The NARSTO Data and Information Sharing Tool

Les A. Hook and Sigurd W. Christensen
NARSTO Quality Systems Science Center

Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN, U.S.A. 37831-6335

NARSTO 2000
Tropospheric Aerosols: Science and Decisions in an International Community
Queretaro, Mexico

October 25, 2000

Setting the NARSTO Stage

NARSTO *is a public/private alliance, whose membership spans government, the utilities, industry, and academe throughout Mexico, the United States, and Canada.*

The **NARSTO mission** *is to plan, coordinate, and facilitate comprehensive, long-term, policy-relevant scientific research and assessment of primary and secondary pollutant species emitted, formed, transformed, and transported in the troposphere over the North American continent.*

How does NARSTO really work?

Member organizations agree to support the stated mission. Support may be in the form of participation in workshops and meetings, financial support, or contribution of in-kind resources.

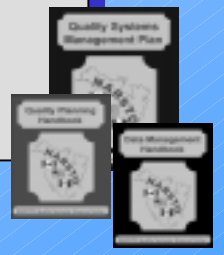
Research and assessment initiatives that support the mission may request to become NARSTO Technical Programs. ***Technical Programs agree to follow certain quality assurance and data management guidelines and send their data to the NARSTO data archive.***

The **Quality Systems Science Center (QSSC)** *coordinates quality systems and data management activities, provides assistance to technical programs, and archives NARSTO data.*

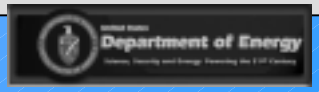
QSSC Web Site:
htttp://cdiac.esd.ornl.gov/programs/NARSTO/

NARSTO Web Site:
http://www.cgenv.com/Narsto/

Permanent Data Archive Web Site:
http://eosweb.larc.nasa.gov/

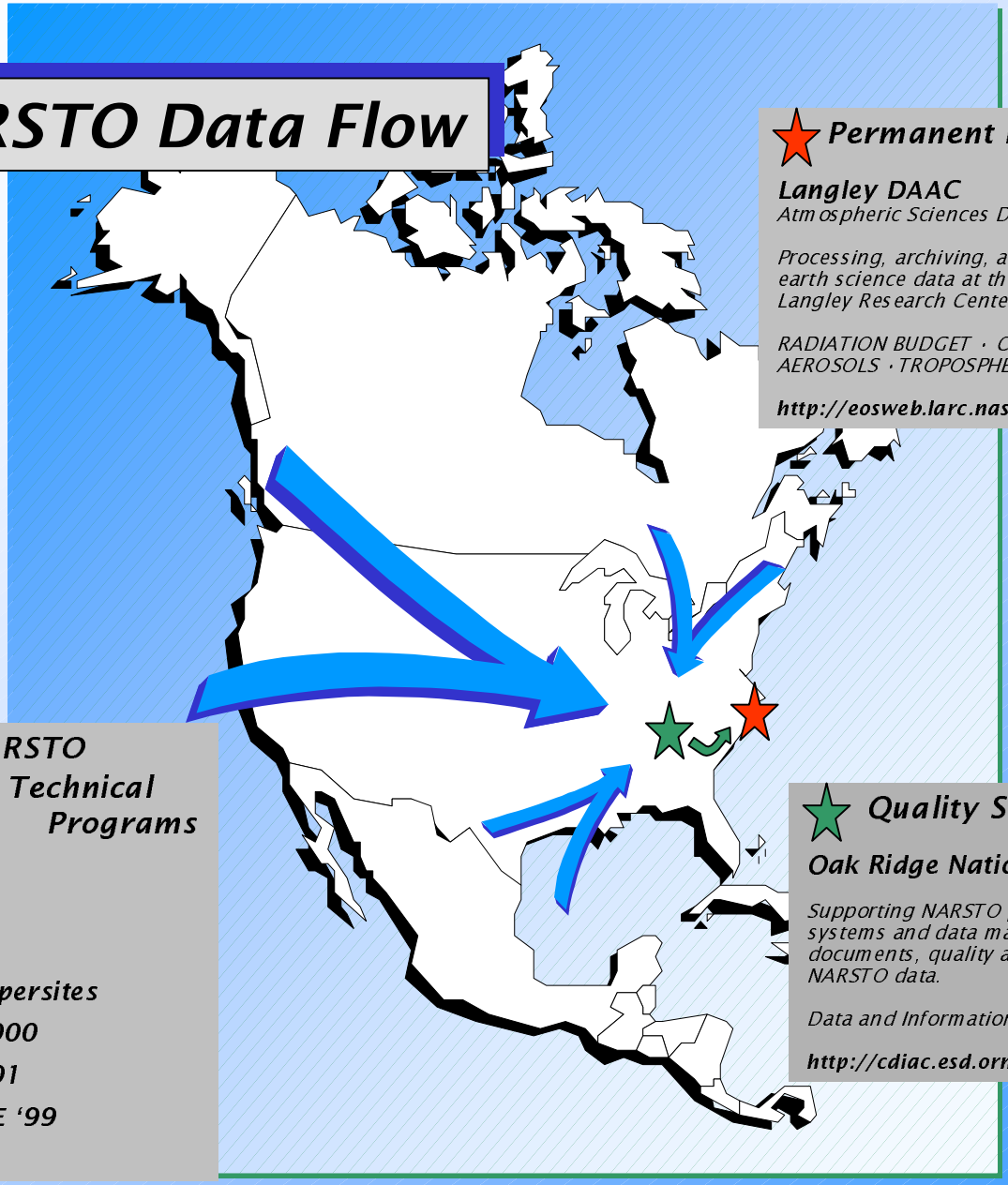


The **U.S. Department of Energy** *sponsors the Quality Systems Science Center at CDIAC, Environmental Sciences Division, Oak Ridge National Laboratory, as a portion of its contribution to NARSTO.*



QSSC Contacts
Les Hook / Director and QA Coordinator / 865-241-4846 / hookla@ornl.gov
Meng-Dawn Cheng / Chief Scientist / 865-241-2049 / chengmd@ornl.gov
Sig Christensen / Data Coordinator / 865-574-7394 / christensen1@ornl.gov

NARSTO Data Flow



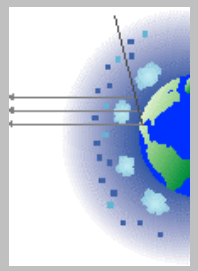
Permanent Data Archive

Langley DAAC
Atmospheric Sciences Data Center

Processing, archiving, and distributing earth science data at the NASA Langley Research Center.

RADIATION BUDGET · CLOUDS · AEROSOLS · TROPOSPHERIC CHEMISTRY

http://eosweb.larc.nasa.gov/



NARSTO Technical Programs

- CRPAQS
- CCOS
- NE-OPS
- EPA PM Supersites
- TexAQS 2000
- Pacific 2001
- NASHVILLE '99

Quality Systems Science Center

Oak Ridge National Laboratory

Supporting NARSTO projects with quality systems and data management planning documents, quality assuring and archiving NARSTO data.

Data and Information Sharing Tool (DIST)

http://cdiac.esd.ornl.gov/programs/NARSTO/

Data Archiving is Important to the Success of the NARSTO Program.

- Timely sharing of and access to quality assured data and research products by the scientific community is essential.
- Long-term usefulness will be realized by ensuring that data are adequately described and in consistent and documented formats.

Data archive documentation should be sufficient to satisfy the "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.)

Given the scale of current and future data collection efforts, the data volume and diversity, the number of NARSTO studies, and data archiving requirements, we must have a very efficient archiving process or we will soon be buried by data that will not be readily identifiable, retrievable, or understandable and of little use to future data users.

I will describe our approach for meeting this challenge.

A data and information management tool that:

- Supports data and metadata from collection through archiving.
- Facilitates sharing of data among project investigators and sponsors.
- Keeps access control in the hands of investigators or other data providers.
- Is simple to use with good help files.
- Is a web-based system that is inexpensive to implement and maintain.
- Utilizes templates with picklists for entry of metadata.
- Minimizes re-entry of metadata -- enter it once and be done.
- Links data, metadata, and documents. Tool is data-format independent.
- Facilitates flow of data and metadata from projects to the QSSC and the Permanent Data Archive.

NARSTO Data and Information Sharing Tool (DIST)

What are Metadata?

- Metadata are data about data. Metadata help you locate data.
- Metadata tell the who, what, when, where, why, and how the data were collected or generated by an investigator.
- A metadata standard is a list of metadata elements agreed to by a community of users.

Proposed

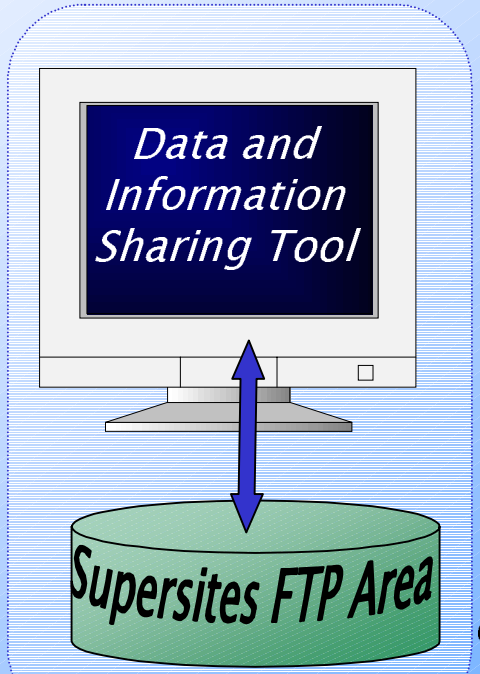
U.S. EPA Particulate Matter Supersites Program Data Flow

PM Research Sites

- New York
- Baltimore
- Pittsburgh
- Houston
- St. Louis
- Fresno
- Los Angeles
- Atlanta

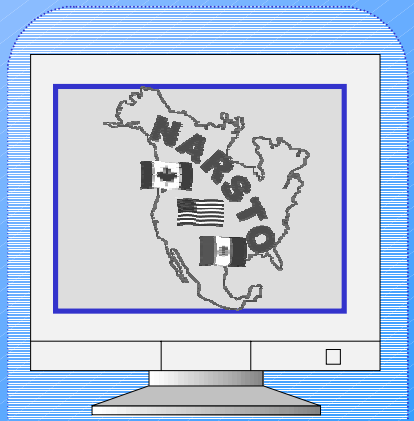
Consistent Metadata

Quality Assured Data



Consistent Metadata

Consistently Formatted Data



Permanent Data Archive

Integrated Data Analysis Project

Keys to Success

- Defined Consistent Metadata
- Documented Data Formats
- Efficient Sharing of Data
- Searchable Index for Retrieving Data
- Web-based Technology

Data Processing Steps

Sites Compile Data

Sites Link Data and Metadata to Share in DIST

Final QA Checks of Data and Metadata by QSSC

Sites Input Metadata in DIST

Sites Output Data and Metadata in Archive Format

Archive Data and Metadata at Langley DAAC

Metadata

Consistent Metadata → YES → YES

Data Format

Native > Site Specific, Special Purpose → Site Specific Data Exchange Standard Special Purpose → Data Exchange Standard Special Purpose

Data Validation Level

Level 0 and 1 → Level 1 and 2 → Level 2 +

Timing

As Collected → 9 Months → 15 Months

Data Access

Site Data Originators and Users → Site and Supersites Program → Public

Using the DIST

*Investigator uses **Metadata Editor** to create **metadata files** with links to data products.*

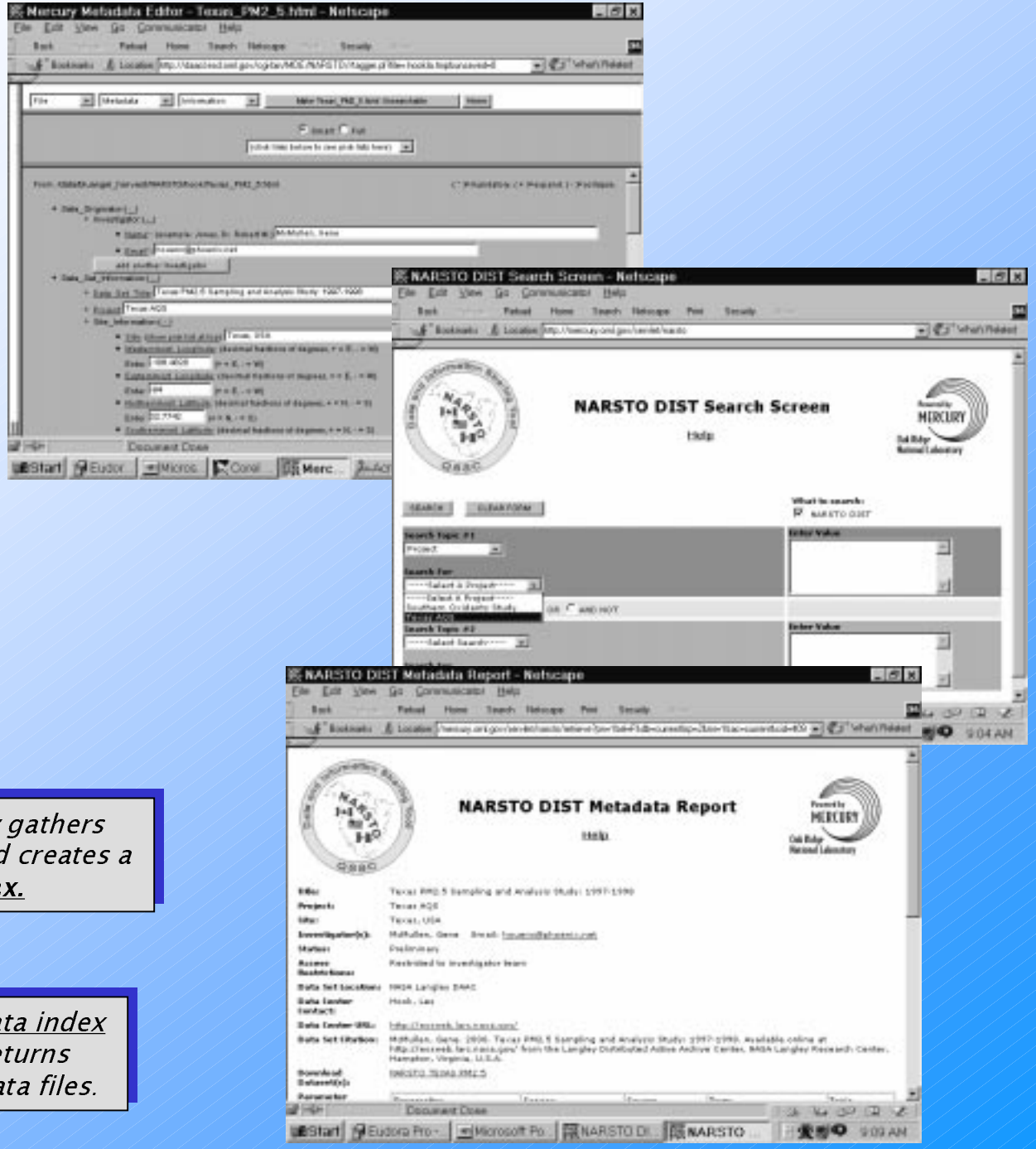
***Controls access** to metadata and data files.*

*Creates **Data Exchange Standard template** for input of data into archive format.*

*Creates **Archive Metadata Files**.*

Metadata Harvester** periodically gathers information from metadata files and creates a **Searchable Metadata Index.

***Search Tool** searches the metadata index using various strategies and returns metadata records with links to data files.*



DIST Components

Metadata Editor

- Provides web-based entry tool with picklists for entering metadata.
- Encourages input of metadata early in data management process.
- Provides control of picklist contents for project specific entries.
- Creates metadata files that link data files and documentation.
- Outputs metadata in format for input by Permanent Data Archive.
- Outputs Data Exchange Standard template for formatting and inputting data.

Metadata Harvester

- Extracts or "harvests" metadata from the metadata files created with the Metadata Editor.
- Originator controls access to the metadata files and linked data files.
- Data files may be located anywhere on the Internet.
- No special software is needed by the data provider.
- Metadata are automatically loaded into the Searchable Metadata Index.

Searchable Metadata Index

- Search engine performs free-text searches of the entire metadata index.
- Search engine performs fielded searches as selected from the Search Topic picklists.
- Boolean searches are permitted.
- Temporal and spatial field search capabilities are supported.
- Fields are as entered in the Metadata Editor.
- Project specific "web-browse trees" can be created.
- Search results are metadata records and respective links to data files.

DIST System Specifics

- Includes file manager and good help files.
- Convenient interface is similar to many existing Internet search engines.
- Is implemented using Internet standards, including XML.
- Supports international metadata standards, including FGDC and Z39.50.
- Is based on COTS software, adds value to the COTS software.
- Part of the ORNL Mercury Consortium, a group of independent data project that work together to reduce individual costs.
- Is inexpensive to implement and maintain.
- Tool is data-format independent -- can link to data, documents, images, etc.
- Metadata entered by a project flows through to Permanent Data Archive.
- Facilitates flow of metadata and data to QSSC and Permanent Data Archive.

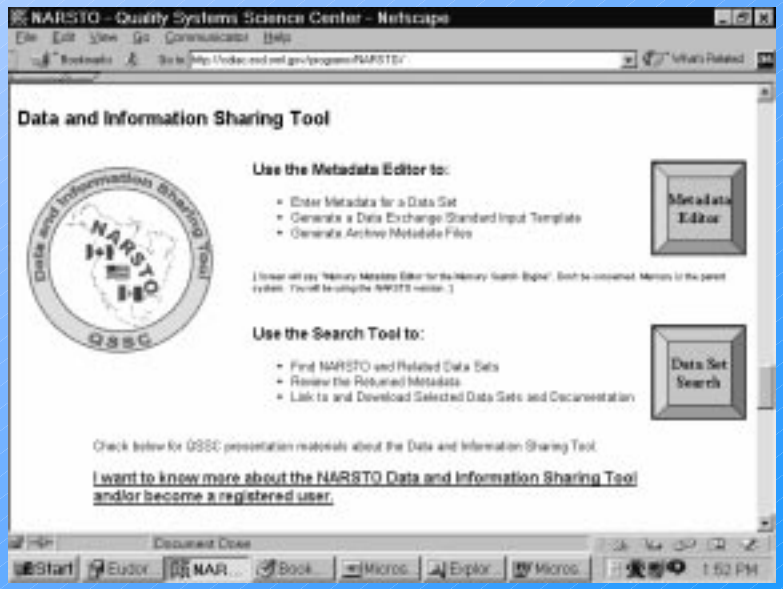
Check Out the Data and Information Sharing Tool

Access DIST from QSSC Web Site

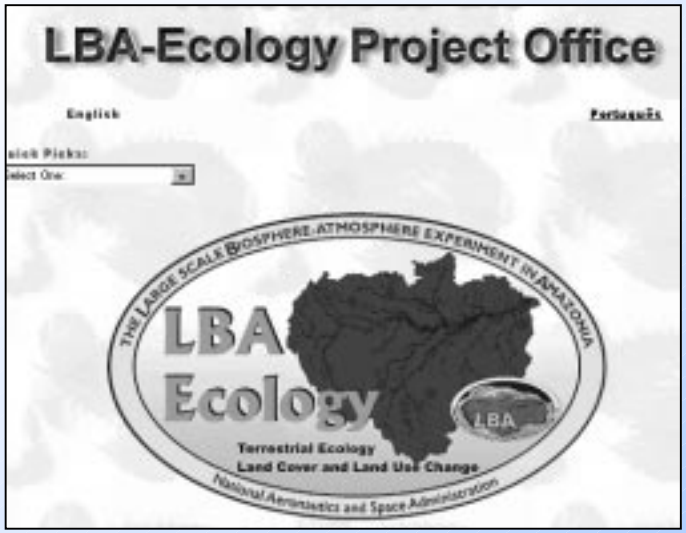
- You are welcome to enter the Metadata Editor as a guest, but be aware that any file you create will not be saved.
- The Search Tool does not require a userid or password.
- There are links to data at the Permanent Data Archive (e.g., Project=Texas AQS). Try the Langley DAAC's search and order system.
- Consider how the Data and Information Sharing Tool might benefit your project and enable more efficient data sharing and archiving.
- Customization for a project is possible. Projects can administer their own DIST implementation, control access, and customize picklists.
- We welcome suggestions for improvements and enhancements.

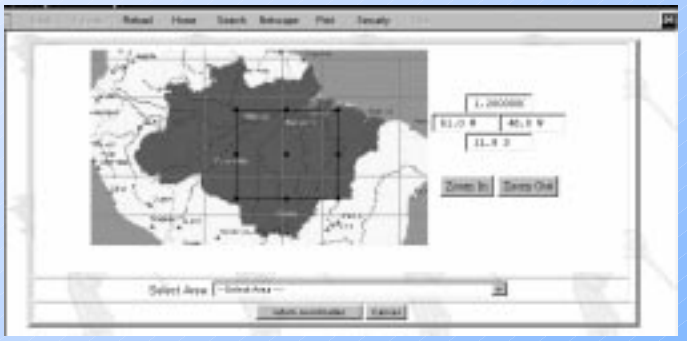
QSSC Web Site:

htttp://cdiac.esd.ornl.gov/programs/NARSTO/



Search with Temporal and Spatial Capabilities





Beija-flor – The Search Engine for the LBA Project - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Stop

Bookmarks Location: http://mercury.ornl.gov/lba/ What's Related

Welcome to *Beija-flor*
The Search Engine for the LBA Project
V1.1

Help Users Guide (PDF)

Other commonly used Links are available at the bottom of this page.

Where to search: LBA
SEARCH

Full Text or Fielded Search
CLEAR TEXT

Search Topic	Search For	Enter a Value	Connector
Entire Document	-----Not Available----		AND
-----Select Search-----	-----Not Available----		

(Ignore Location) Spatial Search
CLEAR SPATIAL

Spatial Search Method: Overlaps Enclosed Within

Northernmost 90
Westernmost -180 Map... Easternmost 180
Southernmost -90

Select Area
---Select Area ----

(Ignore Date) Temporal Search
CLEAR TEMPORAL

during : Oct. 05 1980 through Oct. 05 2000

SEARCH

LBA-Ecology Project Office | LBA Project Office | LBA-E Investigations | LBA-E Activities | LBA-E/ESIP | LBA Metadata Editor
ORNL DAAC | Calendars | Related Web Sites | Find NASA People | Find ORNL People | Acronyms | Email Us

Document: Done

Start Eud... NAR... Boo... Micr... Expl... Micr... Beij... 2:39 PM