Continuously Enhancing Usability of Remote Sensing Data

GO-ESSP 2011

Asheville, NC

Jianfu Pan *, jianfu.pan@nasa.gov
Christopher Lynnes **
Gregory Leptoukh **

Affiliations:
* Adnet Systems, Inc.
** Goddard Earth Sciences Data and Information Center, NASA

User, Data and Technology

User

- Researchers
- Educators/students
- Citizen scientists

Data

- Large amounts
- Various formats
- Stored at different locations
- Swath (L2) and geospatial (L3)
- Quality varies

Technology

- Web services
- OGC standards (data/map access)
- OPeNDAP
- Giovanni

More about data formats

- NetCDF – most application friendly
- HDF – most commonly used so far
- Customer binary
- ASCII

Access data from live feed

- OPeNDAP
- Grads GDS
- WCS

Some issues using HDF

- Fill values
  - Fill value too large and causing rounding problem
  - Undefined fill values
  - Multiple fill values
- Native API too complex
- Latitude/longitude dimensions not easy to identify
- Not always package friendly
  - Lack of naming convention within HDF files

Typical data preparation tasks

- Subsetting
  - Variable
  - Spatial
  - Temporal

- Regridding
  - Resolutions differ
  - Grids not aligned
  - Irregular grids

- Projection/interpolation
  - δ-levels to pressure
  - Between different pressure levels
  - Swath/grids mapping

- Quality filtering

- Format conversion

On-the-Fly (OTF) Web Services

- Execution on acquisition
  - Server-side data preparation
  - According to specific needs of a request

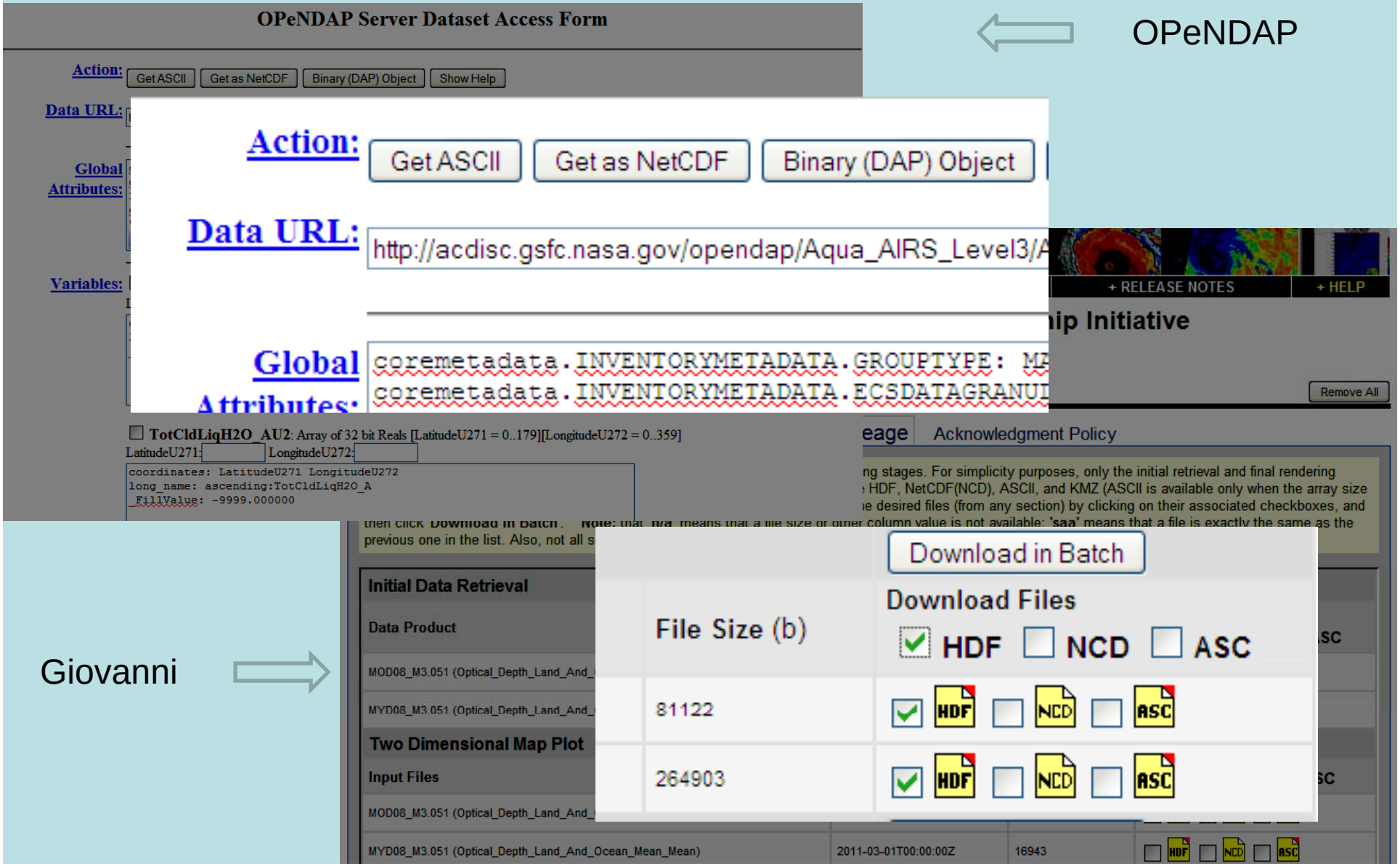
- REST-like: acquire as URLs

- Accommodates any executable that …
  - Takes one file as input
  - Produces one file as output

- Mainly used for parameter subsetting and netCDF conversion

OTF netCDF conversion examples

OPeNDAP

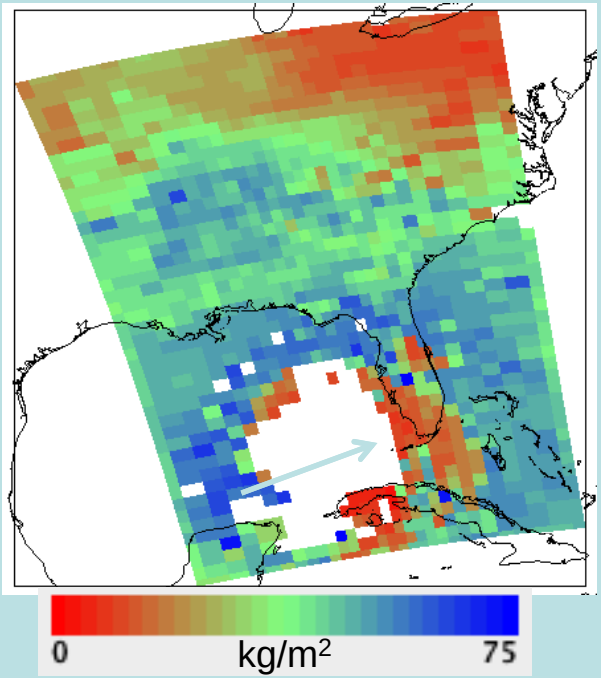


Giovanni

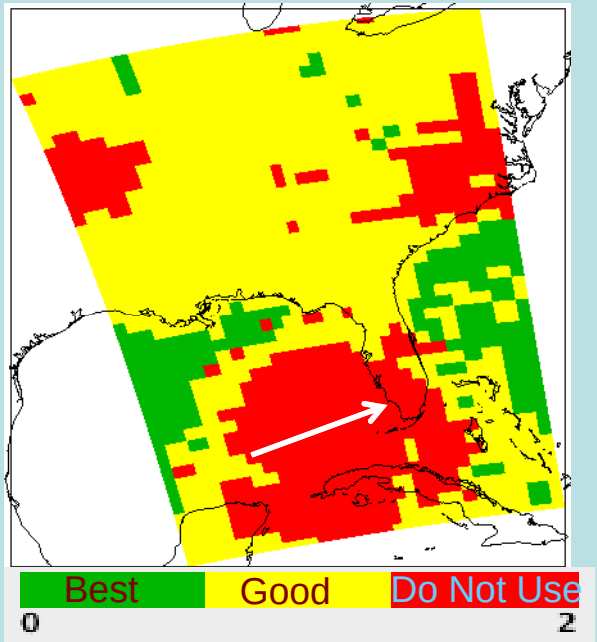
Data Quality Screening Service

- Level 2 Satellite data often comes with quality control flags
- Until now, each user typically had to write his/her own software to filter bad quality data—or ignore them

AIRS (Atmospheric Infrared Sounder)
Total Column Precipitable Water



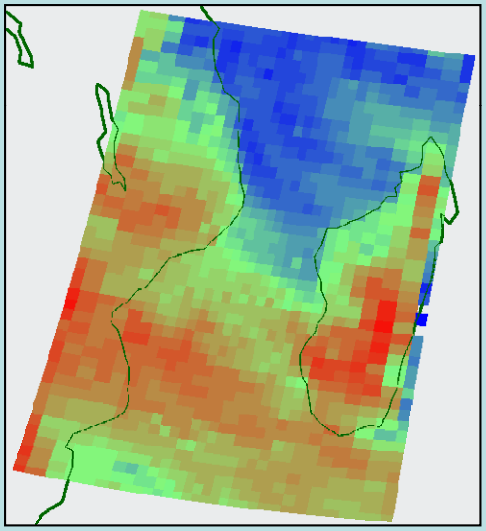
Quality Flag



Hurricane Ike, 9/10/2008

The Data Quality Screening Service for AIRS Level 2 swath data

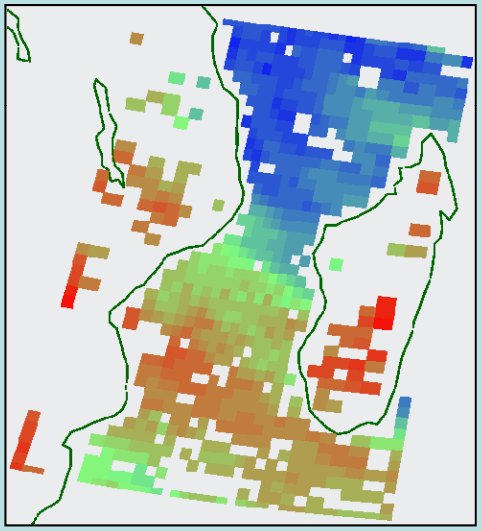
Original data array



Total column precipitable water

Mask based on user criteria Quality flag<2

Good quality data pixels retained



Output file has the same format and structure as the input file, with fill values replacing the low-quality data

OPeNDAP

- A generic and standard way of data access
- Allows simple browsing and access to remote data
- Subsetting capability
- Native file format hidden from users

OPeNDAP at GES DISC

- Hyrax server

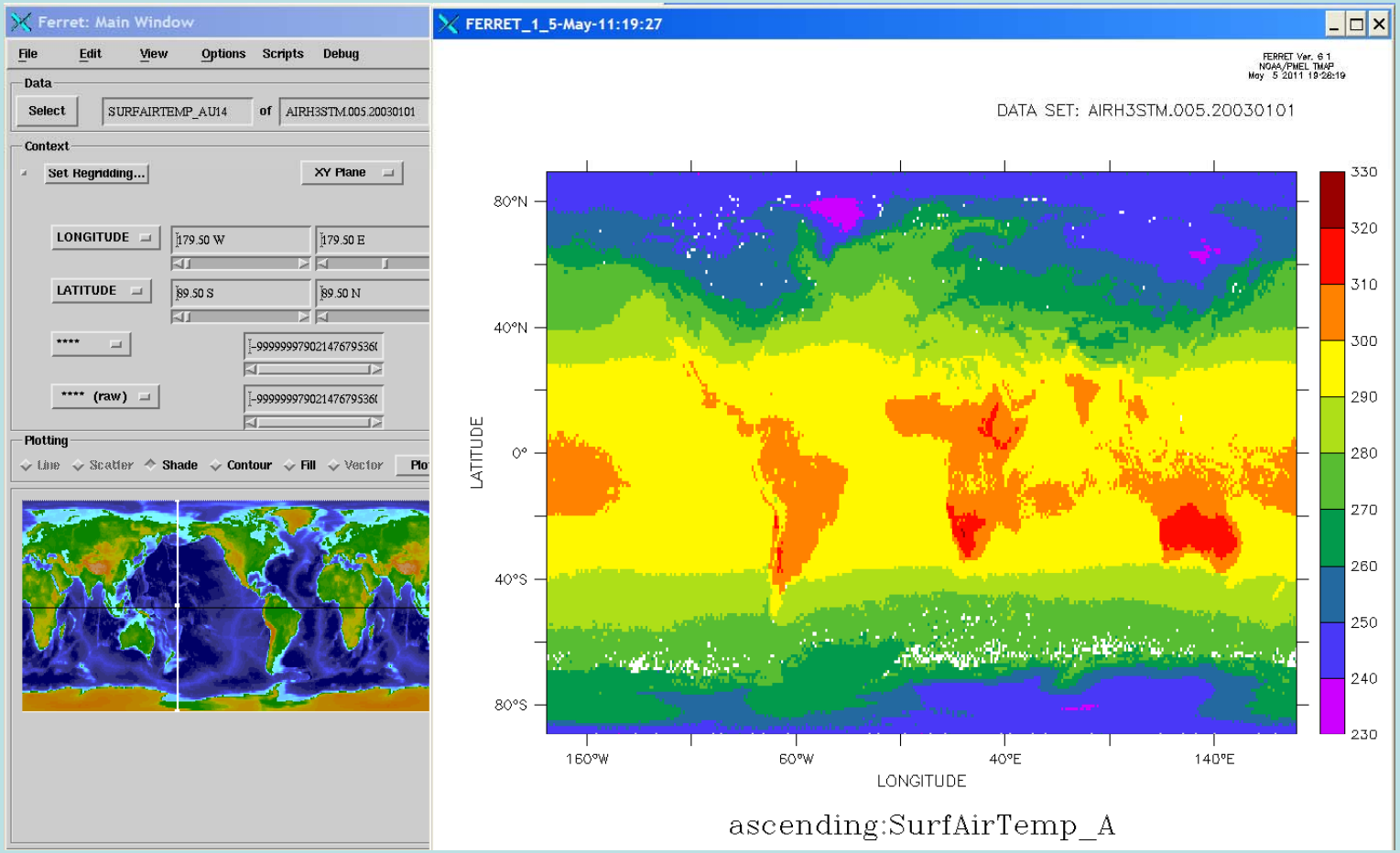
- Giovanni receives OPeNDAP data from NASA Langley
http://l0dup05.larc.nasa.gov/opendap/
  - MISR (Multi-angle Imaging SpectroRadiometer)
  - TES (Tropospheric Emission Spectrometer)
  - CERES (Clouds and the Earth's Radiant Energy System)

- Hosting these OPeNDAP data
  - http://acdisc.gsfc.nasa.gov/opendap/
  - http://agdisc.gsfc.nasa.gov/opendap/
  - http://disc2.nascom.nasa.gov/opendap/
  - …

- Grads GDS support

A Ferret example

- AIRS monthly data AIRH3STM.005 on 1/1/2003
- Data converted to netCDF through OPeNDAP server



Other technologies being integrated

- ❖ IDV
- ❖ Panoply
- ❖ Pomegranate

Particularly interested in …

- capability of producing professional results
- netCDF friendly
- free

Giovanni

❖ An infrastructure integrating data preparation, analysis, and visualization

❖ Accomplished through services and workflows

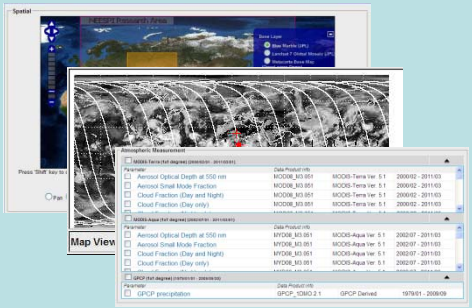
❖ Ultimate simplicity for user

❖Handles everything for user

- Data fetching
- Preparation and quality filtering
- Analysis
- Visualization

❖ Accessible through web as well as WMS/WCS-like URL

Giovanni in a Nutshell



Web access
(portals)

http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/giovanni-wmz.cgi?SERVICE=WMS&WMTVER=1.0.0&REQUEST=GetMap&SRS=EPSG:4326&EXCEPTIONS=INIMAGE&LAYERS=MOD08_M3.005::Cloud_Top_Pressure_Mean_Mean:max=800:min=300&STYLES=latlonplot:top=surface:bottom=surface:ctype=custom:palette=Rainbow&BBOX=109.3359375,20.390625,124.453125,30.9375&TIME=2003-02-01T00:00:00Z/2003-02-28T00:00:00Z&FORMAT=PNG&transparent=true&WIDTH=800&HEIGHT=600

URL access

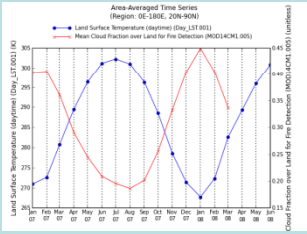
External
Data
Source

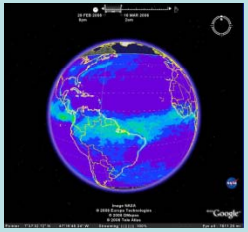
Data description

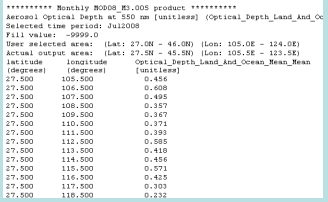
File inventory

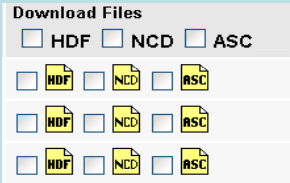
GES
DISC
Archive

Giovanni Workflow
- Fetching
- Subsetting
- Regridding
- …

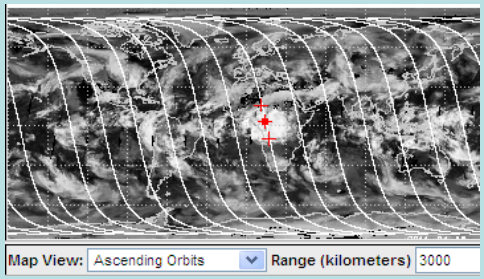


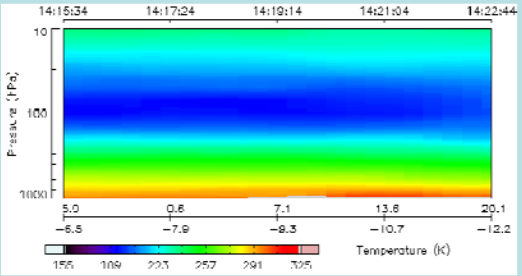


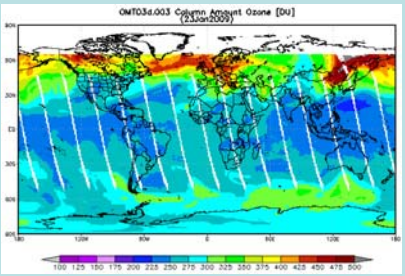


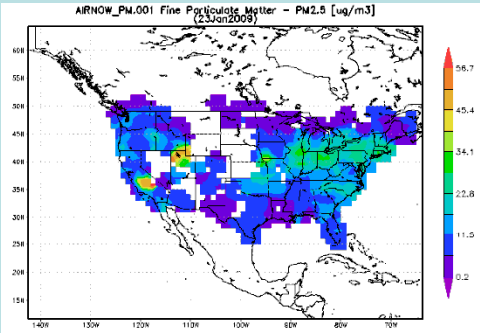


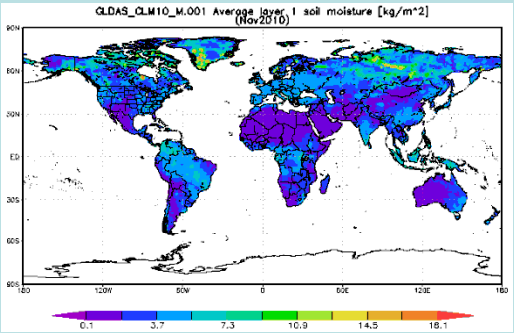
Giovanni Screen Shots

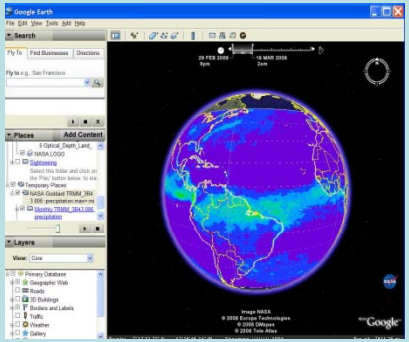


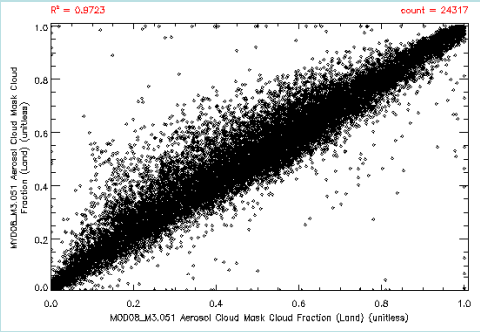


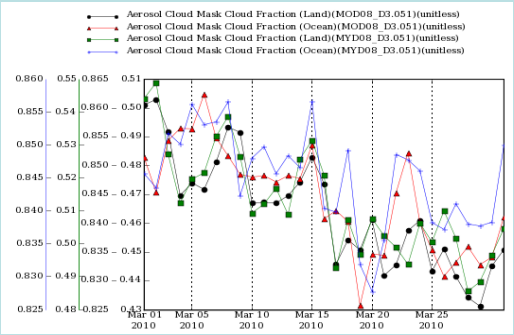


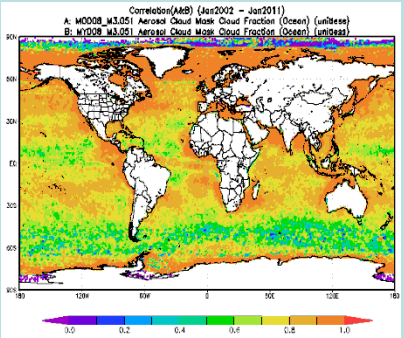












Future Giovanni

- More robust through re-engineering
- Truly service-oriented
- Community-based
- Inter-operable with other services such as data download
- Customized climatology
- Provenance and advisory aspect

Conclusion

Data usability improvement is a continuous process, and is done through…

- Better technology/infrastructure
- Better standardization
- Better data quality control
- Knowledge base and advisory