

Science Data Support for the Research and Applications Communities – a Behind-the-scenes Look at the GES DAAC's Support of TRMM

William Teng
NASA GES DAAC (SSAI)

Outline

- Anatomy of TRMM data support.
- TRMM data mining.
- OGC-compliant web system.

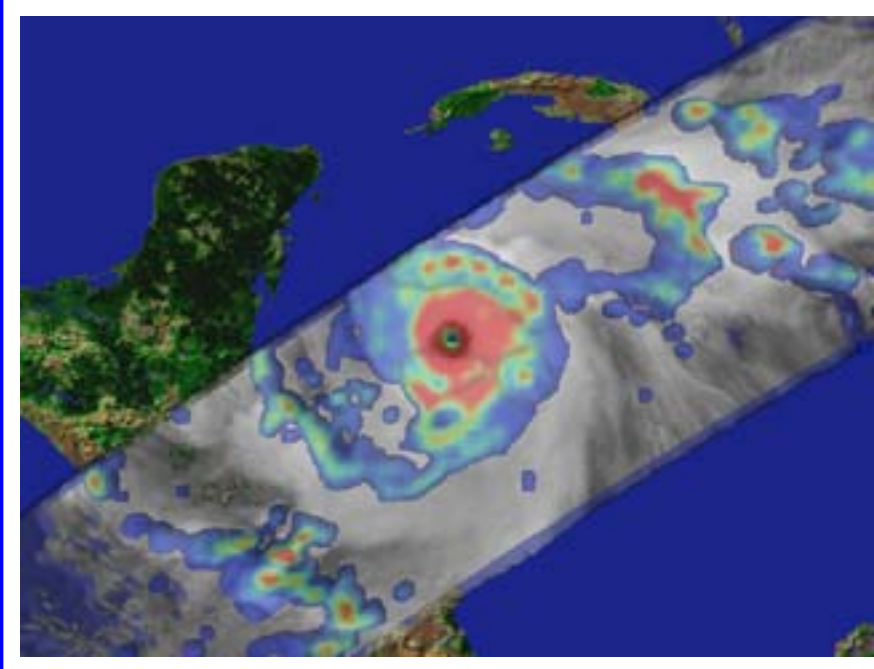
What is the GES DISC DAAC?

- Created as part of NASA's Earth Observing System (EOS) to make data and information about the Earth's environment and climate available to research, education, applications, and information communities.
- Data products are available free through the Web, as direct downloads or orders shipped via media or ftp pickup.
 - <http://daac.gsfc.nasa.gov>

What is TRMM?

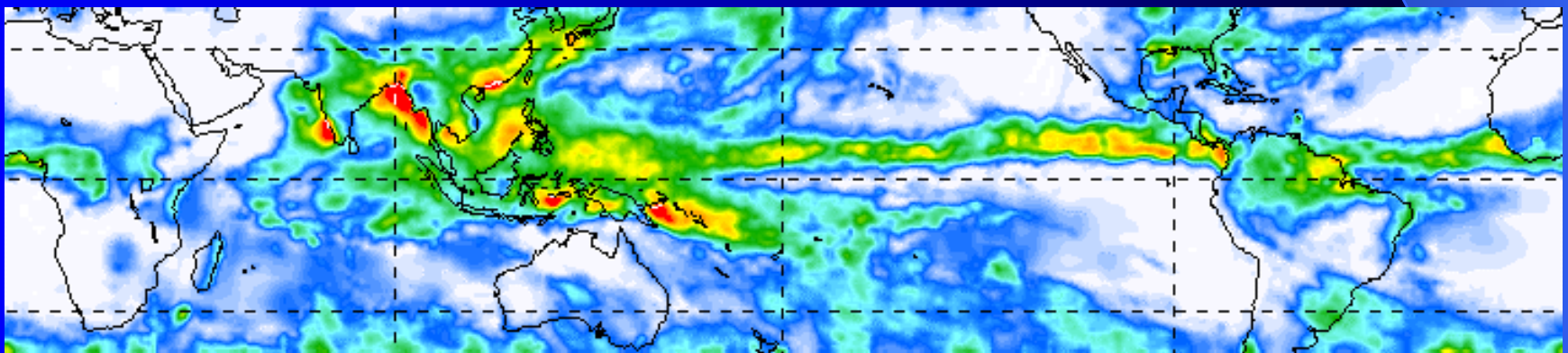
- *Tropical Rainfall Measuring Mission* - a joint effort between NASA and National Space Development Agency (NASDA) of Japan to monitor and study tropical and subtropical rain systems.
- Satellite visible, infrared, and microwave observations; ground radar and rain gauge measurements.
- Launched Nov. 1997.
- Orbit boost Aug. 2001.

What is TRMM cont.



Hurricane analysis
(Mitch, Oct. 1998)

Monthly Rainfall Products



TRMM Merged Precip Jun 2001

(mm/d)

0

4

8

12

16

20+

Anatomy of TRMM Data Support

Poster presented at AMS'02, Orlando

TRMM Data Mining – Why?

- As data volumes increase, the proportion of data that can be distributed to users decrease.
- Users express concerns about ability to manage large amounts of data.
- Many users apply “data mining” techniques to large volumes of data (up to ~1 TB) received from GES DAAC.
- Rapid advances in computer power enabling increases in data processing that are outpacing tape drive performance and network capacity.

TRMM Data Mining – Mitigation

- Migrate more data mining and mining preparation activities into data centers to reduce distributed data volume.
- Offer users more useful, manageable, and information-rich products.

Data Mining Campaign Scenario

- User contacts DAAC, sends code with description
- DAAC analyzes code for compliance
- Code unit tested for benchmarks
- Code integrated
- 1-day, 30-day runs; results checked by user
- Operational run
- Mined products distributed normally; not available to public

TRMM Data Mining Types

- Geo-regional subset
- Parameter subset
- Parameter threshold subset
- Resampling
- New product

Summary of TRMM Data Mining Algorithms

Data Mining ID	Description	Data Mining Type	Prog. Lang.	Input Product ID	Input File Size (MB)	Output File Size (MB)	Size Re-duction Ratio	Total Mined Files
BURN	Burn scar used for diurnal cycle correction	5	C	1B01	86	0.13	661	23780
FIRE	Fire pixel detection from VIRS data	5	C	1B01	86	6.25	13	23823
GAGE	PR surface rain extracted for area surrounding a rain gauge	1, 2	IDL	2A25	13	0.38	34	18071
GBM1B11	9 regional subsets for specific TMI parameters	1, 2	IDL	1B11	13	1.73	7	24175
GBM2A23	9 regional subsets for specific PR parameters	1, 2	IDL	2A23	6	0.54	11	23951
GBM2A25	9 regional subsets for specific PR parameters	1, 3	IDL	2A25	18	0.68	26	23952
MANOS_1B11	8 regional subsets for specific TMI parameters	1, 2	IDL	1B11	13	5.66	2	18150
MANOS_1C21	8 regional subsets for specific PR parameters		IDL	1C21	38	7.36	5	17897
MANOS_2A23	8 regional subsets for specific PR parameters	1, 2	IDL	2A23	6	1.90	3	17897
MANOS_2A25	8 regional subsets for specific PR parameters	1, 3	IDL	2A25	18	2.46	7	18055
MERGJIM	Regional subset	1	IDL	MERG	20	0.87	22	4512
MERGUF	Regional subset and thresholding to reduce size	3	IDL	MERG	20	8.00	2	4512
PRDUMP	Parameter selection	2	C	2A25	18	2.74	6	23952
SFCRAIN	PR surface rain spatial average covering TMI footprint	1, 3, 4	IDL	2A25	18	0.19	94	23952
SSMIRAIN	PR surface rain spatial average covering SSMI footprint	1, 3, 4	IDL	2A25	18	0.12	94	23952

OGC-Compliant Web System

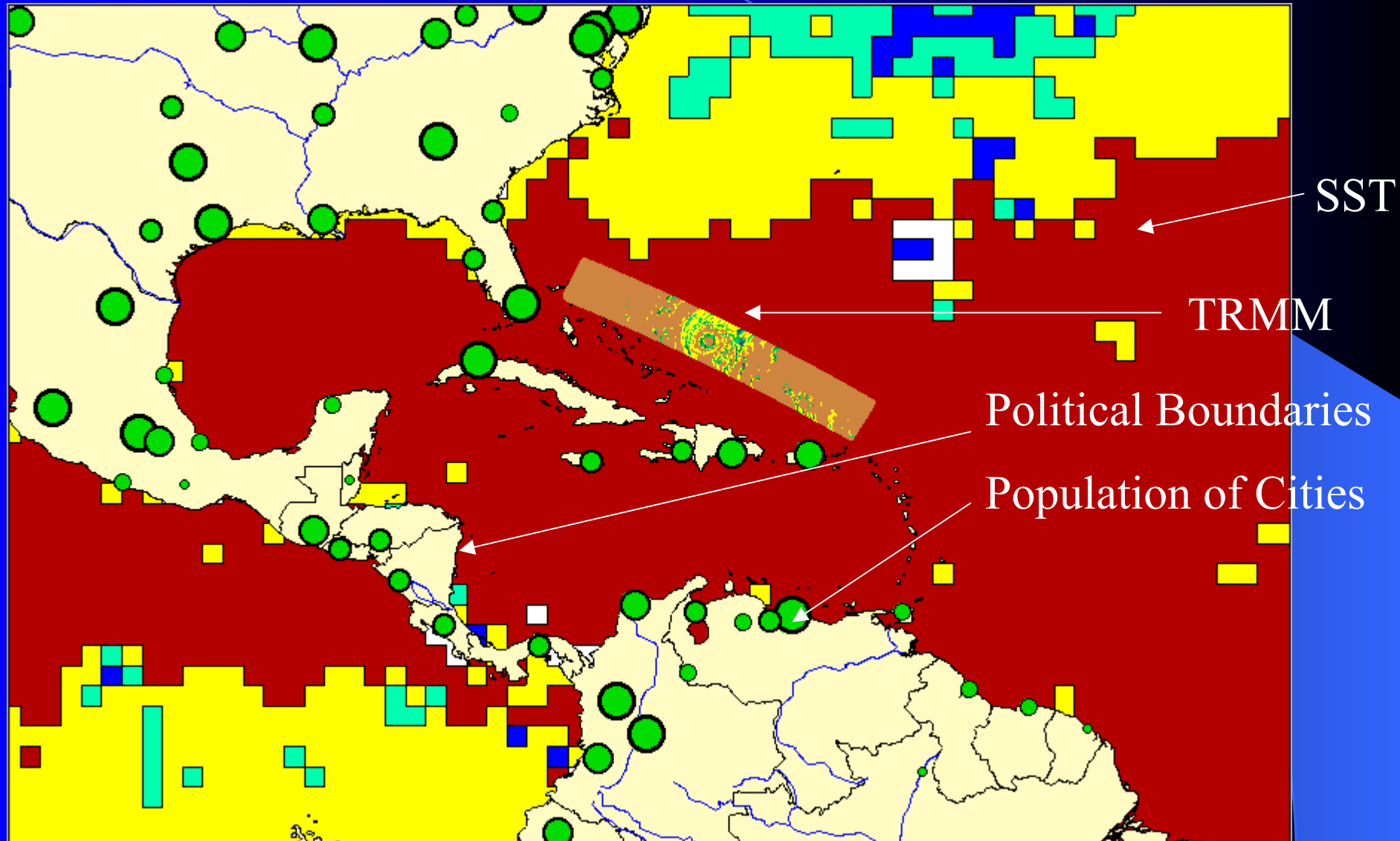
March 23, 2002

COAA Workshop, College Park

Positive/Negative of data storage/distribution in HDF

- Good for data storage
- Combines data and metadata in one file
- Established user community
- Not supported by many commercial packages
- Steep learning curve
- Other types of data are not stored in HDF
- Conversion to other formats requires programming skills

Benefit of TRMM Data in a GIS



OpenGIS as a potential solution

- Open GIS Consortium (OGC)
 - Founded in 1994
 - Partnership of Industry, Government, Educational institutions
 - <http://www.opengis.org>
- Goal is to create interoperability specifications to allow the transparent exchange of Geospatial data regardless of data format.

<http://daac.gsfc.nasa.gov/bin/wmtver=1>

REQUEST=
&LAYER=
&ST=
&h=
&width=
&tran=
except=
&time=

Java Applet Window

File Global Test layers Ancillary data from Globe/DE

Data Provider ▶
Data Set ▶
Ancillary ▶

UpperLeft=-180.0,90.0 LowerRight=180.0,-90.0

Close Window

</Layer>
- <Layer queryable="0" opaque="0" noSubsets="0">

My Computer

Netscape: NASA/GSFC DAAC-RSIP Web GIS Prototype

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: <http://daacdev2.gsfc.nasa.gov/WEBGIS/rsip.html> What's Related

NASA/GSFC DAAC-RSIP Web GIS Prototype

File

101 00011 00101 10111

Surface Rain Rate (mm/hr)

- +7.5
- 6.0-7.5
- 4.5-5.5
- 3.0-4.4
- 0.01-0.
- 0.00

NDVI

- >0.6
- 0.45-0.
- 0.3-0.4
- 0.15-0.
- < 0.15

NASA

Create maps from satellite data stored at multiple nodes in the RSIP network.

Search for data either by desired RSIP or data type

Ancillary layers can be added to the images you are interested in for reference

Zoom in on the area you are interested in: use your mouse to draw a box around a region, then use the magnify tool to zoom in on that part of the image.

View the actual data values: use the Get Coverage button to view the actual data values for a region or entire image.

Turn layers on/off: using the checkbox next to the name of the layer in the scrollable panel at the bottom.

Restart the process, select the clear tool to wipe away all layers

National

Mapping Data

Conclusions

- GES DAAC's TRMM support has enabled users to benefit from the use of TRMM data in solving problems.
- TRMM data mining can efficiently mitigate problems of storage for users, limitation of network distribution capacity, and cost of tape handling.

Conclusions (cont.)

- GIS and OGC-compliant initiatives can provide TRMM users easier access to analysis tools and emerging technologies that show great potential for enhancing TRMM data analysis.
- Will allow a more diverse user community to access TRMM data.
- Increase the diversity of potential applications.

Contact Information

- GES DISC DAAC
 - <http://daac.gsfc.nasa.gov>
 - E-mail: daacuso@daac.gsfc.nasa.gov
- Hydrology Data Support Team
 - http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/hydrology/hd_main.shtml
 - E-mail: hydrology@daac.gsfc.nasa.gov
- GIS-related efforts
 - <http://daac.gsfc.nasa.gov/WEBGIS>