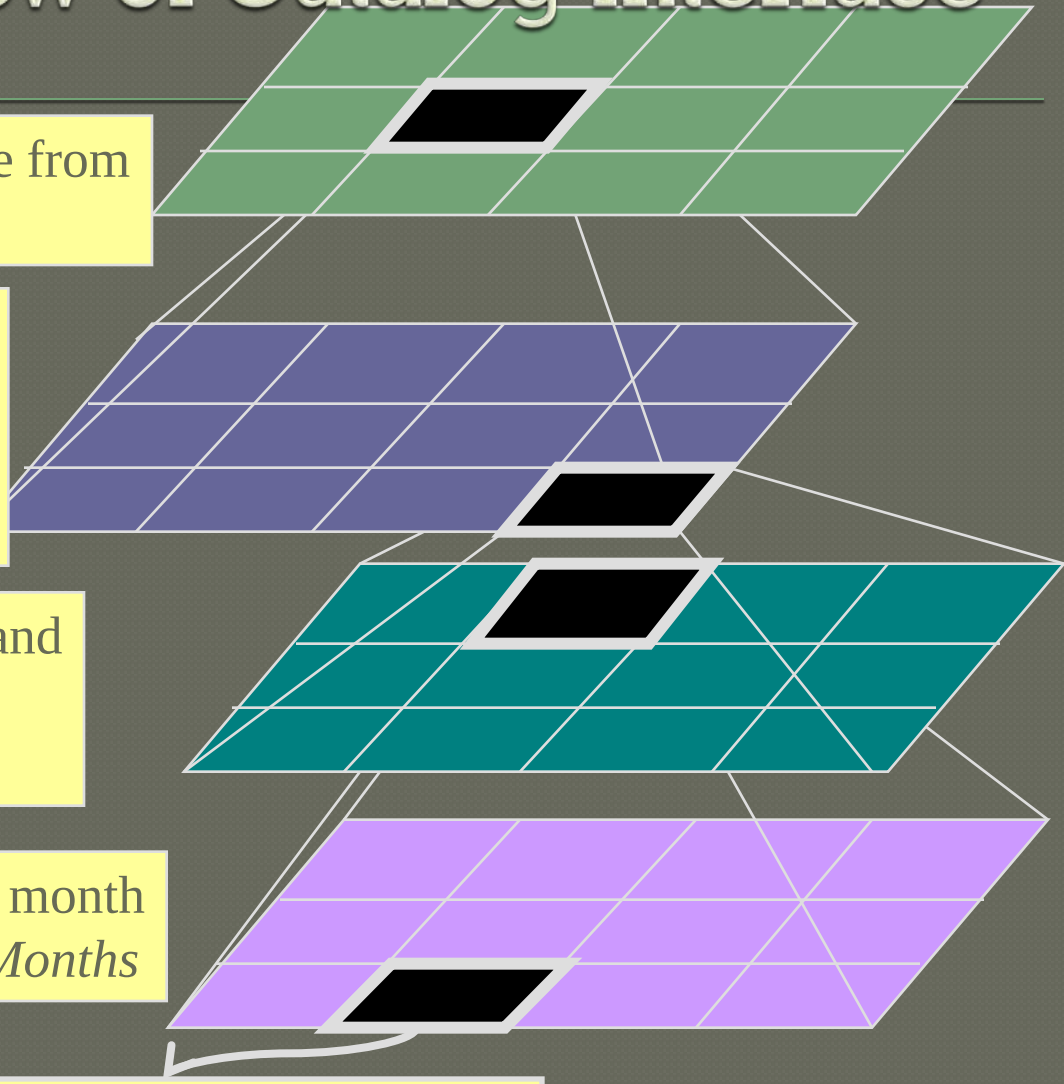
Logical Flow of Catalog Interface

Step1: select a year and a site from a table of *Years x Sites*

Step 2: select a facility type and instrument category from a table of *Instrument category x Facility type*

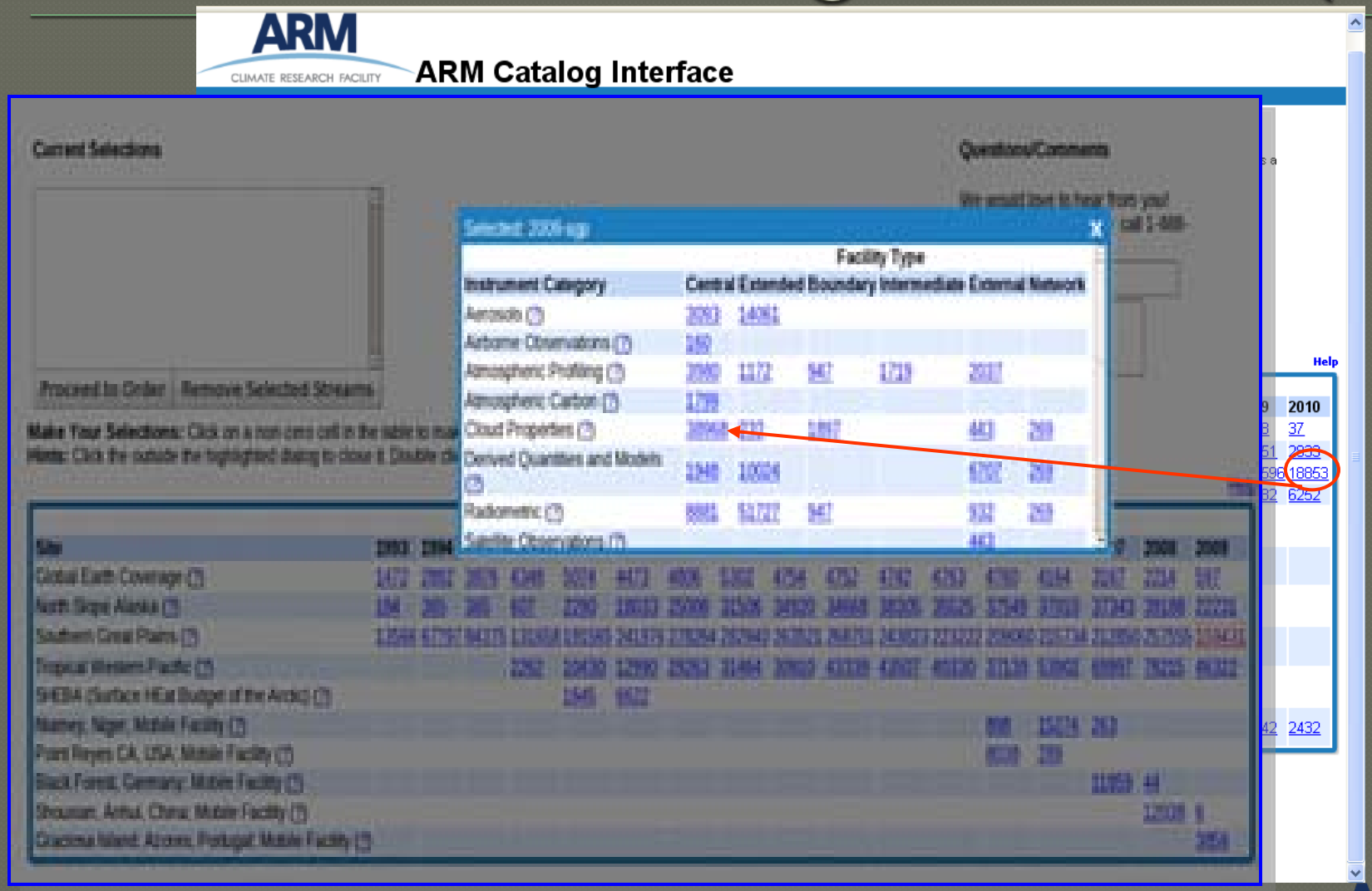
Step 3: select an instrument and a data level from a table of *Instruments x Data levels*

Step 4: select a facility and a month from a table of *Facilities x Months*



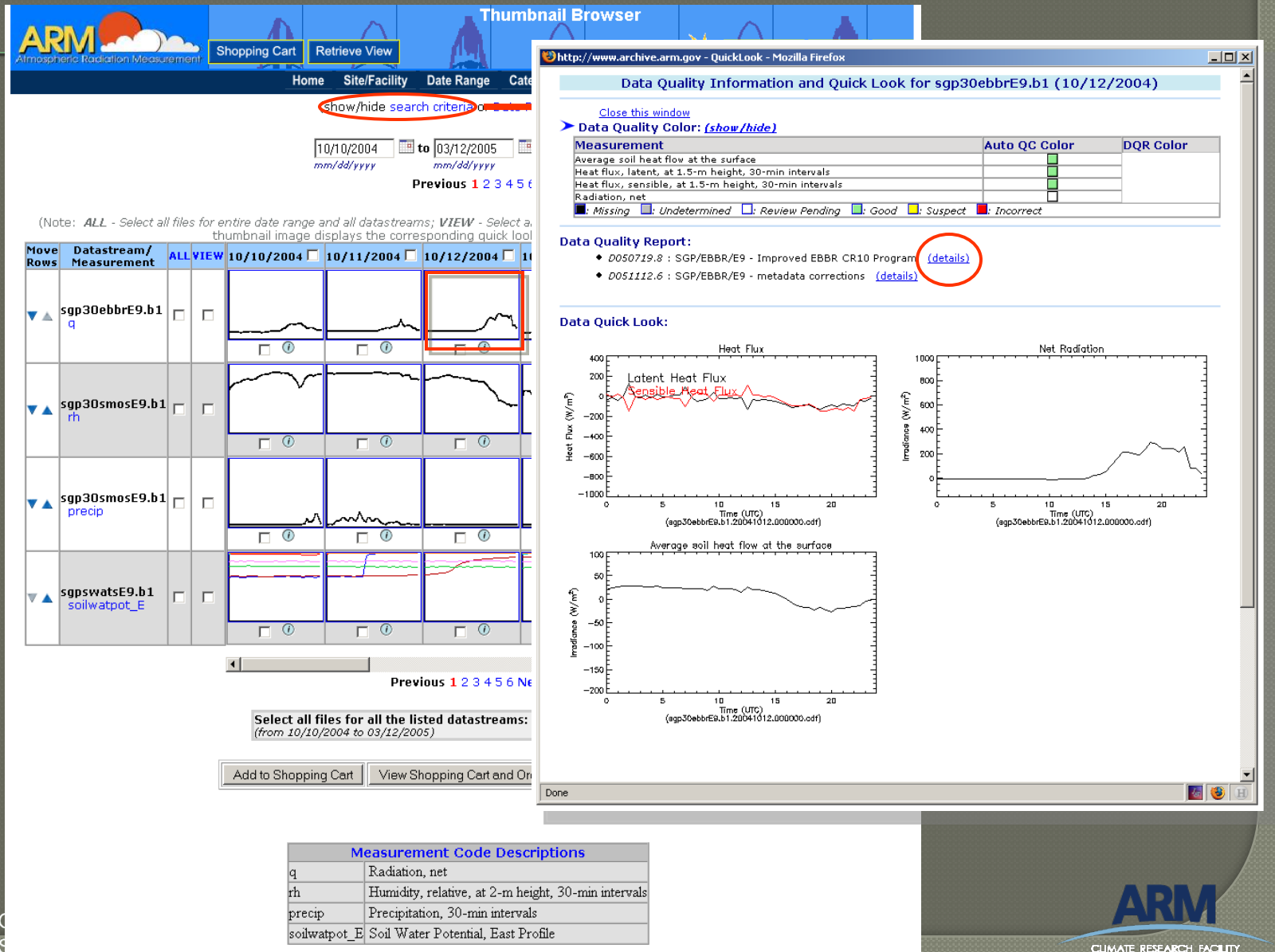
Data added to 'Add to cart'

Catalog Browser (1)



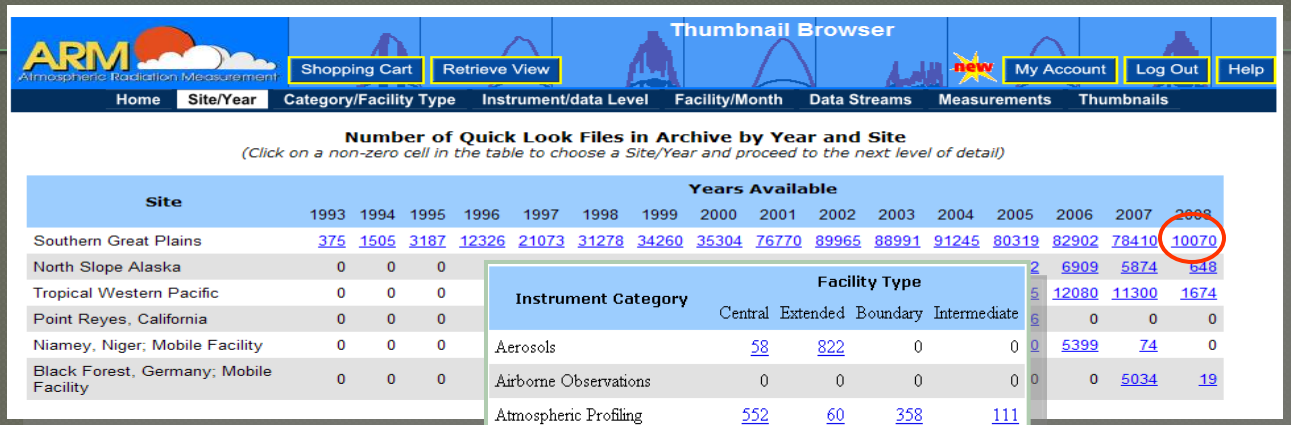
ARM Thumbnail Browser

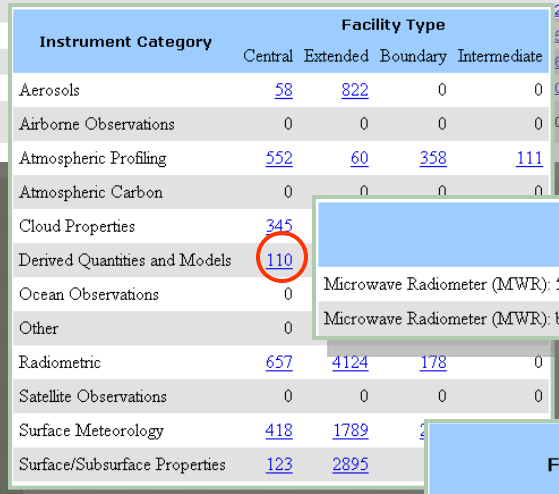
Display Thumbnails

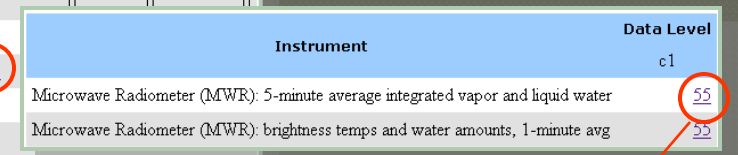


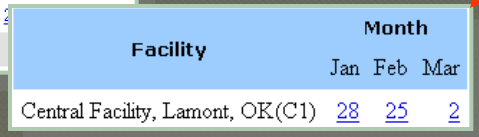
Measurement Code Descriptions	
q	Radiation, net
rh	Humidity, relative, at 2-m height, 30-min intervals
precip	Precipitation, 30-min intervals
soilwatpot_E	Soil Water Potential, East Profile

Thumbnail Browser – Catalog Interface

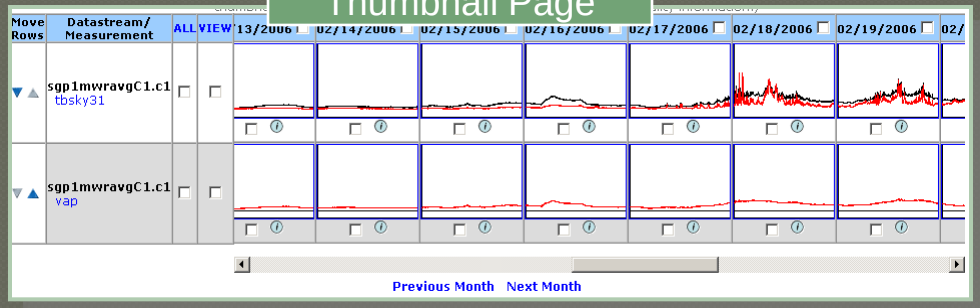








Thumbnail Page



ARM Data Cart Tutorial

- NEEDS NEW LINK!! Available at: http://www.arm.gov/data/orderData_tutorial.stm