Multi-Instrument Tools and Services to Access NASA Earth Science Data from the GSFC Earth Sciences Data and Information Services Center (GES DISC)

http://disc.gsfc.nasa.gov/

March 30, 2010

Steve Kempler, Greg Leptoukh, Chris Lynnes
Steven.J.Kempler@nasa.gov

With contributions from:
ESDIS Project
Presentation Purpose

- Describe multi-instrument tools and services that facilitate access and usability of NASA Earth science data at Goddard Space Flight Center (GSFC)
EOSDIS Facilities

Data centers, collocated with centers of science discipline expertise, archive and distribute standard data products produced by Science Investigator-led Processing Systems (SIPSs).

Key

- Data Center
- SIPs

EOSDIS – Earth Observing System Data and Information System
Definitions

- **Data Center** - Ingests, processes, archives, distributes, and manages data acquired from remote sensing instruments. Data “has no significance beyond its existence” *

- **Information Center** - Ingests, processes, archives, distributes, and manages value-added data products, and value-added data processes. “Data that has been given meaning by way of relational connection” *

- **Knowledge Center** (not yet) - “Knowledge is the appropriate collection of information, such that its intent is to be useful” *

- **Data and Information Services Center** - Provides tools that further the use and usefulness of data and information

Each further enables researchers to do their work

The GES DISCs mission is to maximize NASA’s investment benefit by providing data and services that enable people to fully realize the scientific and educational potential of global climate data.

In Short...

The GES DAACs mission is to:
ENABLE EARTH SCIENCE, APPLICATIONS, and EDUCATION
Summary of Expertise

Software engineering - In-house expertise and experience that understands the best advanced technologies to further mature data management system usability and efficiency

- Science data management expertise – Doctorate-level scientists in interdisciplinary Earth sciences who collaborate with researchers to develop sophisticated Web-based tools to facilitate comprehensive information management, access, analysis and visualization

- Mission Support - In-house expertise who understands the requirements for costing and sizing information management systems for new or existing missions

- Operational active archive and distribution system with complete user services - 8 X5 (24 X 7 on call) staff that understands the importance of, and ensures, continuous data ingest, processing, archive and distribution
Missions Supported

Atmospheric Composition
- Total Ozone Mapping Spectrometer (TOMS)
- Upper Atmosphere Research Satellite (UARS)
- Solar Radiation and Climate Experiment (SORCE)
- Aura: Ozone Monitoring Instrument (OMI), High Resolution Dynamics Infrared Sounder (HIRDLS), Microwave Limb Sounder (MLS)
- Soon: GLORY: Aerosol Polarimetry Sensor (APS), Total Irradiance Monitor (TIM), Cloud Camera

Atmospheric Dynamics
- TIROS Operational Vertical Sounder (TOVS)
- Aqua Atmospheric Infrared Sounder (AIRS)

Modeling
- Data Assimilation Office (DAO), Global Modeling Assimilation Office (GMAO) – 30 year Merra dataset
- Land Data Assimilation System (LDAS)

Precipitation
- Tropical Rainfall Measuring Mission (TRMM)
- Hydrology Data Data Collections
Core Technology Components

User Community
Science Investigators/Partners

Data Mining
- Algorithm Build
- Data Selection
- Data Subscription
- Processing Engine
- Archive
- GES DISC

Giovanni: Data Visualization and Analysis
- Data from multiple sensors
- Single- and multi-parameter statistics
- Multiple output formats & protocols
- Data lineage support (in development)

Data & Information Web Portals
- Community and project based portals
- Tailored to the users being served

Data Processing with S4PM
- Simple, Scalable, Script-based Science Processor for Measurements

http://s4pm.sci.gsfc.nasa.gov

Data Archive and Distribution with S4PA
- Simple, Scalable, Script-based, Science Product Archive (Same kernel as S4PM)
- MERRA
- A-Train Data Depot
- GSFC Hydrology
- OMI, MLS HIROS
- TRMM
- SORCE
- Aqua AIRS

Mirador Data Search

Science Teams
Multi-Instrument Tools and Services at the GES DISC

- Mirador Data Search
- Giovanni Data Exploration
- Multi-Instrument Data Exploration
- Google Earth\textsuperscript{tm}
- Data Merging
- Applications
Mirador Data Search

http://mirador.gsfc.nasa.gov
Mirador Data Search

- Based on Google
- Fast, easy to use
- Gazetteers for places and events
- Can support portals
- Also available as Web Service
**Mirador Projects and Science Areas Tabs**

- **Designed for simple and quick data access**

### Projects

<table>
<thead>
<tr>
<th>Project</th>
<th>Description</th>
<th>Start Year</th>
<th>End Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Tran</td>
<td>Collocated with CloudSat subsets of MODIS/Aqua, AMSR-E/Aqua, OMI/Aura, and POLDER/PARASOL.</td>
<td>2004</td>
<td>2010</td>
</tr>
<tr>
<td>AIRS</td>
<td>The Atmospheric Infrared Sounder (AIRS) is a facility instrument aboard the second Earth Observing System (EOS) polar-orbiting platform, EOS Aqua. In combination with the Advanced Microwave Sounding Unit (AMSU) and the Humidity Sounder for Brazil (HSB), AIRS constitutes an innovative atmospheric sounding group of visible, infrared, and microwave sensors. Global coverage will be obtained twice daily (day and night) on a 1.330am sun synchronous orbit from a 705-km altitude. For processing convenience, the data is divided into 8-minute files for Level 1 and 2 data.</td>
<td>2002</td>
<td>2010</td>
</tr>
<tr>
<td>GLDAS</td>
<td>The Global Land Data Assimilation System (GLDAS) is generating a series of land surface state (e.g., soil moisture and surface temperature) and flux (e.g., evaporation and sensible heat flux) products simulated by four land surface models (CLM, Mosaic, Noah and VIC). Current data holdings include a set of 1.0 degree resolution data products from the four models, covering 1979 to the present; and a 0.25 degree data product from the Noah model, covering 2000 to the present.</td>
<td>1979</td>
<td>2010</td>
</tr>
<tr>
<td>GOCART</td>
<td>The Goddard Chemistry Aerosol Radiation and Transport (GOCART) model simulates major tropospheric aerosol components, including sulfate, dust, black carbon (BC), organic carbon (OC), and sea-salt aerosols. The following is a brief description of the model. The GOCART model uses the assimilated meteorological fields of the Goddard Earth Observing System Data Assimilation System (GEOS DAS), generated by the Goddard Global Modeling and Assimilation Office. The model has a horizontal resolution of 3 deg latitude by 2.5 deg longitude or 1 deg by 1 deg, and 20 vertical sigma layers (depending on the version of GEOS DAS).</td>
<td>2000</td>
<td>2007</td>
</tr>
<tr>
<td>HIRDLS</td>
<td>The High Resolution Dynamics Limb Sounder (HIRDLS) aboard the EOS Aura spacecraft (launched July 15, 2004) measures infrared emissions in 21 channels ranging from 6.12 to 17.76 microns. These measurements are used to derive vertical profiles of Ozone, HNO3, Water Vapor, Methane, N2O, NO2, N2O5, CFCl1, CFCl2, aerosols, and aerosol extinction.</td>
<td>2004</td>
<td>2008</td>
</tr>
</tbody>
</table>

### Science Areas

#### Atmospheric Composition
- Aerosols
- High Energy Particles
- Trace Gases

#### Carbon Cycle and Ecosystems
- Land Biomass
- Trace Gases

#### Climate Variability and Change
- Atmosphere
- Temperature
- Clouds
- Surface Radiation
- Trace Gases

#### Earth Surface and Interior
- Rain
- Snowfall
- Soil
- Terrestrial Water
- Heat Flux

#### Water and Energy Cycles
- Atmospheric Radiation
- Clouds
- Precipitation
- Rain
- Surface Radiation
- Land Impact
Mirador Features

- Space-time hit estimator for datasets
- Search by location name
- Search on geophysical event (e.g. Hurricane Katrina)
- Data availability calendars (Projects Tab only)
- Descriptive filenames
- Multiple batch download methods
  - Java downloader application
  - URL List
  - DownThemAll
  - FTP script
- Orbit number search (TRMM only)
Mirador Access to Services

- KML - points to WMS in many cases*
- OPeNDAP - data access / transfer framework*
- Subsetting*
- NetCDF format conversion*

*Access provided through Mirador
NetCDF Conversion
OPeNDAP Example
WMS Example
Giovanni

Goddard Interactive Online Visualization ANd aNalysis Infrastructure

http://giovanni.gsfc.nasa.gov/
About Giovanni

- Provides a simple and easy way to explore, visualize, analyze, and access vast amounts of Earth science remote sensing and model data.
- Is a Web-based application.
- Supported by NASA EOSDIS and several NASA-funded projects.
Grand view

Data Inputs

- AIRS
- MODIS
- MISR
- Parasol
- CloudSat
- CALIOP
- TOMS
- OMI
- MLS
- HIRDLS
- HALOE
- TRMM
- AMSR-E
- SeaWiFS
- Models
- and more...

Giovanni Instances

Particulate Matter (PM 2.5) from AIRNow
Aerosol from MODIS and GOCART model
Carbon Monoxide from AIRS

Water Vapor from AIRS
MODIS vs SeaWiFS Chlороphylli
Ozone Hole from OMI
Giovanni Allows Scientists to Concentrate on the Science

The Old Way:
- Pre-Science
  - Find data
  - Retrieve high volume data
  - Learn formats and develop readers
  - Extract parameters
  - Perform spatial and other subsetting
  - Identify quality and other flags and constraints
  - Perform filtering/masking
  - Develop analysis and visualization
  - Accept/discard/get more data (sat, model, ground-based)

Web-based Services:
- Jan
  - Read Data
  - Extract Parameter
  - Subset Spatially
  - Filter Quality
  - Reformat
  - Reproject
  - Visualize
  - Explore
  - Analyze

The Giovanni Way:
- Minutes
- Days for exploration
- Use the best data for the final analysis
- Derive conclusions
- Write the paper
- Submit the paper

- Giovanni and other web-based tools allow scientists to compress the time needed for pre-science preliminary tasks:
  - data discovery, access, manipulation, visualization, and basic statistical analysis analysis.

Scientists have more time to do science!
Giovanni Now

- Almost 30 customized Giovanni portals
- Thousands of geophysical parameters
- Data from:
  - ~20 space-based instruments
  - ~50 models
  - EPA and Aeronet stations
- Multiple visualization and statistical analysis functionalities including data intercomparison
- Data lineage
- Subsetted data downloads in multiple formats
Capabilities

One-parameter:
- **Area plot** – averaged or accumulated over any data period for any rectangular area (various map projections)
- **Time plot** – time series averaged over any rectangular area
- **Hovmöller plots** – longitude-time or latitude-time cross sections
- **Image animation** – for area plot
- **Vertical profiles**
- **Vertical cross-sections, zonal means**
- **Download in ASCII, netCDF, HDF, KMZ format**
- **Lineage/provenance**

Multi-parameter/Multi-sensor
- **Area plot** - geographical intercomparison between two parameters
- **Time plot** - an X-Y time series plot of several parameters
- **Scatter plot of parameters in selected area and time period**
- **Scatter plot of area averaged parameters**
- **Temporal correlation map**
- **Temporal correlation of area averaged parameters**
- **Difference plots**
- **Anomaly plots**
- **Acquiring parameter and spatial subsets** in a batch mode through Giovanni
Giovanni Features

- Uses only a Web browser
- Accesses data from multiple sensors, models and ground-based measurements from one place
- No need to learn multiple data formats
- No need to write various data readers
- No need to download large amounts of data.
- Data from multiple sources are already quality screened and harmonized
- Obtain customized data and analyses with only a few mouse clicks

Caution: Giovanni is a data exploration tool!
Giovanni is a web-based application developed by the GES DISC that provides a simple and intuitive way to visualize, analyze, and access vast amounts of Earth science remote sensing data without having to download the data.

Giovanni is comprised of a number of interfaces, called instances, each tailored to meet the needs of different Earth science research communities. To access a Giovanni instance, click on one of the four categories below.

- **Atmospheric Instances:** A-Train along CloudSat Track; Aerosol Optical Thickness Measurement and Model Comparison Daily and Monthly; Aqua/AIRS Global Daily and Monthly; Aura High Resolution Dynamics Limb Sounder (HIRDLS); Aura Microwave Limb Sounder (MLS); Aura OMI Level 3 and Level 2D; MISR Daily and Monthly; Modern Era Retrospective-Analysis for Research and Applications (MERRA) 3D Monthly and 2D Monthly; MODIS Terra and Aqua Daily and Monthly; Earth Probe and Nimbus-7 TOMS; Upper Atmosphere Research Satellite (UARS) Halogen Occultation Experiment (HALOE)

- **Environmental Instances:** Agriculture; Air Quality; Northern Eurasia Earth Science Partnership Initiative (NEESPI) Daily and Monthly

- **Ocean Instances:** Ocean Color Radiometry (SeaWiFS, MODIS, and derived and model products); Ocean Model Daily and Monthly.

- **Hydrology Instances:** Modern Era Retrospective-Analysis for Research and Applications (MERRA) 3D Monthly and 2D Monthly; MODIS Terra and Aqua Daily and Monthly; Northern Eurasia Earth Science Partnership Initiative (NEESPI) Daily and Monthly; TRMM Online Visualization and Analysis System (TOVAS); Global Land Data Assimilation System (GLDAS) Monthly.

If you already know which instance to choose, please select it from the table below.
Tracking dust storm

1. MODIS
   - High AOD

2. AIRS
   - Dry air

For area (Lat: 11.0N-20.0N, Long: 24W-14.5W)
Multi-Parameter Scatter Plots

July 2003

(For 14.0N-24.0N, Lon: 106.0E-122.0E)
(Jul2003-Jul2003)

Scatter plots also demonstrate increased chlorophyll concentration, decreased SST during winter upwelling regime.
Time-series of Accumulated Rainfall for Hurricanes Katrina and Rita
Accumulated Rainfall Tracks of Hurricanes Katrina and Rita

Katrina

Rita
Studying correlations between Chlorophyll-a and SST in the northern East China Sea using MODIS-Aqua

Temporal correlation map

Time-series
California’s Wildfires from Space Using

23-27 October 2007

Data from NASA's Aura OMI (Tropospheric NO2 and UV Aerosol Index), Aqua AIRS (Total Column CO) and Terra MODIS (Aerosol Small Fraction, Cloud Optical Thickness and Aerosol Mass Concentration Over Land)
Multi-Instrument Data Exploration
A-Train Data Depot

Comparison of A-Train Datasets Along CloudSat Orbital Track
CloudSat, and coregistered MODIS/Aqua, AIRS/Aqua, CALIPSO lidar, and OMI/Aura Atmospheric Measurements
Plots of vertical profiles of clouds, temperature, humidity, cloud and aerosol classification, horizontal swaths of cloud characteristics and total column aerosols, collocated with CloudSat track; line over-plots of cloud pressures.

Select Constraints:

Spatial

Temporal

Parameters

Select Visualization:

1. Pick Date
2. Select Location
3. Choose Parameters
4. Generate Visualization
Studying Aerosols...

Vertical Profiles (curtain plots) and Horizontal Strips ±100 km from CloudSat and CALIPSO nadir
Comparison of Aura with other A-Train Satellite Datasets
Google Earth™
Importing Giovanni Data into Google Earth™
Importing AIRS Data into Google Earth™

July 21, 2006

July 30, 2006

KMZ files for AIRS Level-3 products (2D variables only) can be downloaded from GES DISC. These example images track CO transport at the beginning and end of Siberian fires between July 21 - 30, 2006.
Data Merging
Data Fusion in Giovanni (prototype)

MODIS Terra and Aqua AOT Data Fusion

Click and drag to select area; or input latitudes (-90, 90) and longitudes (-180, 180)

Dust event, May 23, 2007
Merged multi-sensor aerosol data
March 13, 2007
Applications
Multi-Sensor Air Quality

**Data Sources**  Multi-sensor Giovanni  **End Users**

- Satellite Data
- Surface Monitor Data
- Model Data

Giovanni
Online Visualization and Analysis

- Scientists
- Air Quality Forecasters
- Policy Makers and Regulators
- Other Professionals
- Students
- Public
Giovanni: PM2.5 Gridded Data

PM$_{2.5}$ data from EPA AirNow provided in GIOVANNI as 1 deg gridded product, which makes easy to compare with other satellite & ground observation.

PM$_{2.5}$ Air Quality Index

Gridded GIOVANNI PM$_{2.5}$ (µg/m$_3$)
Prototyping PM25 data in Giovanni

PM2.5 (EPA DataFed Giovanni)

Deep Blue MODIS Aerosol Optical Depth

The standard MODIS AOT

GOCART AOT
Giovanni Air Quality Services: AOD/PM$_{2.5}$ Correlation Maps and Time Series

May 2007 AOD/PM$_{2.5}$ correlation map over the U.S

Moderate to good correlation in the eastern U.S

No significant differences were found when using the Fine Mode MODIS AOD.

May 2007 - AOD and PM$_{2.5}$ Time series over the southeast
Air Quality Data (July 7th, 2006)

- MODIS and OMI imagery show smoke aerosols over the northeast, southeast and Great Lakes.
- CALIOP Aerosol Flag (yellow) confirms that aerosols are above the boundary layer.
- EPA AirNow PM$_{2.5}$ doesn’t show anything around Great Lakes, i.e. aerosols are primarily above the boundary layer.
Thank You

http://disc.gsfc.nasa.gov/

QuikSCAT wind vector data overlaid upon TRMM precipitation data
Data Archive and Distribution with S4PA

Simple, Scalable, Script-based, Science Product Archive
- Radically simplified architecture for archive and distribution
- Features
  - Public and restricted-access
  - Subscriptions
  - Automated data integrity checking

Interfaces to:
- S4PM processing system
- Science Investigator Processing Systems (SIPS)
- EOS Data Operations (EDOS)
- Mirador search tool
- Giovanni
- Global Change Master Directory
- EOS Clearinghouse

MERRA
TRMM
SORCE
GSFC Hydrology
A-Tran Data Depot
Aqua AIRS
OMI, MLS HIRDLS
Data Processing with S4PM

Simple, Scalable, Script-based Science Processor for Measurements

- In-house developed open-source software
- Runs all data processing at the GES DISC since 2002
- Near-real-time processing: AIRS, MLS/Aura (in progress)
- Reused by LaRC for CALIPSO, FlashFlux, MISR, EDC for ASTER On-Demand

http://s4pm.sci.gsfc.nasa.gov
Data & Information Web Portals

Community and project based portals
Accessible from http://disc.gsfc.nasa.gov
Tailored to the users being served
  • Multi-mission science research
  • Discipline specific portals
  • Remote data access
Mirador Data Search

- [http://mirador.gsfc.nasa.gov](http://mirador.gsfc.nasa.gov)
- Based on Google
- Fast, easy to use
- Gazetteers for places and events
- Can support portals
- Also available as Web Service
Giovanni: Data Visualization and Analysis

- Data from multiple sensors
- Single- and multi-parameter statistics
- Multiple output formats and protocols
  - JPEG and PNG
  - WMS
  - HDF
  - netCDF
  - ASCII
  - KML for Google Earth
- Multiple input formats & protocols
  - HDF4 and HDF5
  - OPeNDAP
  - WCS
  - Web Services (including SciFlo)
  - Standard FTP
  - Data lineage support (in development)

http://giovanni.gsfc.nasa.gov
Data Mining

Data mining services available in S4PM
- Users submit and execute data mining algorithms
- Simple Web interface
- Subscriptions process new data as they arrive
- Mining results are made available to user via FTP

Coming Soon: Mining Web Services
- ADaM mining algorithms from Univ. Alabama-Huntsville
- Invoked via Web Services interface
Some Handy Acronyms

- DAAC – Distributed Active Archive Center (the collective name for all NASA Earth science data centers)
- GES DISC – GSFC Earth Sciences Data and Information Services Center (the name of the DAAC located at NASA’s Goddard Space Flight Center)
- EOS – Earth Observing System
- EOSDIS – EOS Data and Information System
- ESDIS – Earth Science Data and Information Systems Project
- GSFC – Goddard Space Flight Center
# EOSDIS Evolution at the GES DISC: 2006-2007

<table>
<thead>
<tr>
<th>Pre-Evolution System Characteristics</th>
<th>Evolved System Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized interface(s)</td>
<td>Discipline-specific interfaces in addition to generalized interface(s)</td>
</tr>
<tr>
<td>Tape archive</td>
<td>Disk archive</td>
</tr>
<tr>
<td>- All products archived</td>
<td>- Some products processed on demand (virtual products)</td>
</tr>
<tr>
<td>- Order data for delivery</td>
<td>- Download data automatically upon choosing</td>
</tr>
<tr>
<td>Search and order tools</td>
<td>Tools to find, explore, and analyze data</td>
</tr>
<tr>
<td>Distribute standard products</td>
<td>Distribute lower volume tailored products</td>
</tr>
<tr>
<td>System changes require long lead time</td>
<td>System changes implemented quickly according to priority</td>
</tr>
<tr>
<td>Steward data</td>
<td>Steward data</td>
</tr>
</tbody>
</table>
# EOSDIS Evolution at the GES DAAC: 2009 – 2013

<table>
<thead>
<tr>
<th>Evolved System Characteristics</th>
<th>Future System Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discipline-specific interfaces in addition to generalized interface(s)</td>
<td>Discipline-specific, multi-mission, services oriented interfaces in addition to generalized interface(s)</td>
</tr>
</tbody>
</table>
| Disk archive  
  - Some products processed on demand (virtual products)  
  - Download data automatically | Disk archive  
  - Access provided to all products archived locally and remotely  
  - Download data automatically upon choosing |
| Tools to find, explore, and analyze data                 | Web-based tools to provide for comprehensive discovery, access, visualization and analysis of coherently related Earth science datasets (satellite, ground-based and model output) to enable interdisciplinary Earth science research |
| Distribute lower volume tailored products                | - Distribute lower volume tailored products  
  - Broker hard to get data |
| System changes implemented quickly according to priority | System changes implemented quickly according to priority within given budget cap |
| Steward data                                            | Steward data                                                      |