Obtaining Data from the ACRF Archive

Orientation for new Science Team Members

Office of Science
U.S. Department of Energy

One click to the ACRF Data Archive

Office of Science
U.S. Department of Energy

ACRF Data Archive – Newly Designed Home Page

Data Access Tools

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. Help

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. Help

Catalog Browser — The catalog based user interface presents, in an interactive sequence of tables, a hierarchical summary of available data files organized in a way that will be useful to the inexperienced, as well as the expert Archive user. Help

Thumbnail Browser — View prepared plots of data to quickly find data of interest to you. The thumbnail browser uses location, measurement type and date range selections to retrieve data plot thumbnails that the user can browse. You can also download high-resolution images of the data plots, or download the data files. Help

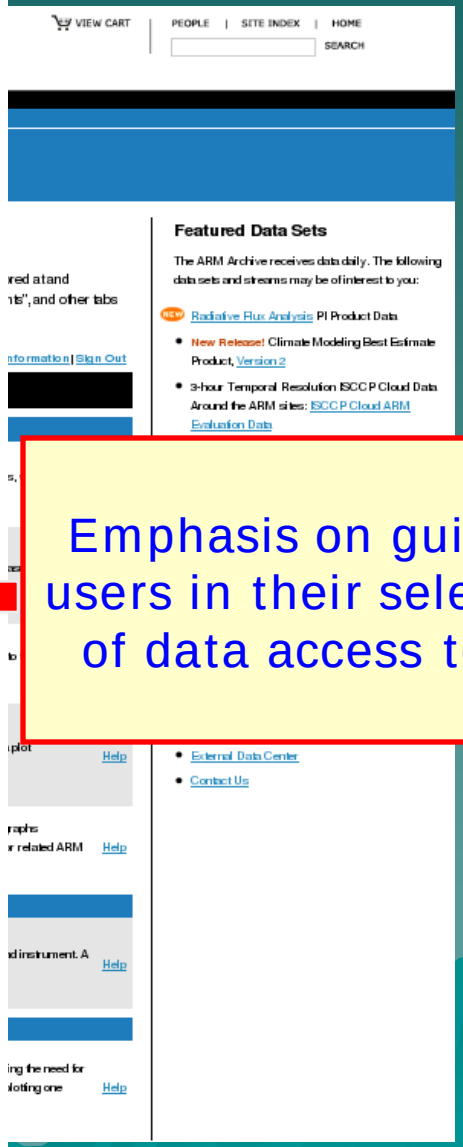
Statistical Browser — Users select a location and measurement and then drill down through time scales ranging from the full period of record to individual months. In addition to viewing graphs displayed by this interface, access to extractions of data behind the statistical graphs, obtain the measurements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

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. Help

Plot previously ordered data

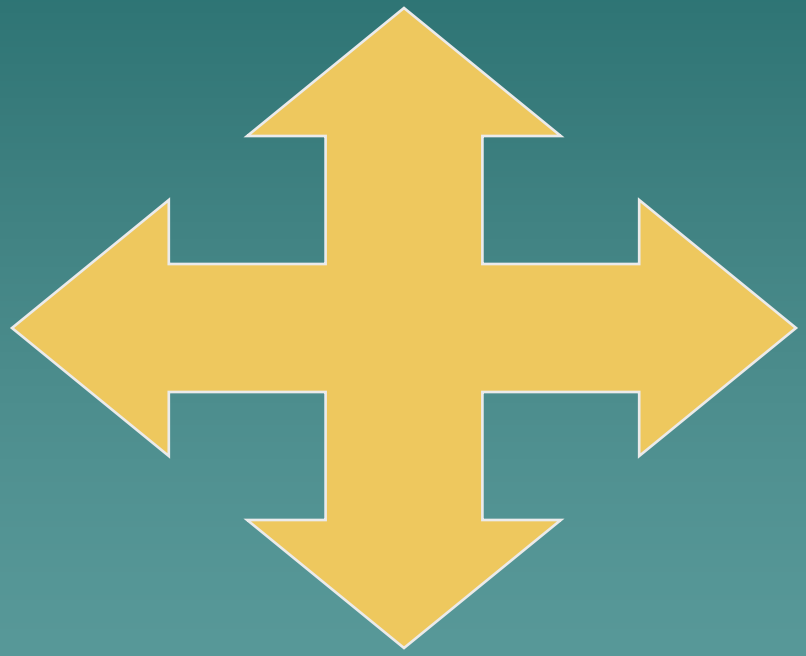
NCV Web — NCVWeb is an interactive NetCDF data plotting tool users can use to plot the data they have ordered from the archive, or plot regular standing data orders, eliminating the need for separate visualization software. It has many powerful features such as producing detailed tables of NetCDF file contents, data extraction, generating statistics, and plotting one variable Help



Emphasis on guiding users in their selection of data access tools

Office of Science
U.S. Department of Energy

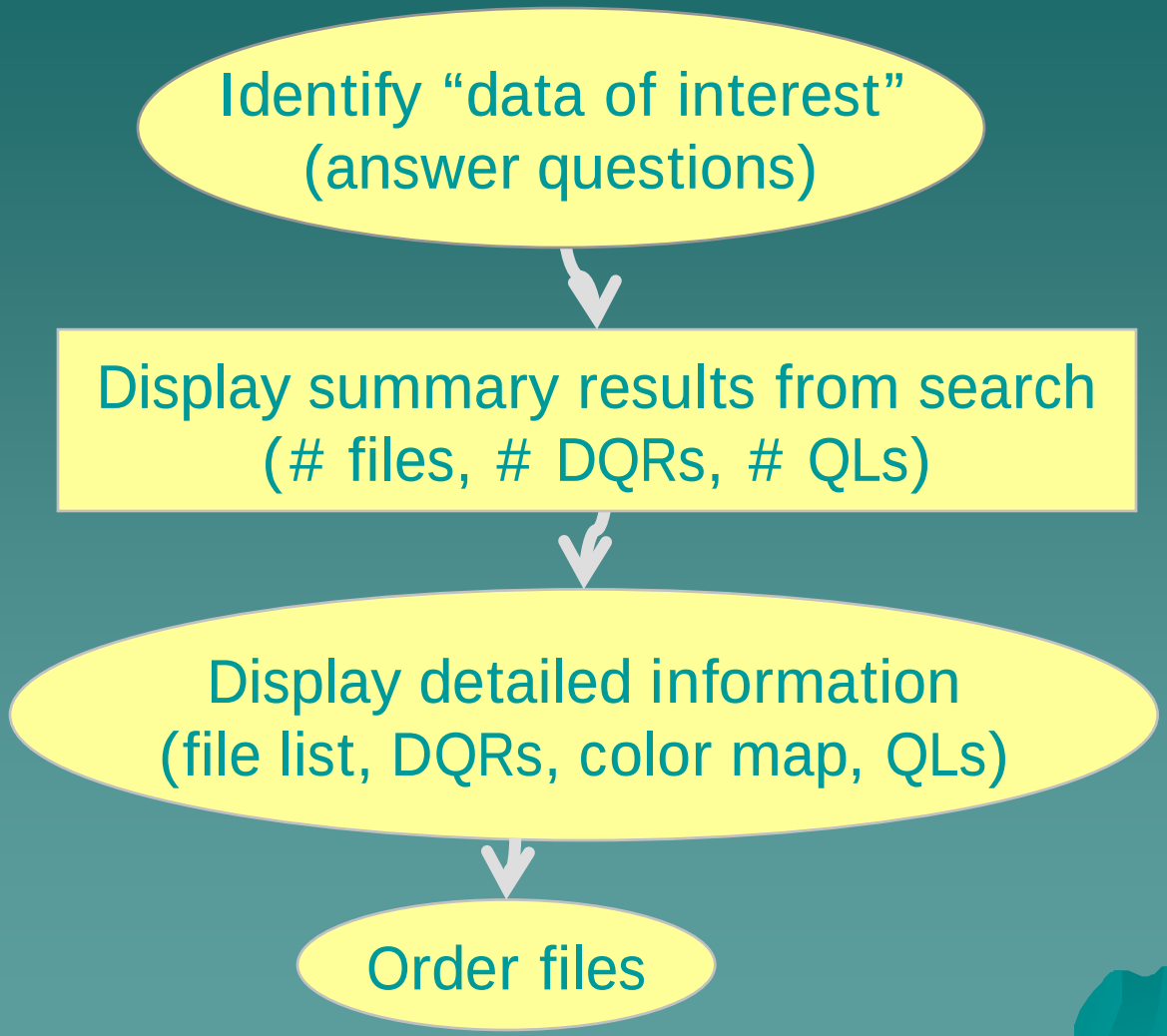
Accessing ARM Data: Options



Comparison of Browser/Interface Options

Interface name	Accessible data	"*Shopping*" approach (armarchive@ornl.gov, 1-888-ARM-DATA)
ARM 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
Web Shopping 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.
Statistical Browser	Special Data (CMBE, QCRAD, CONSTRVARANA)	*"I' need to see climatological summaries 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 and their underlying measurements and data files.
IOP Data Browser	IOP, special, PI, and 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



ACRF Archive – Data Access Examples

Data Access Tools

ARM

VIEW CART | PEOPLE | SITE INDEX | HOME

SEARCH

Cloud base height

Measurement Categories: Cloud Properties

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

Note: *Click the "▸" to expand/collapse the list.*

Instruments that produce the measurement

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
- External Instruments
- Field Campaign Instruments

Value-Added Products (VAPs) including the measurement

- (ARSCL) Active Remotely-Sensed Cloud Locations Process
- (MPL) Micropulse Lidar Process
- (MPLCBH1SCOTT) Cloud Base Height from MPL Using Scott-Spinhirne Algorithm Process
- (MPLNOR) Normalized Backscatter Profiles from the Micropulse Lidar Process
- (RLPROF) Raman LIDAR Vertical Profiles Process

Order Data

BUILD AN ORDER

Or use the datastream interface at the ARM Archive.

Questions/Comments?

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

Email Address

SUBMIT

The ARM Program gathers cloud and radiation measurements via land-based instruments, ships, and satellites.

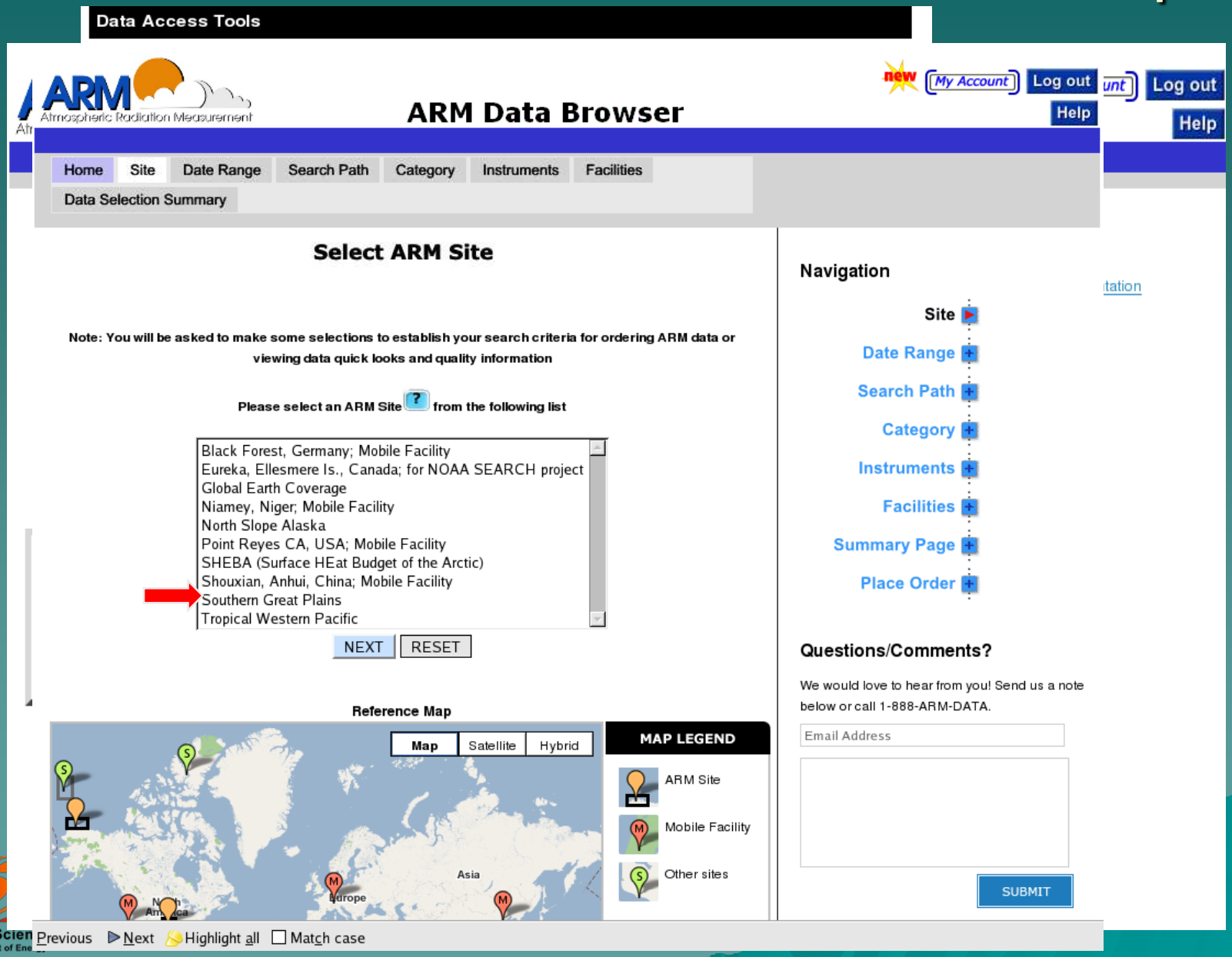
NCV Web

the data they have ordered from the archive, or plot regular standing data orders, eliminating the need for separate visualization software. It has many powerful features such as producing detailed tables of NetCDF file contents, data extraction, generating statistics, and plotting one variable

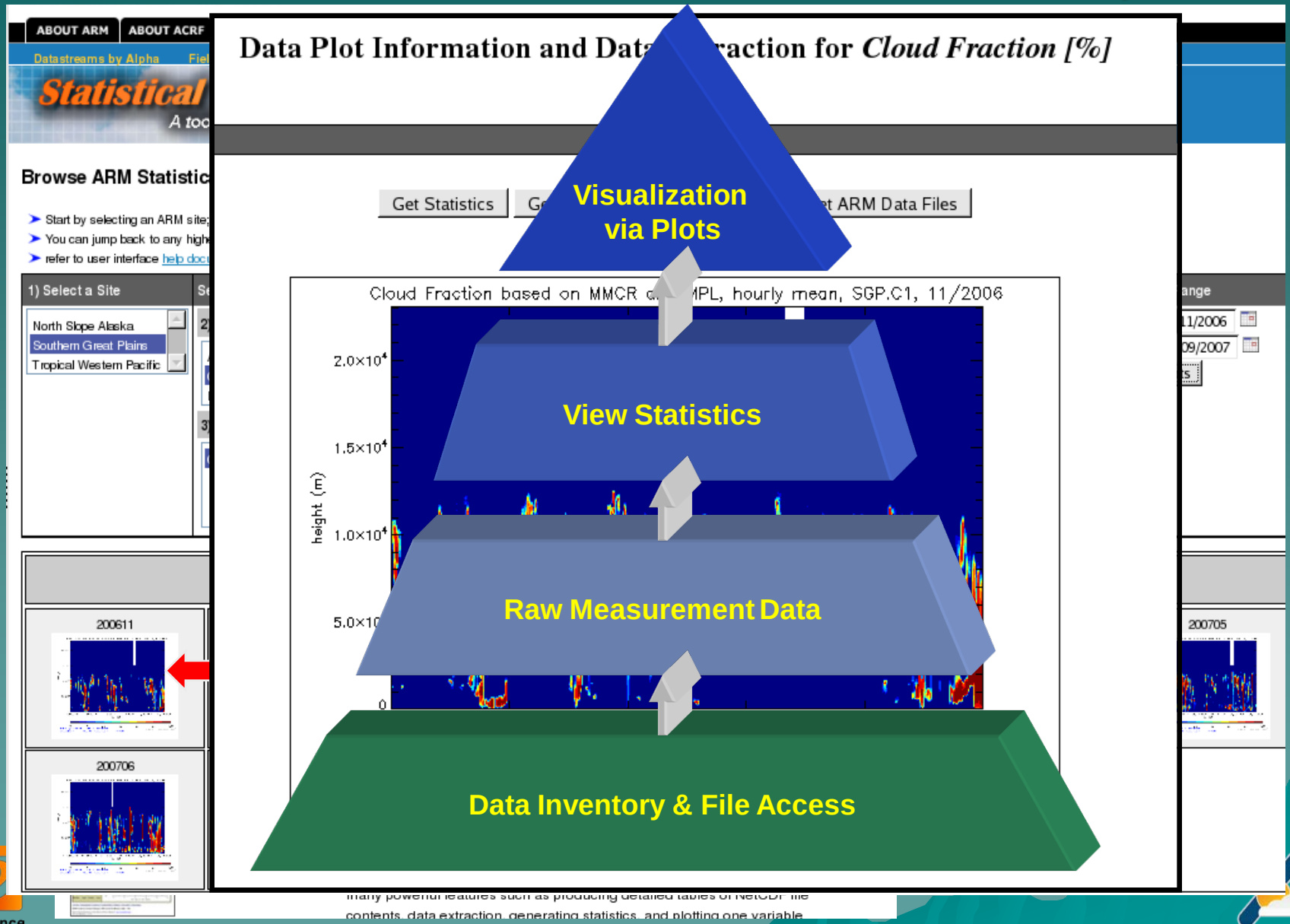
Help

Office of Science
U.S. Department of Energy

ACRF Archive – Data Browser Example

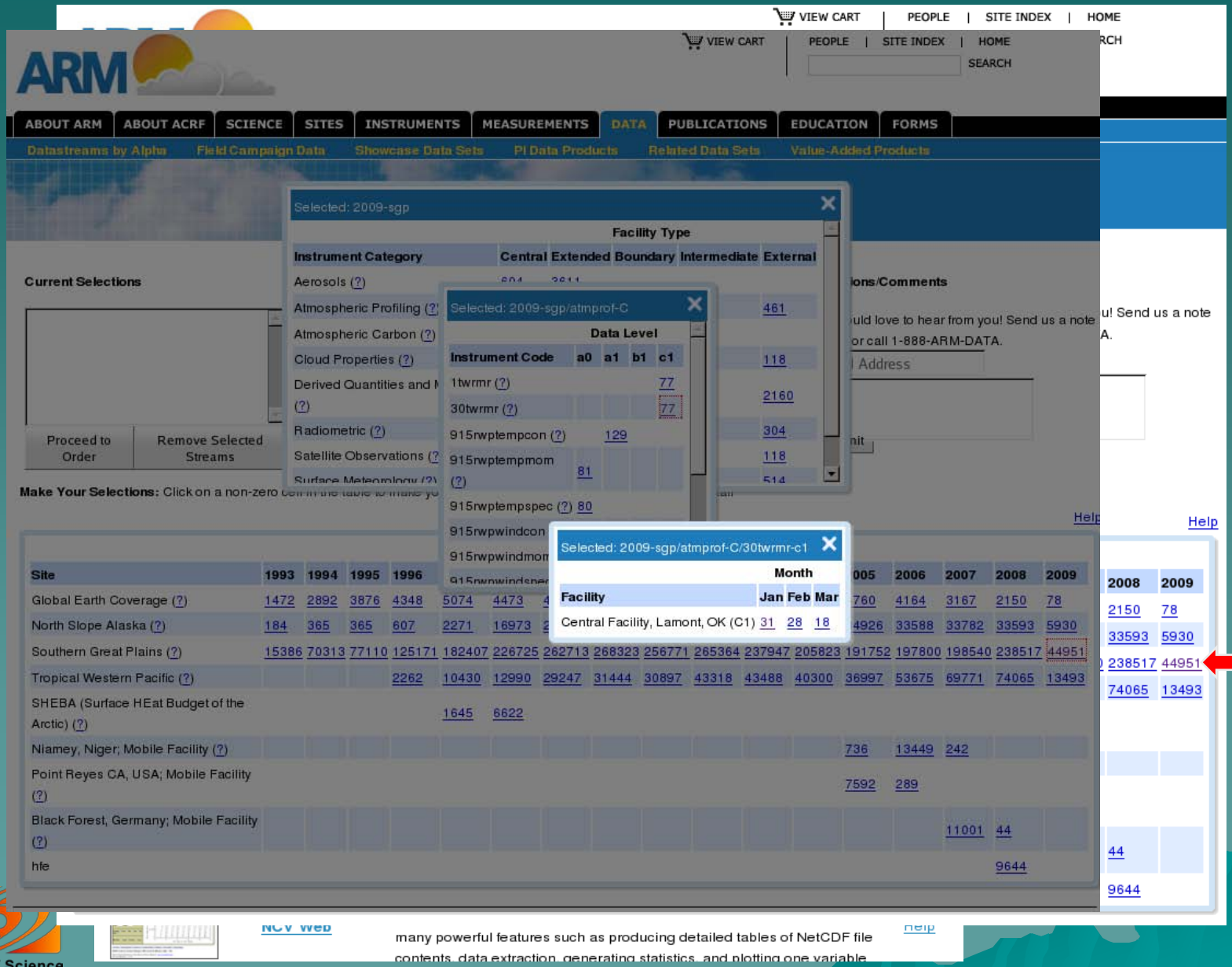


ACRF Archive – Statistical Browser Example



Office of Science
U.S. Department of Energy

ACRF Archive – Catalog Browser Example (new version)



Office of Science
U.S. Department of Energy

ACRF Archive – NCVWEB Example

Data Access Tools

Apply Changes | X Y Plot

X Axis
gmthour [1440] (hours)
Autoscale Manual
Xmin 0 Xmax 0

Y Axis
BestEstimate_down_short_hemisp [14
Autoscale Manual
Ymin 0 Ymax 0

Symbol: None Plus Star
Line:

Plot Size: Small Medium La

Apply Changes You may switch to

Choose New File | Choose New Req

Variable Details | Statistics

Send Comments/Questions to Sean Moore

NCV Web

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

contents, data extraction, generating statistics, and plotting one variable

Office of Science
U.S. Department of Energy

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
  - Call “1-888-ARM-DATA”
  - FAX 1-865-574-4665

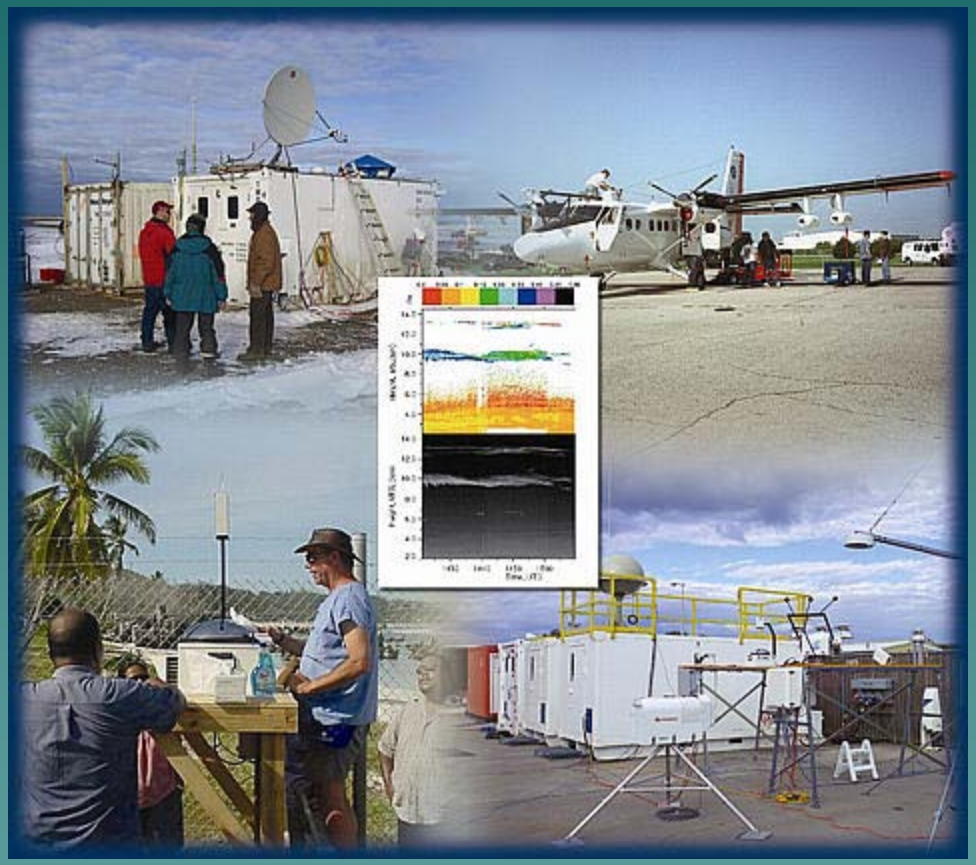
Office of Science
U.S. Department of Energy

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

Office of Science
U.S. Department of Energy

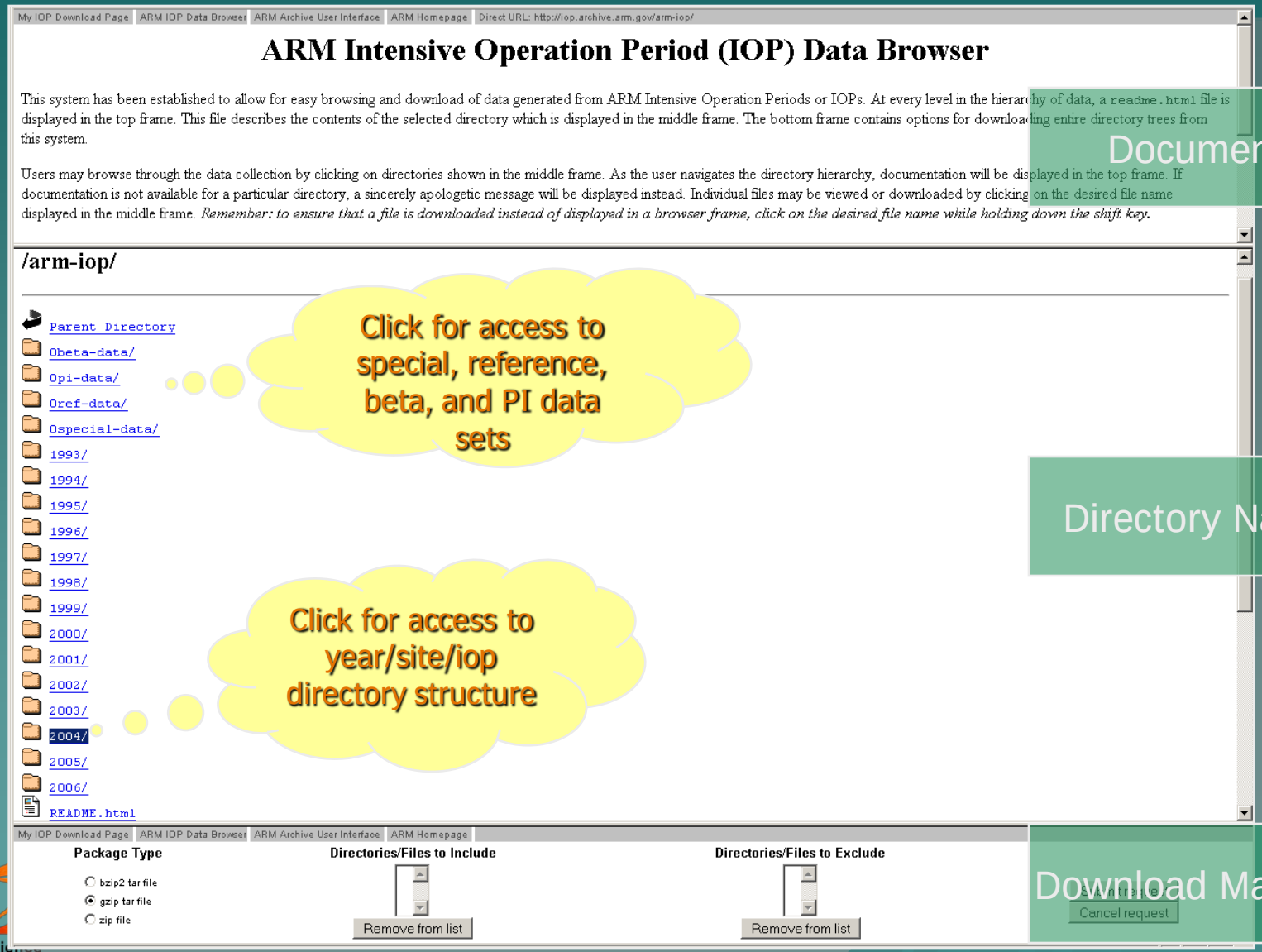
ARM IOP* Data Browser



*IOP == Field Campaign

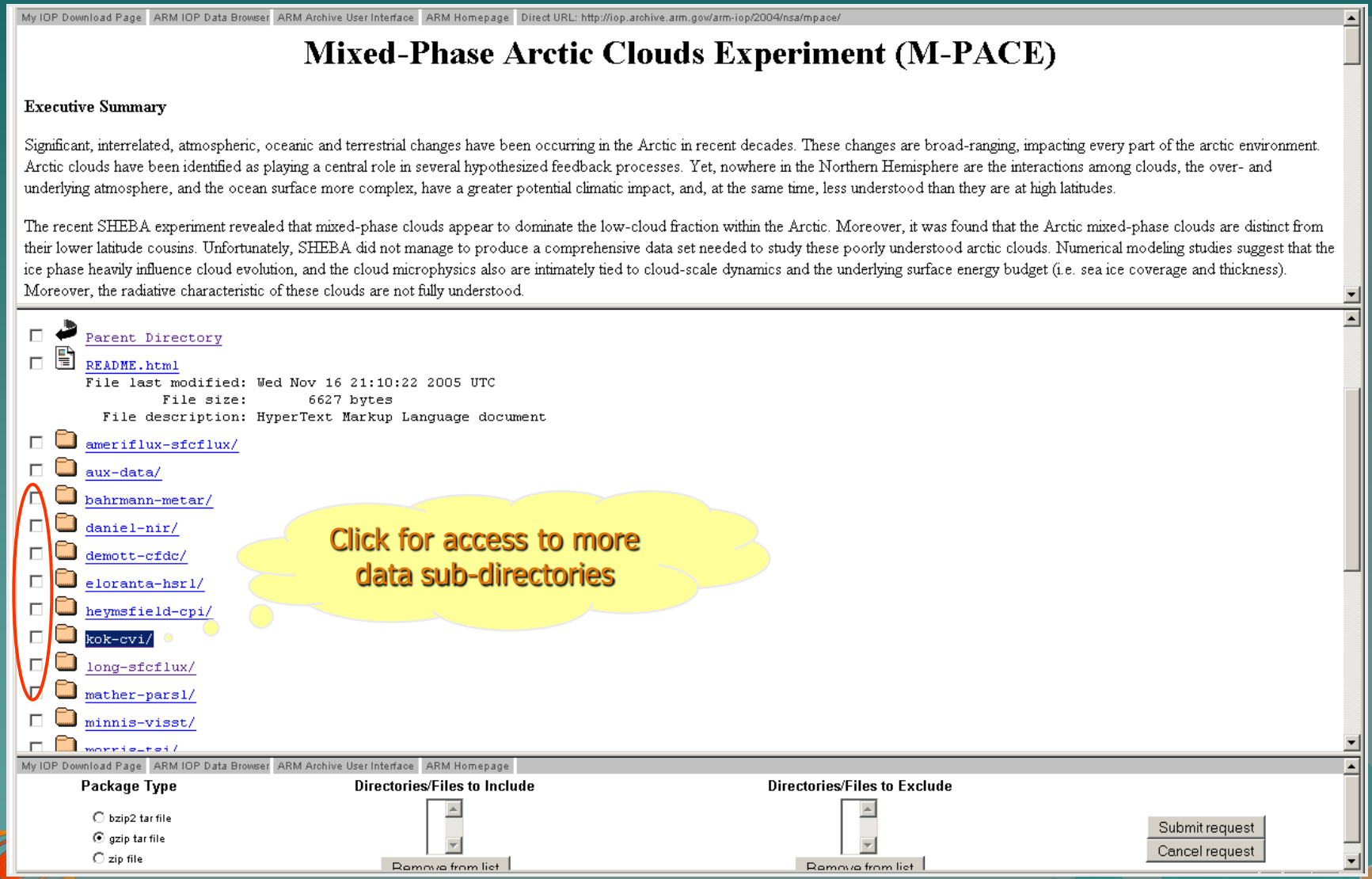
Office of Science
U.S. Department of Energy

IOP Data Browser – “home page”



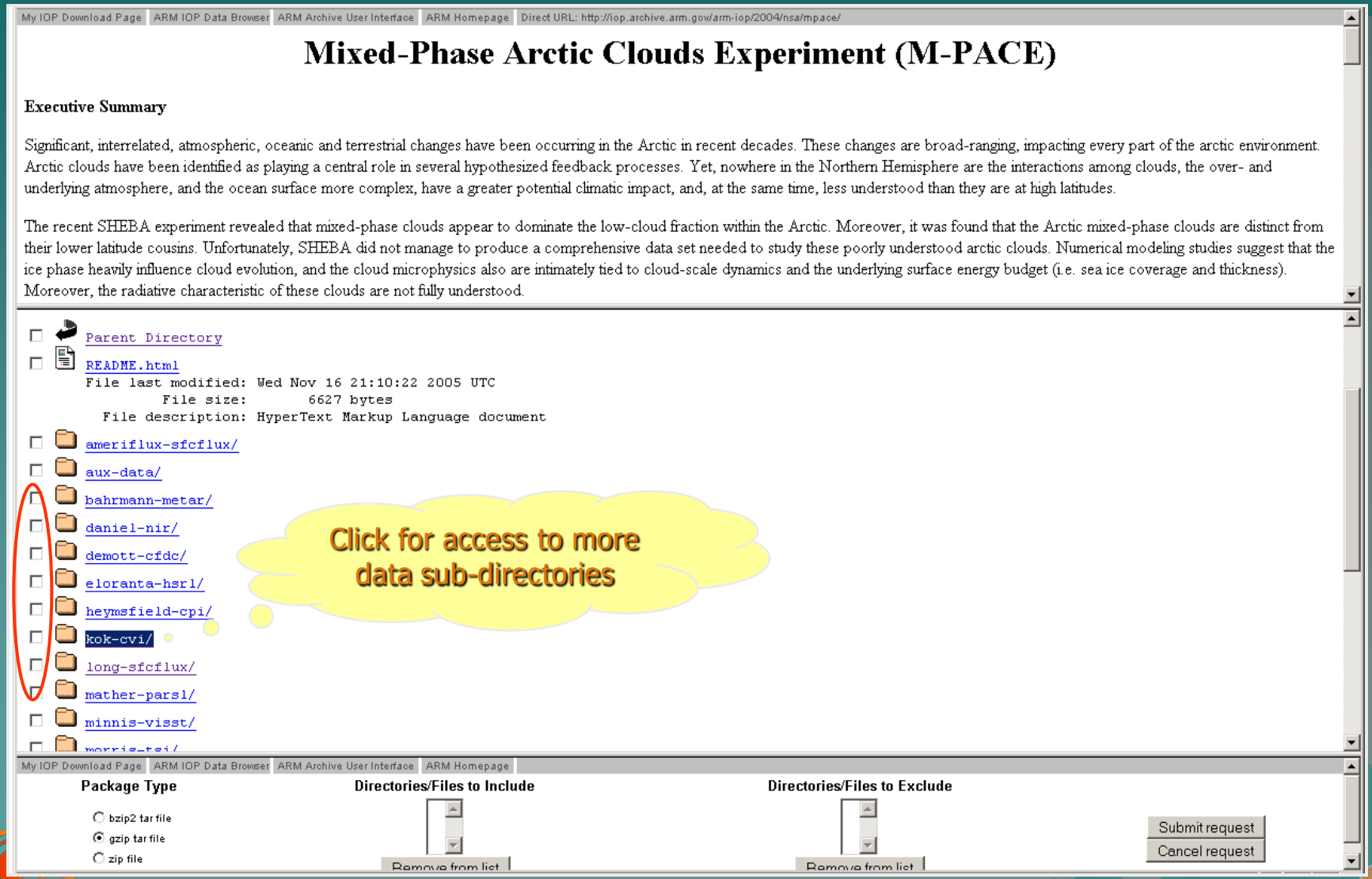
Office of Science
U.S. Department of Energy

IOP Data Browser – IOP View



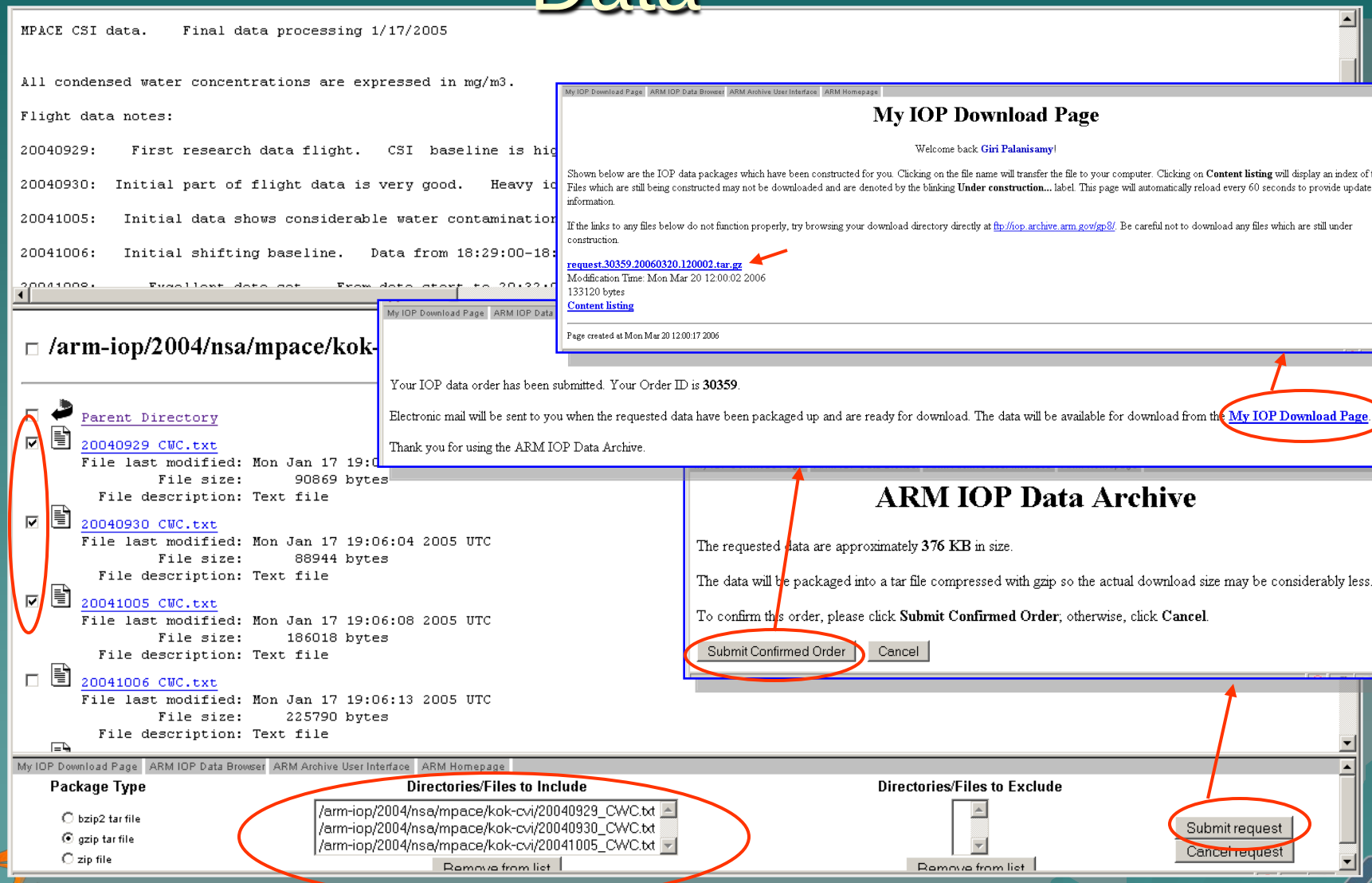
Office of Science
U.S. Department of Energy

IOP Data Browser – IOP View



Office of Science
U.S. Department of Energy

IOP Data Browser – Download Bulk Data



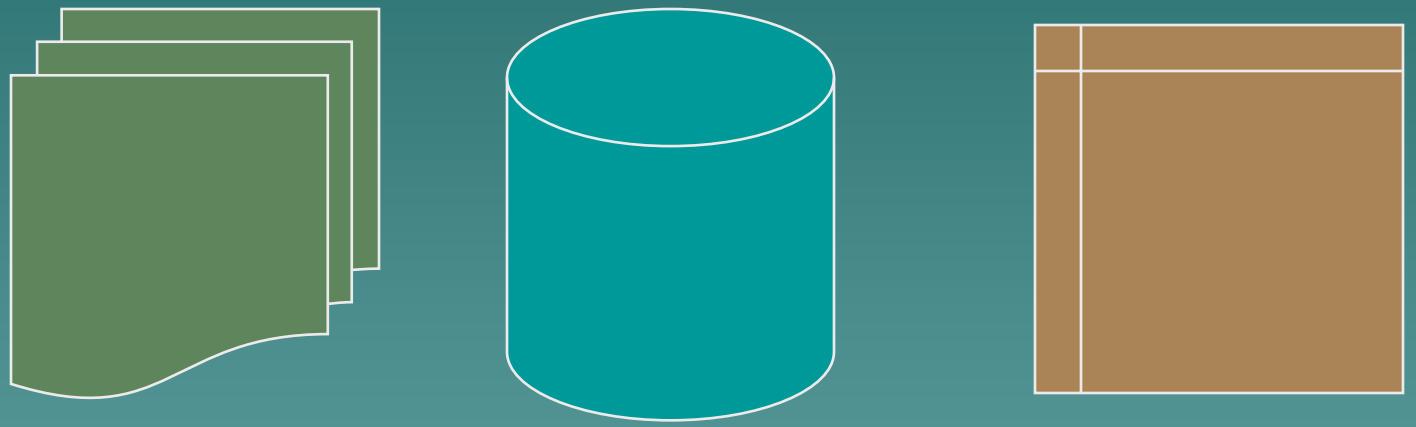
Office of Science
U.S. Department of Energy

Other ACRF 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*

Office of Science
U.S. Department of Energy

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

Office of Science
U.S. Department of Energy

ARM Data Types – more detail

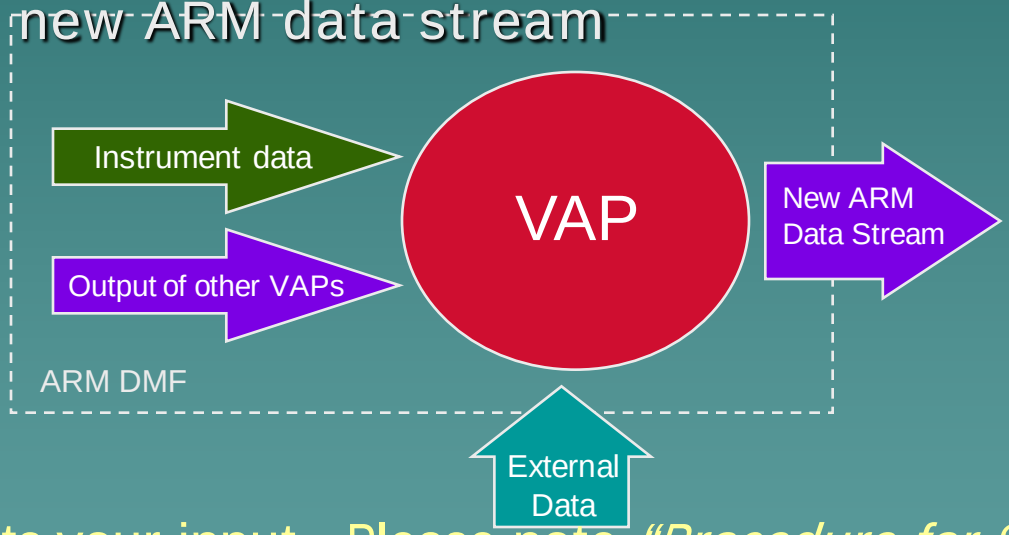
armarchive@ornl.gov
1-888-ARM-DATA

- 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
  - 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.

Office of Science
U.S. Department of Energy

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

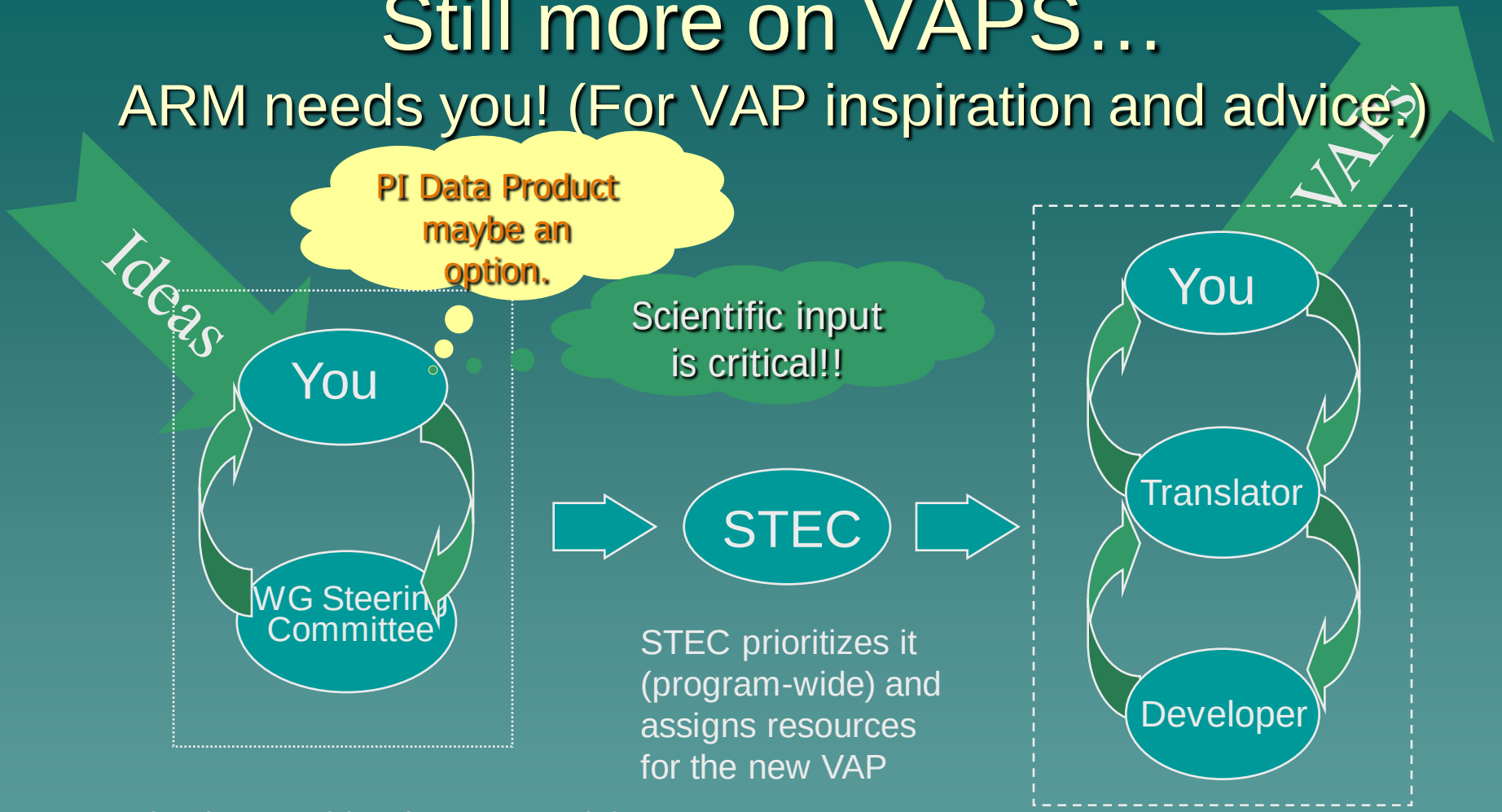


- ARM wants your input. Please note *"Procedure for Submitting Science and Research Products to the Data Archive"* at: http://www.arm.gov/data/pi_procedure.stm

Office of Science
U.S. Department of Energy

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

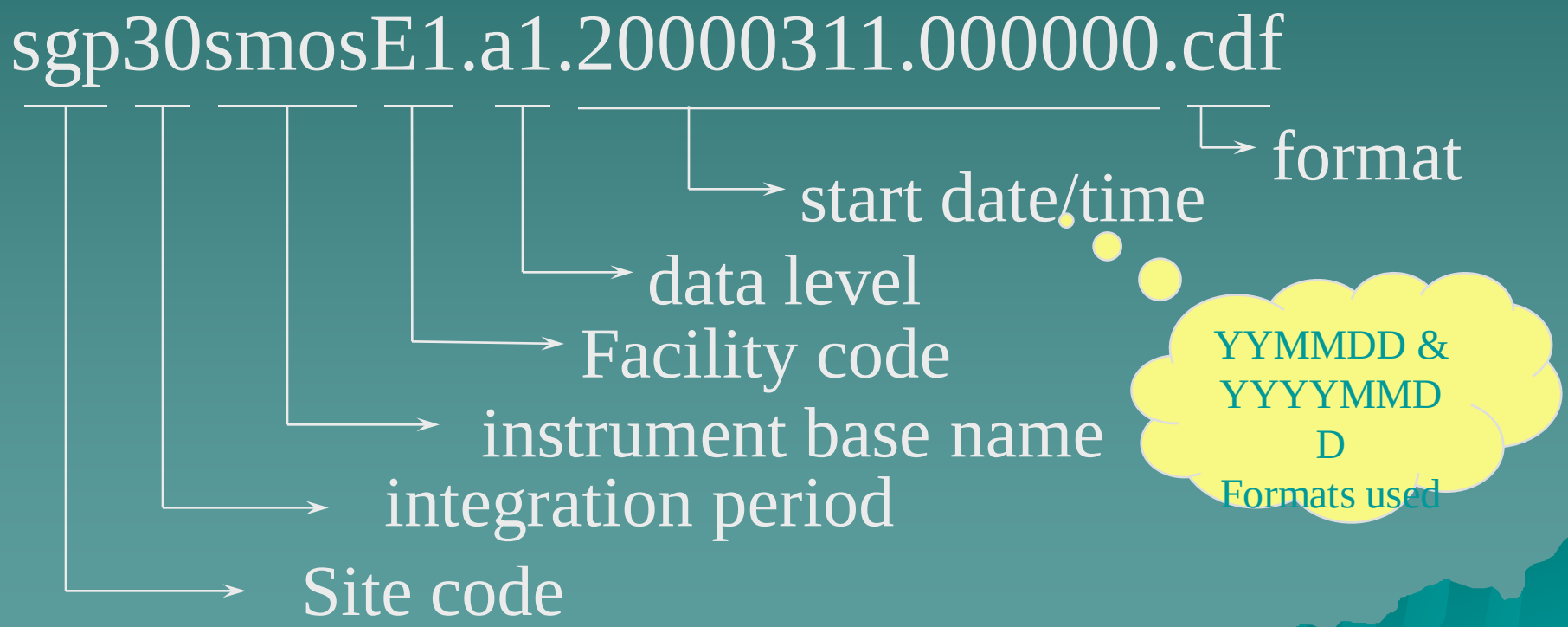
Translator works with the Scientist to further define the algorithm, and then interacts with the Developer to implement the VAP. Translator and the Scientist then evaluate and document.

Office of Science
U.S. Department of Energy

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)



Office of Science
U.S. Department of Energy

Quality Information

Office of Science
U.S. Department of Energy

Types of Quality Information

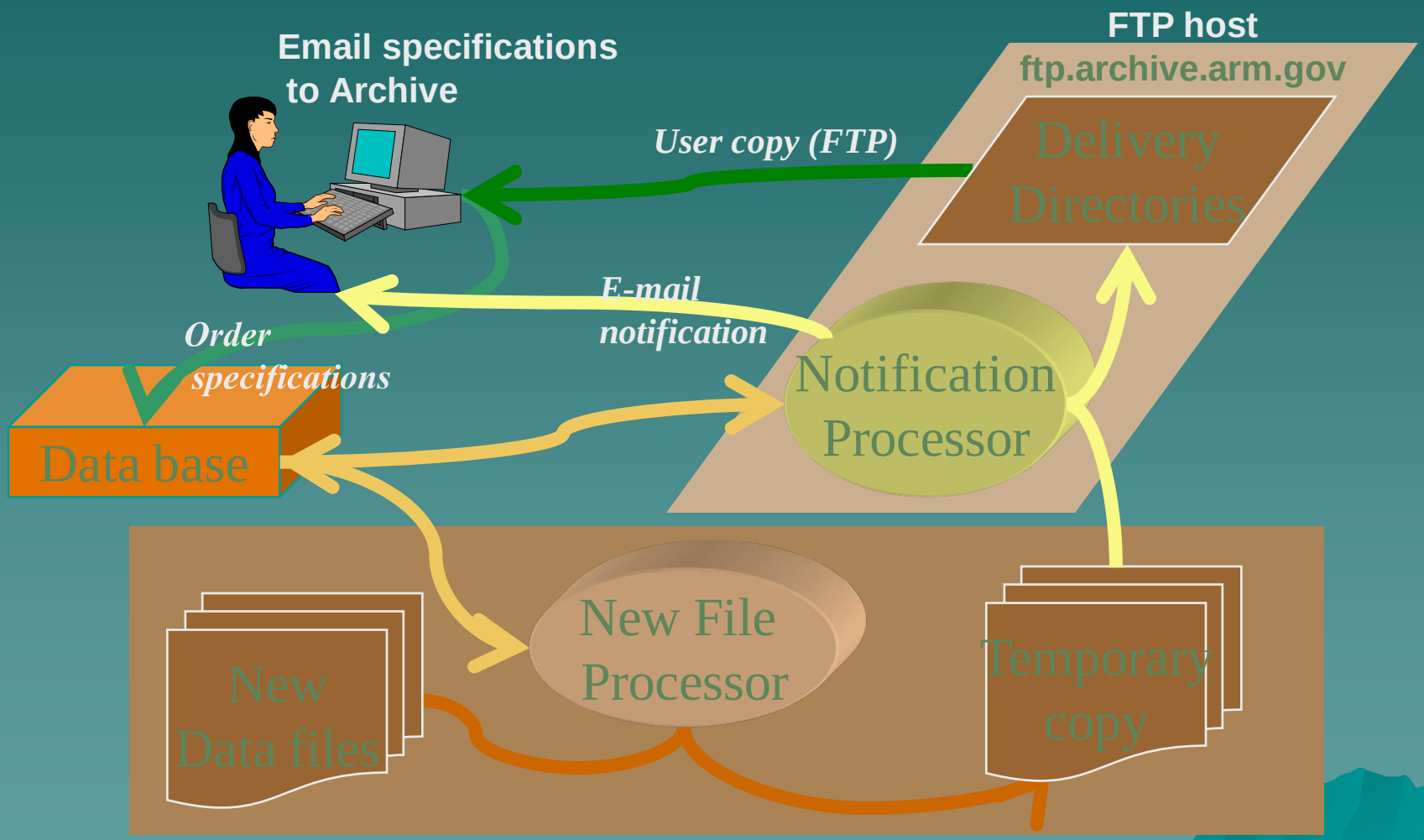
- Automated products
  - QC flags
    - inserted in data files during processing
  - Summaries of flags (data color)
- Manual products
  - Data Quality Reports (DQRs)
    - web accessible reports; delivered as html files after data requests (more later); event driven and problem-based
  - Instrument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



Office of Science
U.S. Department of Energy

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

Office of Science
U.S. Department of Energy

Archive Details

More than you want to know...?

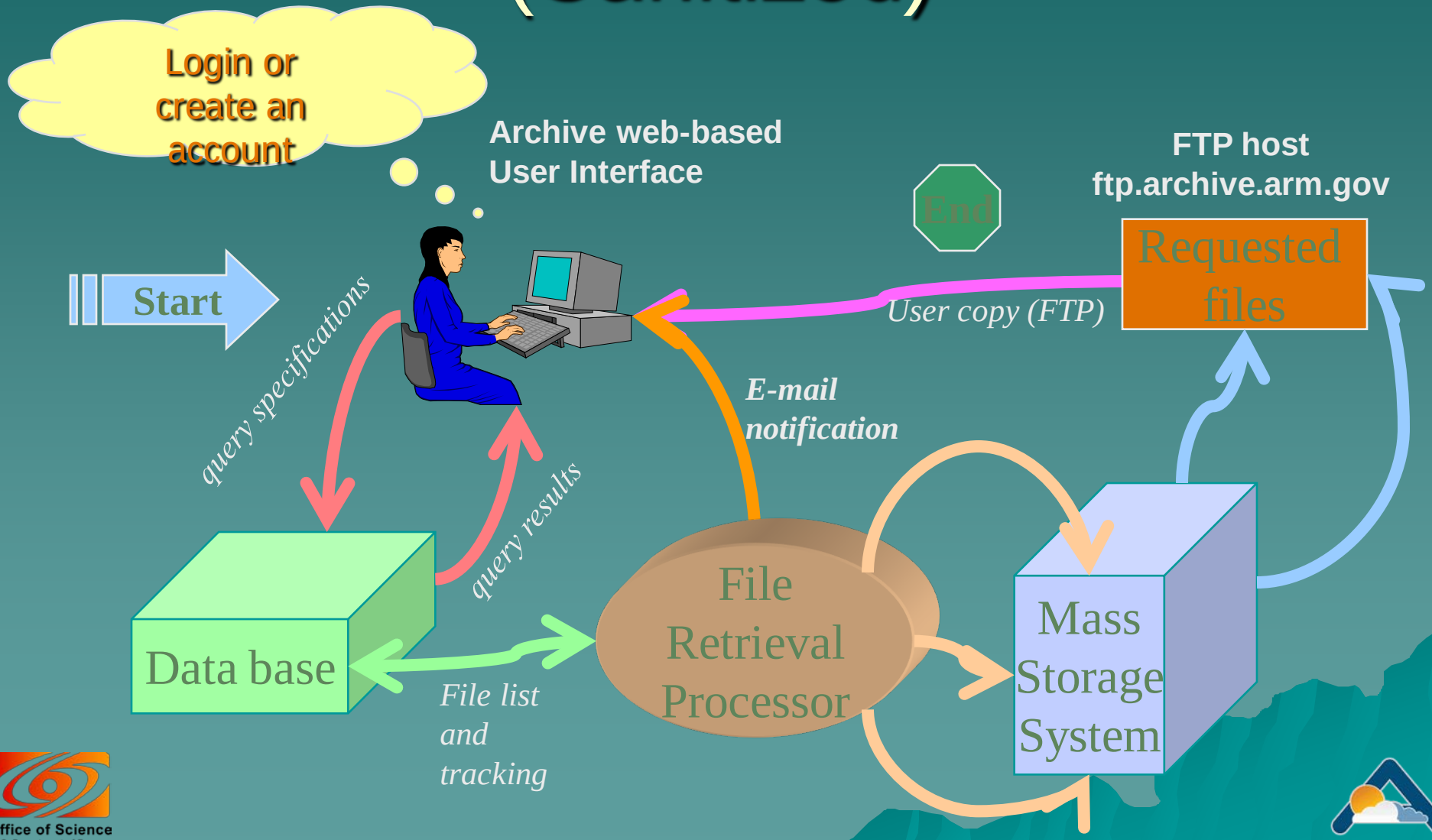
You are NOT alone...

- 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: 50-70,000 files, 2-3 TB per month
  - Total storage: 8.4 million files, 118 TB of data
- Usage
  - Yearly requests – 2 million files, 20 TB of data
  - 800-1000 different users active each year.

Office of Science
U.S. Department of Energy

You and the Archive 'Guts' (Sanitized)



Office of Science
U.S. Department of Energy

Accessing Data from the Archive

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

Office of Science
U.S. Department of Energy

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

New ARM Archive Developments
(Guided by feedback from a recent User Group meeting)

- User Group meeting held October 30, 2007
  - *12 members – ARM and non-ARM*

- The meeting covered:
  - *Planned revisions to existing Archive functionality*
  - *Planned additions to Archive functionality*
  - *New ideas and recommendations from the User Group*

Ongoing Revisions to Existing Archive Functionality

- New Login Page - Allows email address as an alternative to Archive User Name (user's choice)
- Clarifying "Account" Creation Process
- Developing more/better guidance on choosing and navigating the various user interfaces
- Developing categories for IOP data collection
- Clarifying routine (measured vs. derived), IOP, and aircraft designations

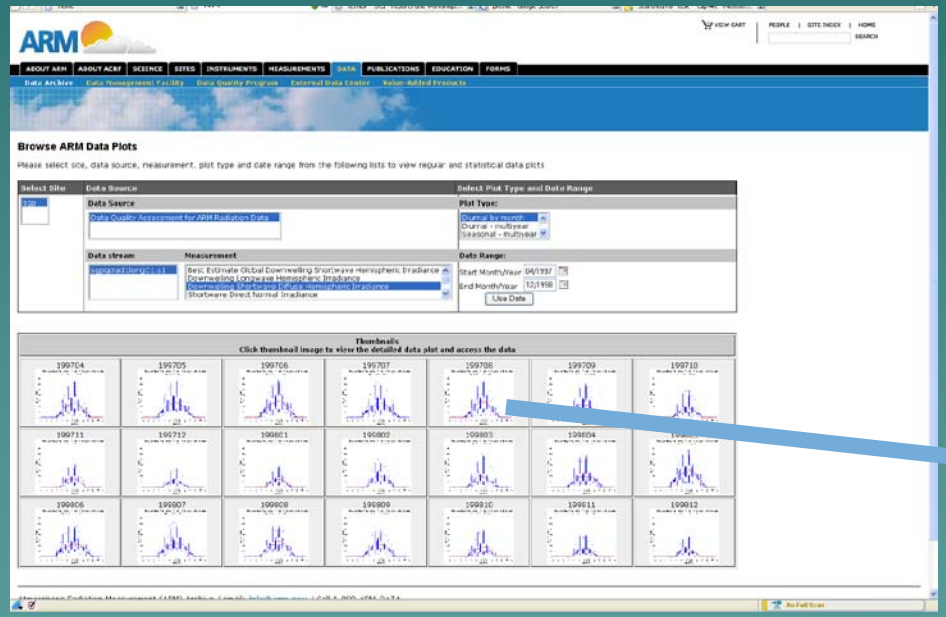
Planned Additions to Archive Functionality

- *Statistical Views of ARM Data*
  - A means of teasing potential data users
  - An alternative to basic data products
  - Being implemented only "by request" for highly polished data products
  - See Poster *4G* on Wednesday (Palanisamy *et al*.), and
  - See demo at

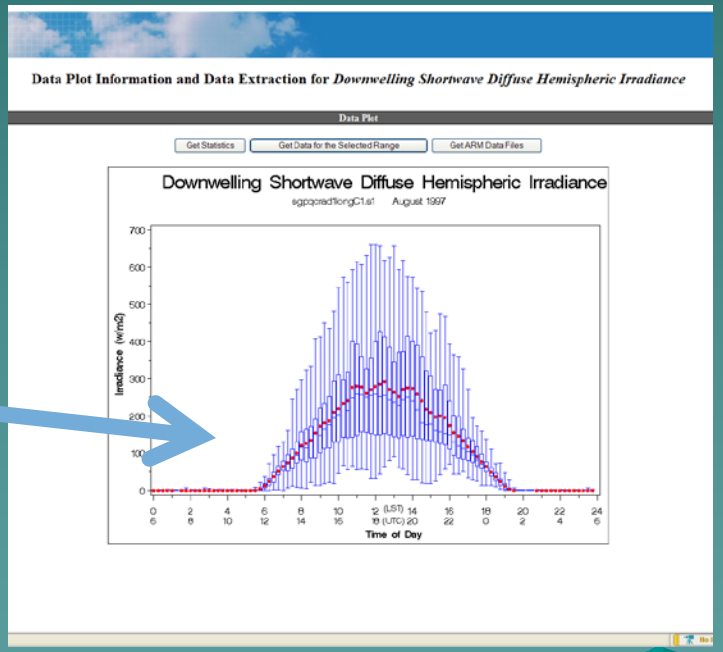
    http://www.archive.arm.gov/arm/stattnb1.jsp

Planned Additions to Archive Functionality

- *Statistical Views of ARM Data*
  - an example using QCRAD data



User interface to select thumbnails of Statistical Views



Detailed view of graph; options to order statistics, data, or data files

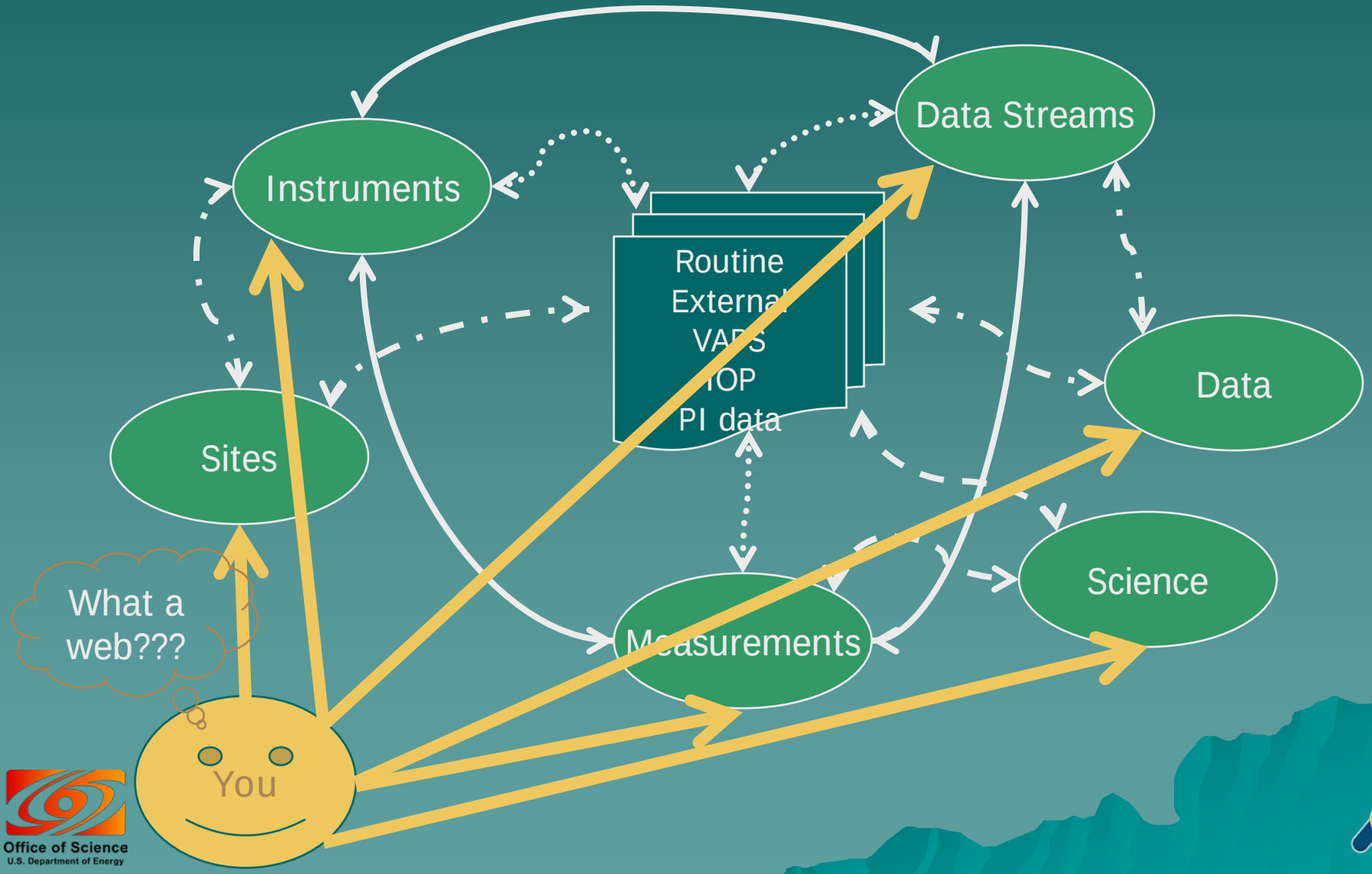
Office of Science
U.S. Department of Energy

Planned Additions to Archive Functionality

Data Extraction

- Developing a more explicit version of NCVweb for data extraction
  - Other data formats
  - Concatenated data files
- Developing an "insertion" into measurements part of Data Browser user interface so that retained measurements from a single data stream can be saved
  - Look at extending this to include the specification of a conditional query for data selection
  - Include (??) a companion file containing DQR "quality mask" to go with selected data

Wandering ARM Web “stuff”



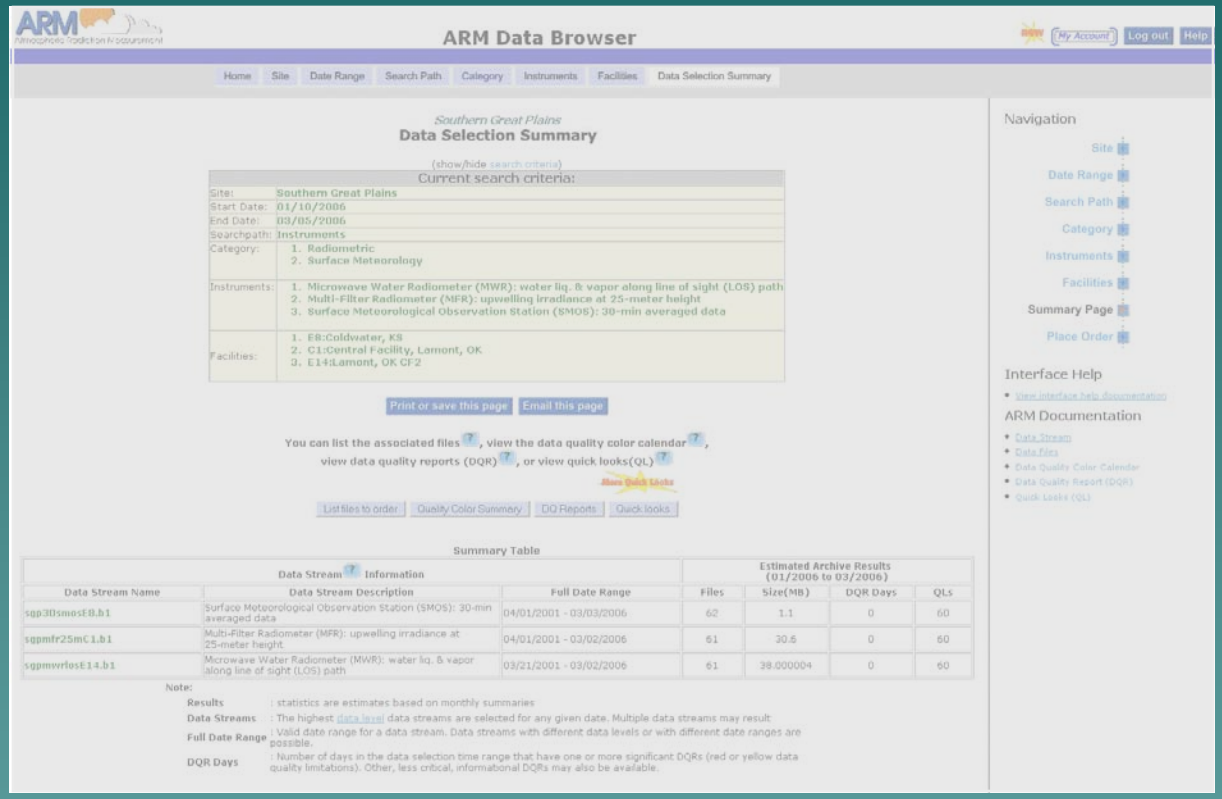
Office of Science
U.S. Department of Energy

ARM Shopping Cart Tutorial

- Available at: http://dev.www.arm.gov/data/arm_orderData_tutorial.html

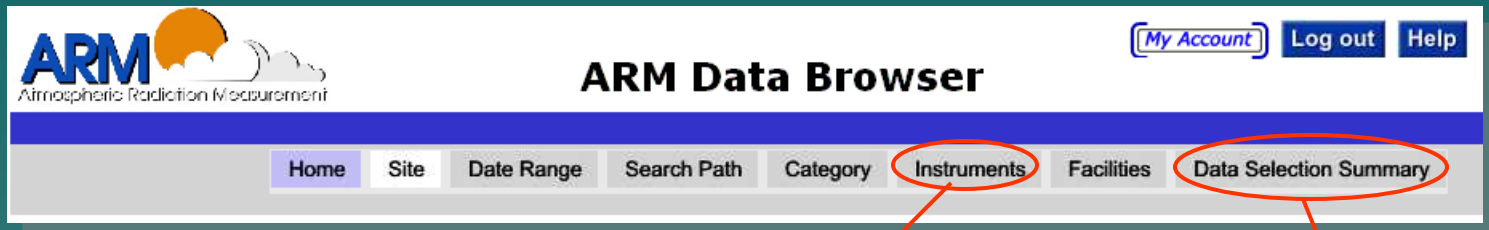
Office of Science
U.S. Department of Energy

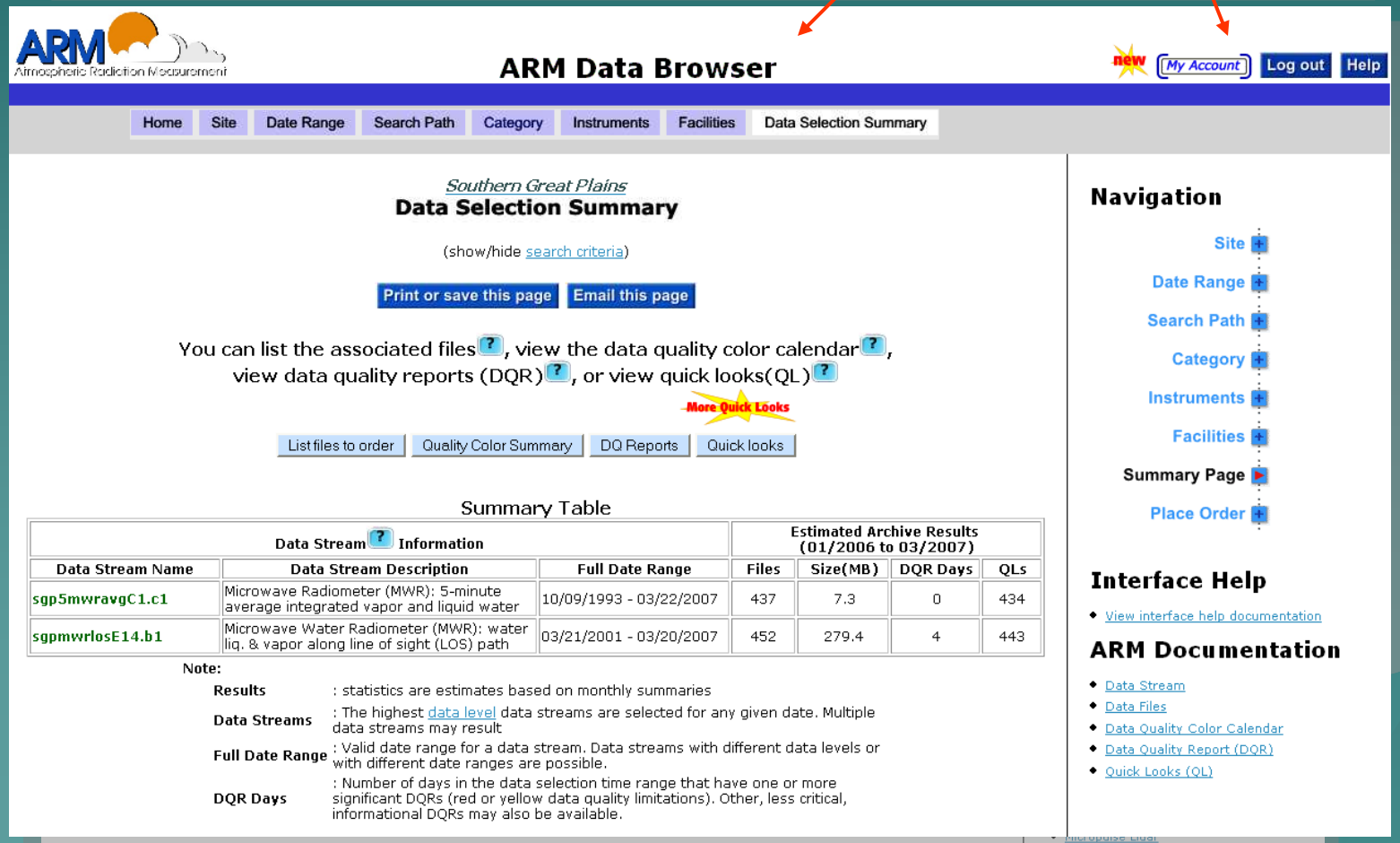
ARM Data Browser



Office of Science
U.S. Department of Energy

ARM Data Browser





Office of Science
U.S. Department of Energy

ARM Catalog Browser

Office of Science
U.S. Department of Energy

Features of Catalog Tables

A: Title of the Table; showing selection ategories

B: Links to higher selection screens

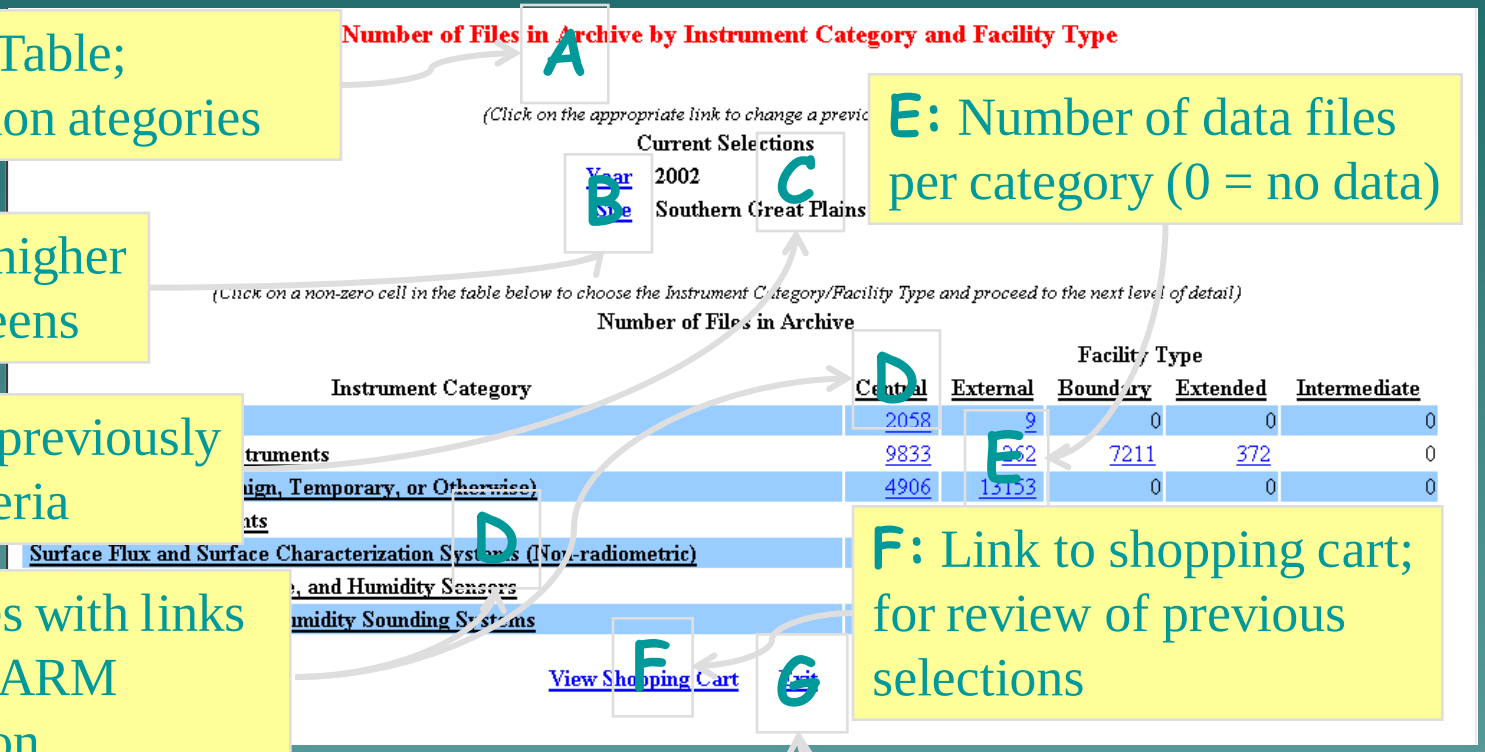
C: values of previously selected criteria

D: Categories with links to additional ARM documentation

E: Number of data files per category (0 = no data)

F: Link to shopping cart; for review of previous selections

G: Exit from interface



Office of Science
U.S. Department of Energy

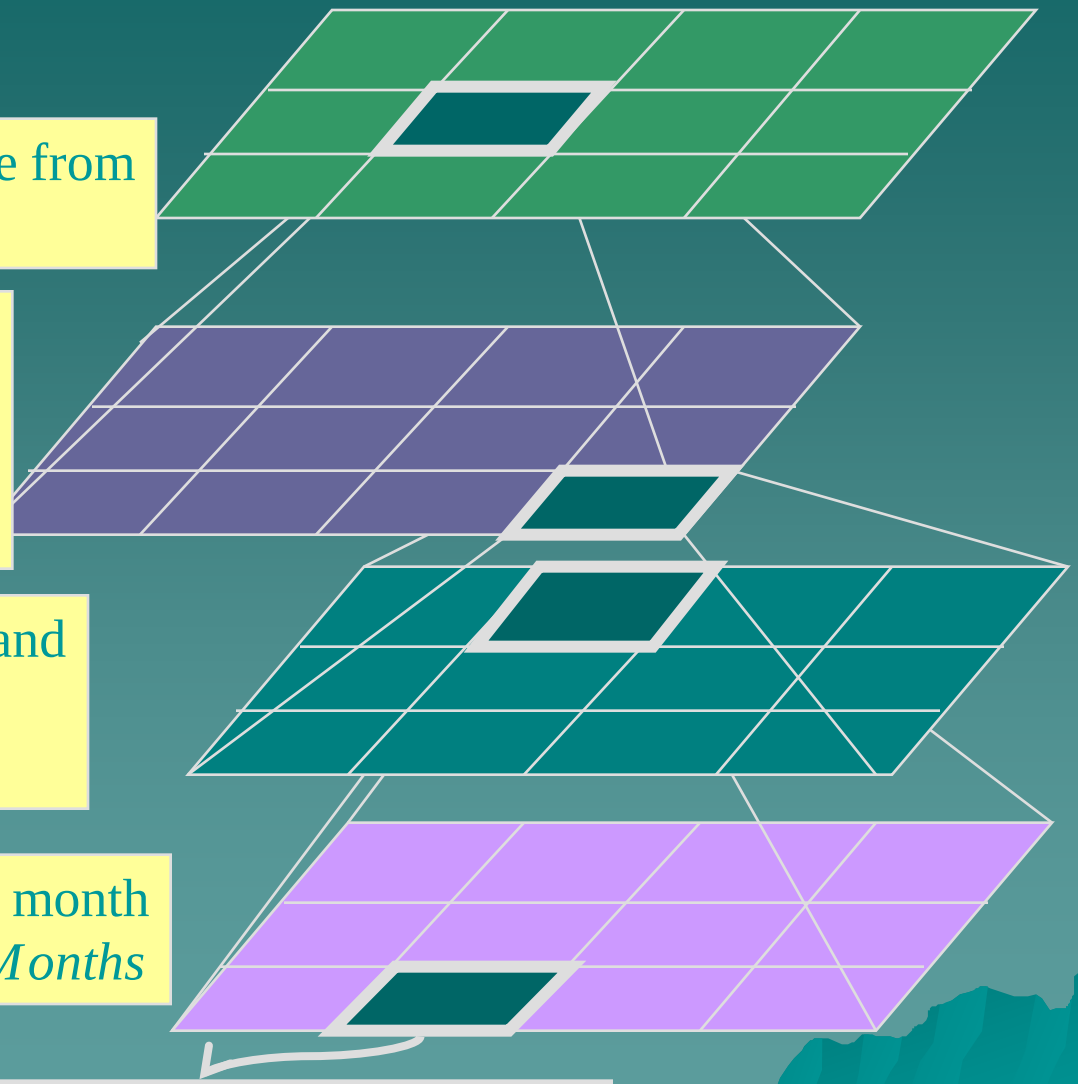
Logical Flow of Catalog Interface

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*

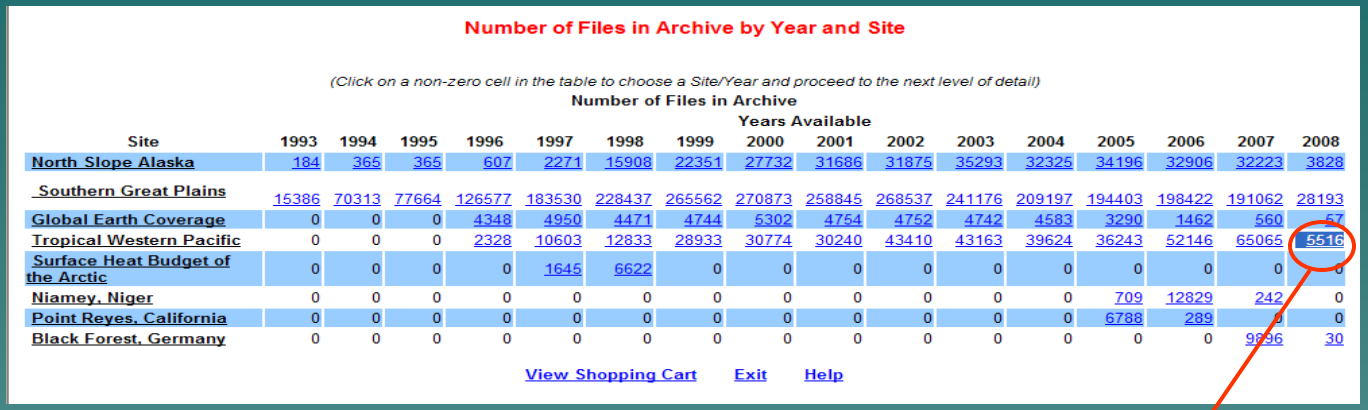
Step 4: select a facility and a month from a table of *Facilities x Months*



Data added to 'shopping cart'

Office of Science
U.S. Department of Energy

Catalog Browser (1)



Number of Files in Archive by Year and Site

(Click on a non-zero cell in the table to choose a Site/Year and proceed to the next level of detail)

Number of Files in Archive

Years Available

Site	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
North Slope Alaska	184	365	365	607	2271	15908	22351	27732	31686	31875	35293	32325	34196	32906	32223	3828
Southern Great Plains	15386	70313	77664	126577	183530	228437	265562	270873	258845	268537	241176	209197	194403	198422	191062	28193
Global Earth Coverage	0	0	0	4348	4950	4471	4744	5302	4754	4752	4742	4583	3290	1462	560	57
Tropical Western Pacific	0	0	0	2328	10603	12833	28933	30774	30240	43410	43163	39624	36243	52146	65065	5516
Surface Heat Budget of the Arctic	0	0	0	0	1645	6622	0	0	0	0	0	0	0	0	0	0
Niamey, Niger	0	0	0	0	0	0	0	0	0	0	0	0	709	12829	242	0
Point Reyes, California	0	0	0	0	0	0	0	0	0	0	0	0	6788	289	0	0
Black Forest, Germany	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9896	30

View Shopping Cart Exit Help

Number of Files in Archive by Instrument Category and Facility Type

(Click on the appropriate link to change a previous selection)

Current Selections

Year 2006

Site Tropical Western Pacific

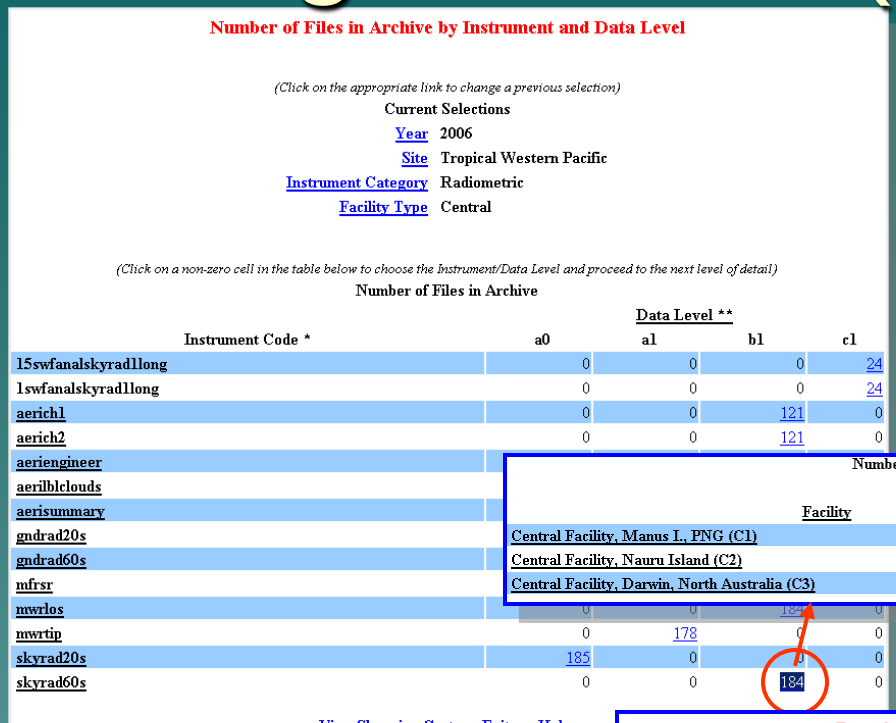
(Click on a non-zero cell in the table below to choose the Instrument Category/Facility Type and proceed to the next level of detail)

Number of Files in Archive

Instrument Category	Facility Type: Central	Facility Type: External
Aerosols	366	0
Atmospheric Profiling	748	59
Cloud Properties	1928	0
Derived Quantities and Models	121	82
Radiometric	2055	0
Surface Meteorology	183	0

Office of Science
U.S. Department of Energy

Catalog Browser (2)



Number of Files in Archive by Instrument and Data Level

(Click on the appropriate link to change a previous selection)

Current Selections

Year 2006
Site Tropical Western Pacific
Instrument Category Radiometric
Facility Type Central

(Click on a non-zero cell in the table below to choose the Instrument/Data Level and proceed to the next level of detail)

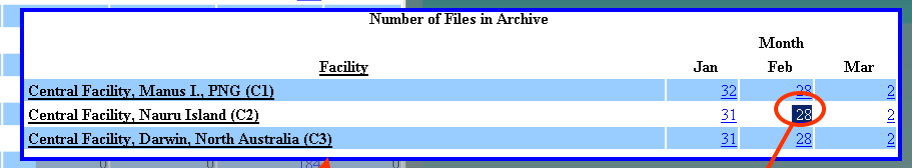
Number of Files in Archive

Instrument Code *	Data Level ** a0	a1	b1	c1
15swfanalskyrad1long	0	0	0	24
1swfanalskyrad1long	0	0	0	24
aerich1	0	0	121	0
aerich2	0	0	121	0
aeriengineer				
aerilblclouds				
aerisummary				
gndrad20s				
gndrad60s				
mfrsr				
mwrlos	0	0	184	0
mwrtip	0	178	0	0
skyrad20s	185	0	0	0
skyrad60s	0	0	184	0

View Shopping Cart Exit Help

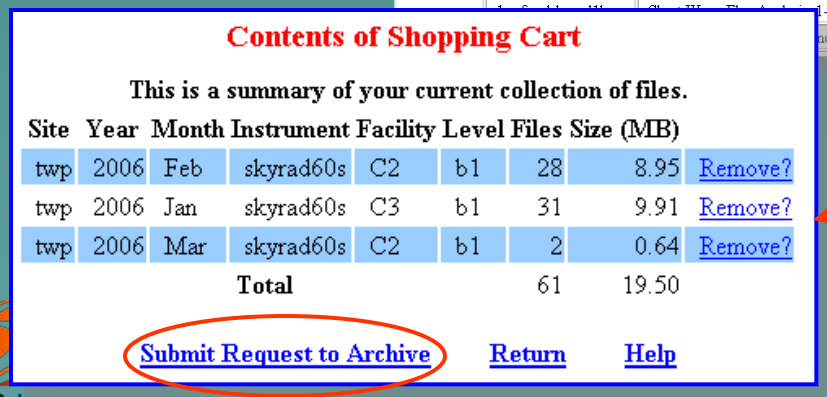
* Instrument Code Descriptions

15swfanalskyrad1long	Short Wave Flux Analysis: 15-min resolution on SKYRA



Number of Files in Archive

Facility	Month Jan	Feb	Mar
Central Facility, Manus I., PNG (C1)	32	28	2
Central Facility, Nauru Island (C2)	31	28	2
Central Facility, Darwin, North Australia (C3)	31	28	2



Contents of Shopping Cart

This is a summary of your current collection of files.

Site	Year	Month	Instrument	Facility	Level	Files	Size (MB)	
twp	2006	Feb	skyrad60s	C2	b1	28	8.95	Remove?
twp	2006	Jan	skyrad60s	C3	b1	31	9.91	Remove?
twp	2006	Mar	skyrad60s	C2	b1	2	0.64	Remove?
			Total			61	19.50	

Submit Request to Archive Return Help

Description of Archive Files

(Click on the appropriate link to change a previous selection)

Current Selections

Year 2006
Site Tropical Western Pacific
Instrument Category Radiometric
Facility Type Central
Instrument skyrad60s
Data Level b1
Facility Central Facility, Nauru Island
Month Feb
Files 28
Total Size (MB) 8.950

Add files to your Shopping Cart?

Click "Yes" to add the current set of files to your collection. Either button returns to the previous page (Facilities/Month).

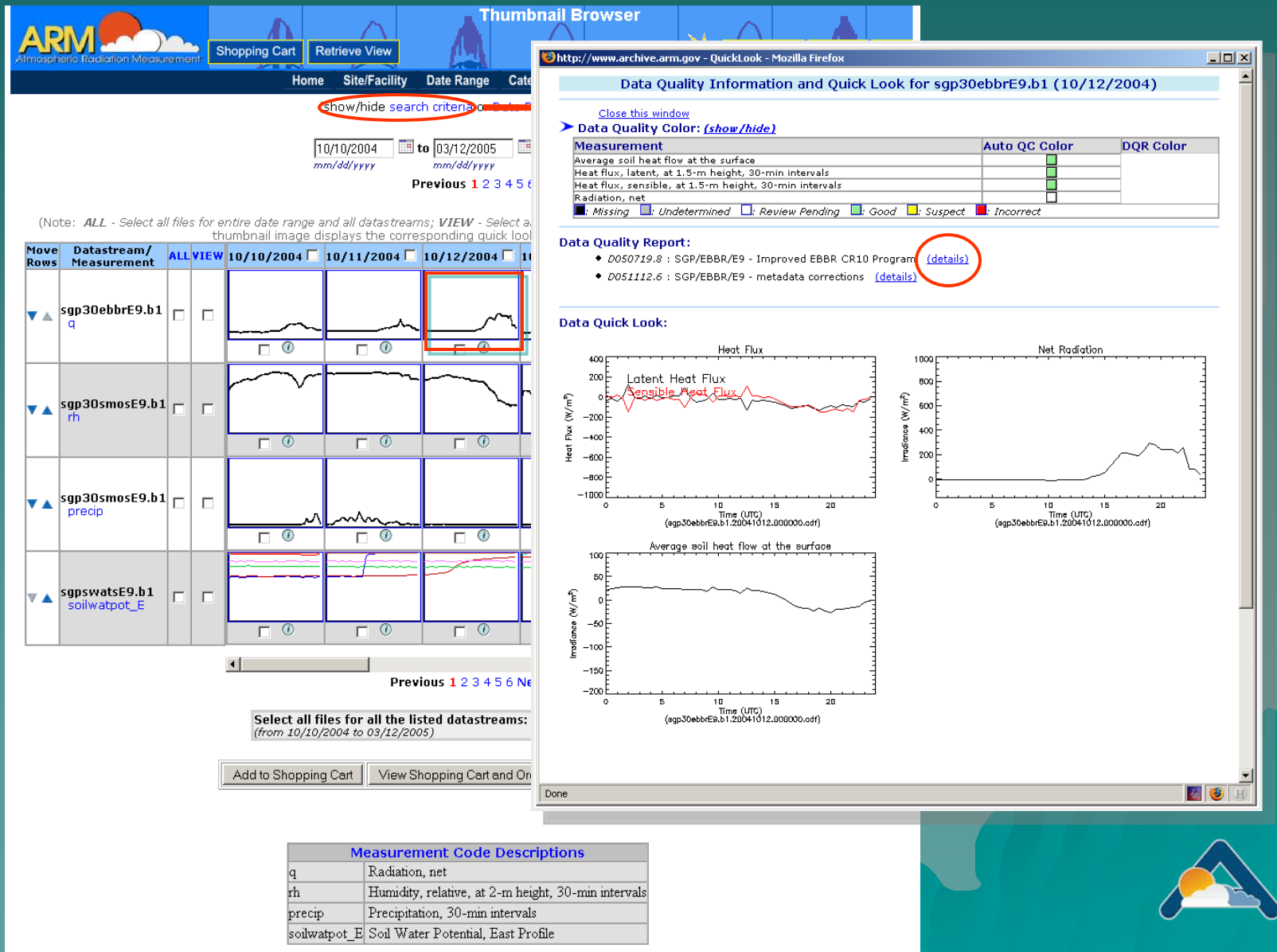
No Yes

Office of Science
U.S. Department of Energy

ARM Thumbnail Browser

Office of Science
U.S. Department of Energy

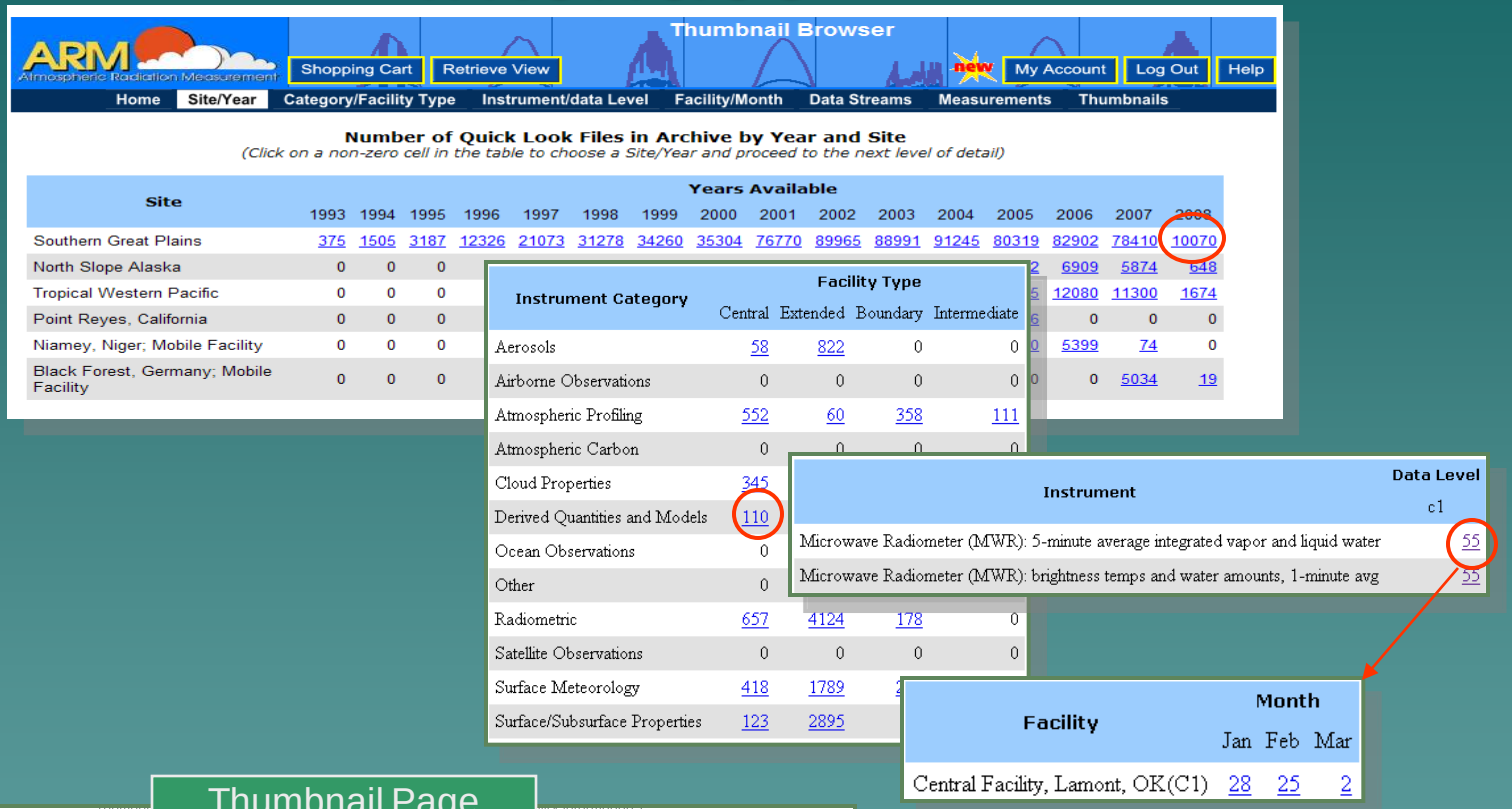
Display Thumbnails



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

Office of Science
U.S. Department of Energy

Thumbnail Browser – Catalog Interface



Office of Science
U.S. Department of Energy