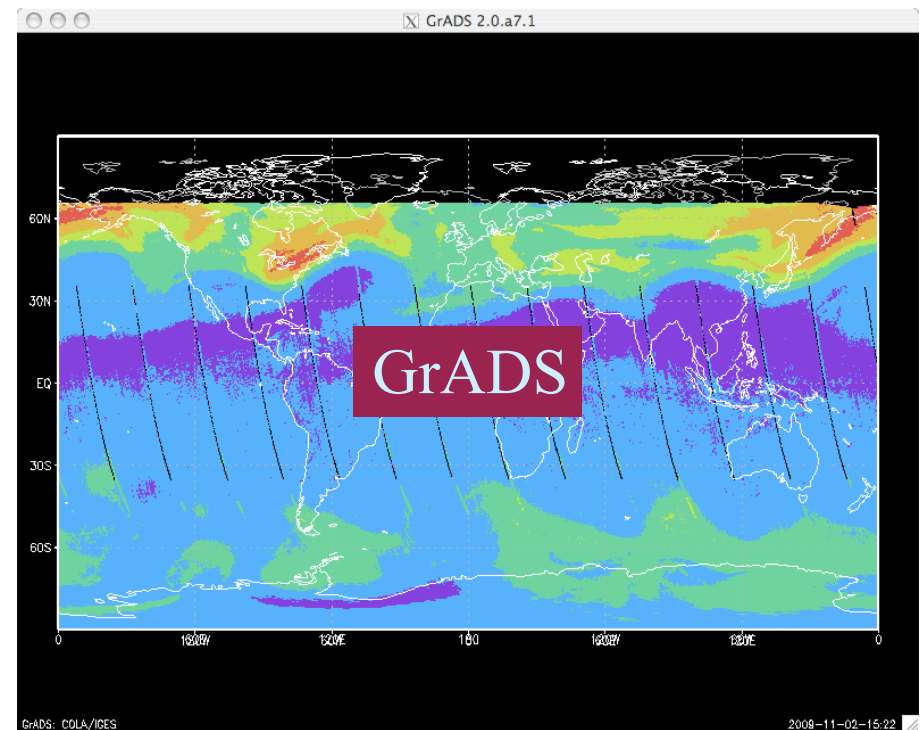
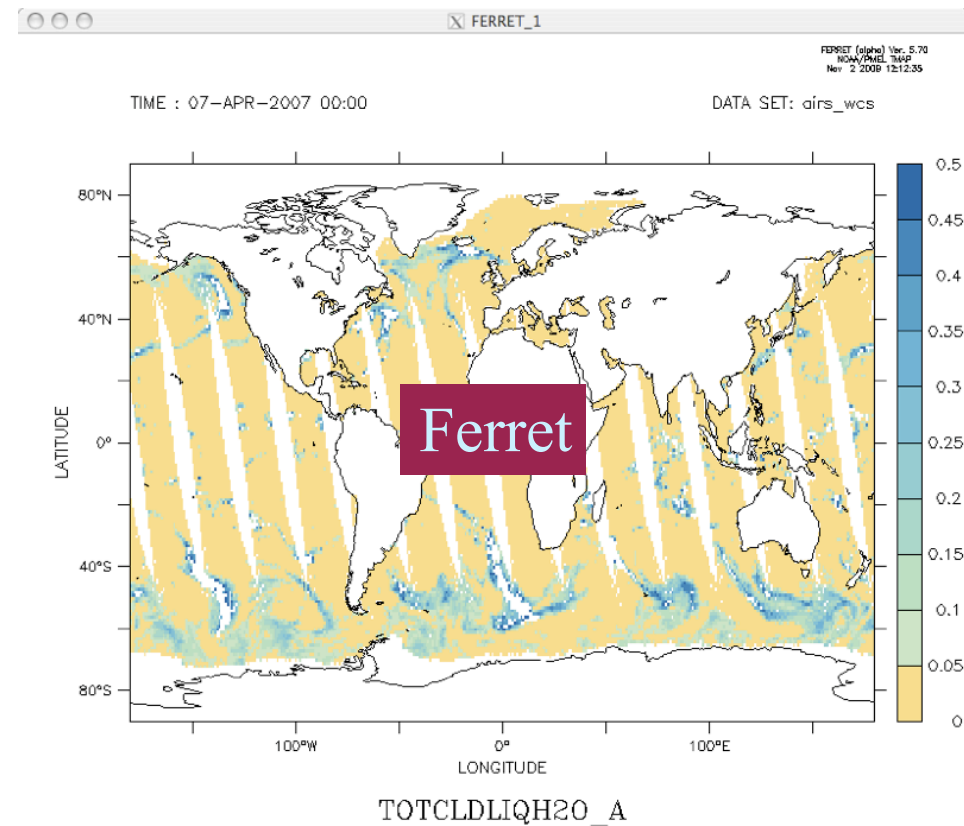
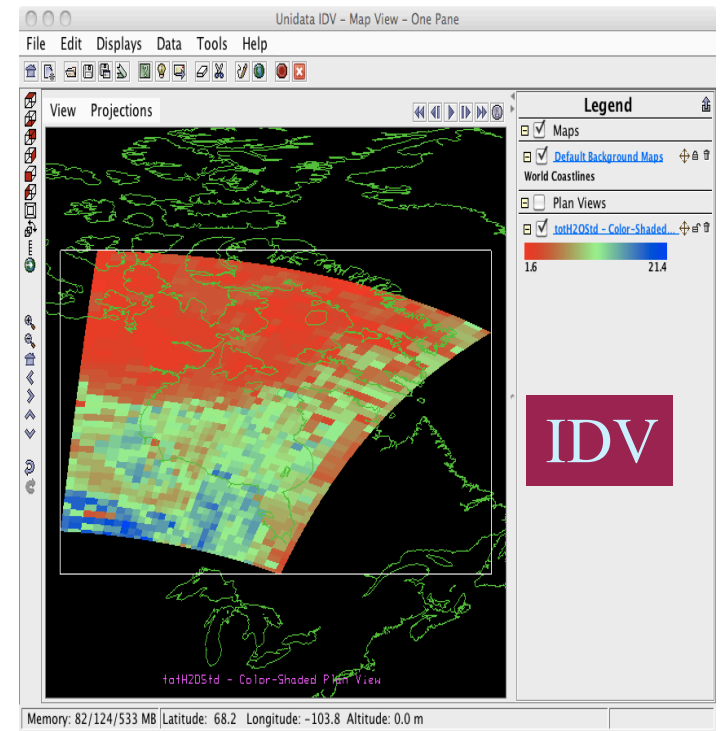
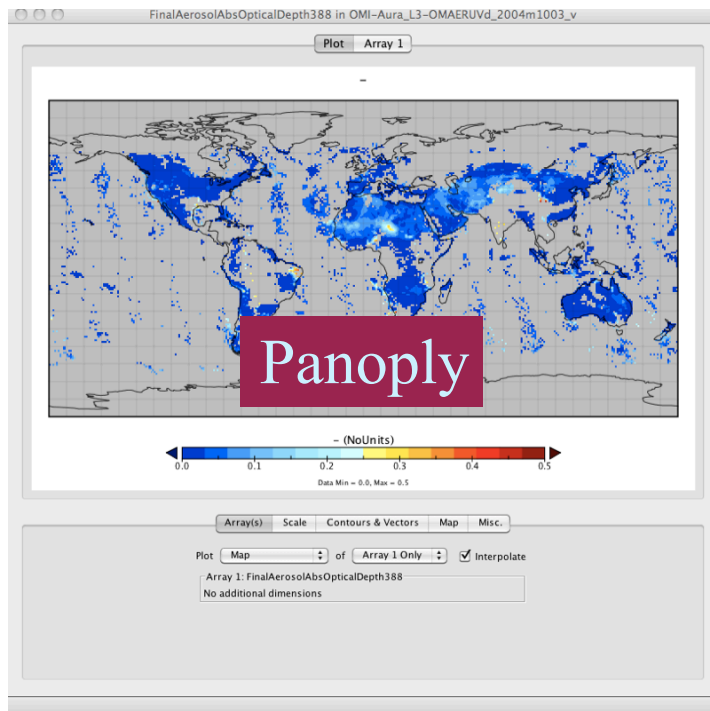




PRACTICAL METHODS FOR MAKING HDF DATA USABLE IN NETCDF-BASED TOOLS

Chris Lynnes and Fan Fang

Goddard Earth Sciences
Data and Information Services Center



5 Nov 2009

Thursday, November 5, 2009

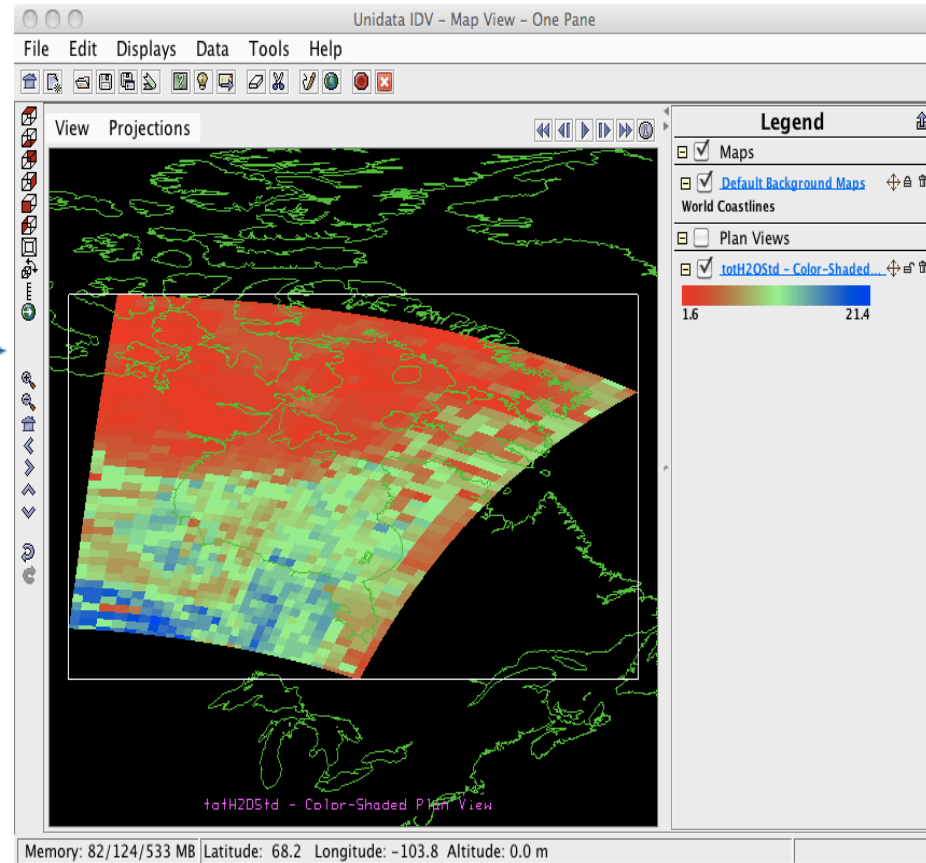


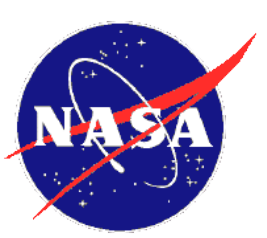
HDF
Data
:
0000
1100
0000
0011
0001
0011
0000
0001
0110
1101
1101
1100
0101
1110
0000
0101
0110

"And then a miracle occurs..."



Unidata IDV





IDV with HDF-EOS5 “self-describing format”



IDV with HDF-EOS5 “self-describing format”

The screenshot shows the IDV software interface. The main window is titled "Dashboard" and has a menu bar with "File", "Edit", "Displays", "Data", "Tools", and "Help". Below the menu bar is a toolbar with various icons. The main area has tabs for "Quicklinks", "Data Choosers", "Field Selector", and "Displays". The "Data Choosers" tab is active, showing a "Data Source Type" dropdown set to "Grid files (netCDF/GRIB/OPeNDAP/GEMPAK)". Below this is a folder selection dropdown set to "HDF_Workshop". A table lists files with columns "Name" and "Date Modified". One file is selected: "OMI-Aura_L3-OMAERUVd_2009m1028_v003-2009..." with a date of "Sunday, November 1, 2009 8:20 PM".

An error dialog box titled "No Gridded Data" is overlaid on the main window. It contains the text: "No gridded data found for:" followed by the file path: `/Users/clynnnes/work/HDF_Workshop/OMI-Aura_L3-OMAERUVd_2009m1028_v003-2009m1030t014455.he5`. Below this, it asks: "Do you want to try to load this as another data type?". There are "OK" and "Cancel" buttons. At the bottom of the dialog, there is a blue link: "Press 'Add Source' to load the selected file" and an "Add Source" button with a refresh and help icon.

At the bottom of the main window, a status bar shows: "Memory: 64/98/533 MB | Loading in data source: Grid files (netCDF/GRIB/OPeNDAP/GEMPAK)" and a progress indicator.



Panoply w/HDF5

Datasets Browser

Create Plot Target

Remove Remove All Hide CDL

Datasets & Variables

Name	Long Name	Type
OMI-Aura_L3-OMAERUVd_2...	OMI-Aura_L3-OMAERUVd_2...	Local File

List: Only Plottable Variables

Dataset/Variable CDL Info

```
:_FillValue = -1.2676506E30f; // float
:Units = "NoUnits";
>Title = "Final Aerosol Extinction Optic
:UniqueFieldDefinition = "OMI-Specific";
:ScaleFactor = 1.0; // double
:Offset = 0.0; // double
:MissingValue = -1.2676506E30f; // float
:_lastModified = "2009-10-30T05:44:55Z";
float FinalAerosolExtOpticalDepth500(YDim=
:_FillValue = -1.2676506E30f; // float
:Units = "NoUnits";
>Title = "Final Aerosol Extinction Optic
:UniqueFieldDefinition = "OMI-Specific";
:ScaleFactor = 1.0; // double
:Offset = 0.0; // double
:MissingValue = -1.2676506E30f; // float
:_lastModified = "2009-10-30T05:44:56Z";
float FinalAerosolSingleScattAlb388(YDim=1
:_FillValue = -1.2676506E30f; // float
:Units = "NoUnits";
>Title = "Final Aerosol Single Scatterin
:UniqueFieldDefinition = "OMI-Specific";
:ScaleFactor = 1.0; // double
:Offset = 0.0; // double
:MissingValue = -1.2676506E30f; // float
:_lastModified = "2009-10-30T05:44:56Z";
```



Ferret and GrADS with HDF-EOS5

```
yes? set data OMI-Aura_L3-OMAERUVd_2004m1001_v003-2009m0714t093948.he5

**TMAP ERR: Error in namelist record
           Error in: FORMAT_RECORD; or file not descriptor file
           Data set: ./OMI-Aura_L3-OMAERUVd_2004m1001_v003-2009m0714
t093948.he5
yes? quit
```

```
ga-> sdfopen OMI-Aura_L3-OMAERUVd_2004m1001_v003-2009m0714t093948.he5
Scanning self-describing file:  OMI-Aura_L3-OMAERUVd_2004m1001_v003-20
09m0714t093948.he5
gaopnc error: nc_open failed to open file OMI-Aura_L3-OMAERUVd_2004m10
01_v003-2009m0714t093948.he5
read_metadata: gaopnc failed
gadsdf: Couldn't ingest SDF metadata.
ga-> █
```

*Can be opened if user defines dimensions and uses **xdfopen**



Solution #1: Convert to netCDF

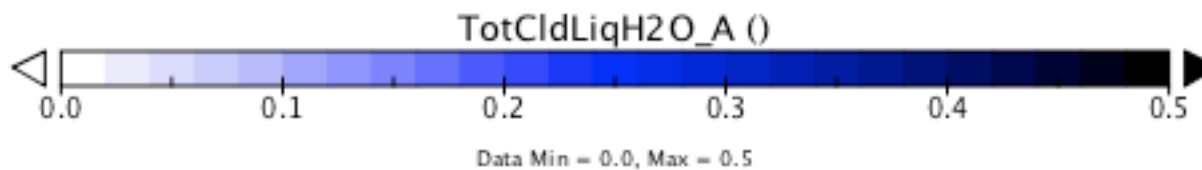
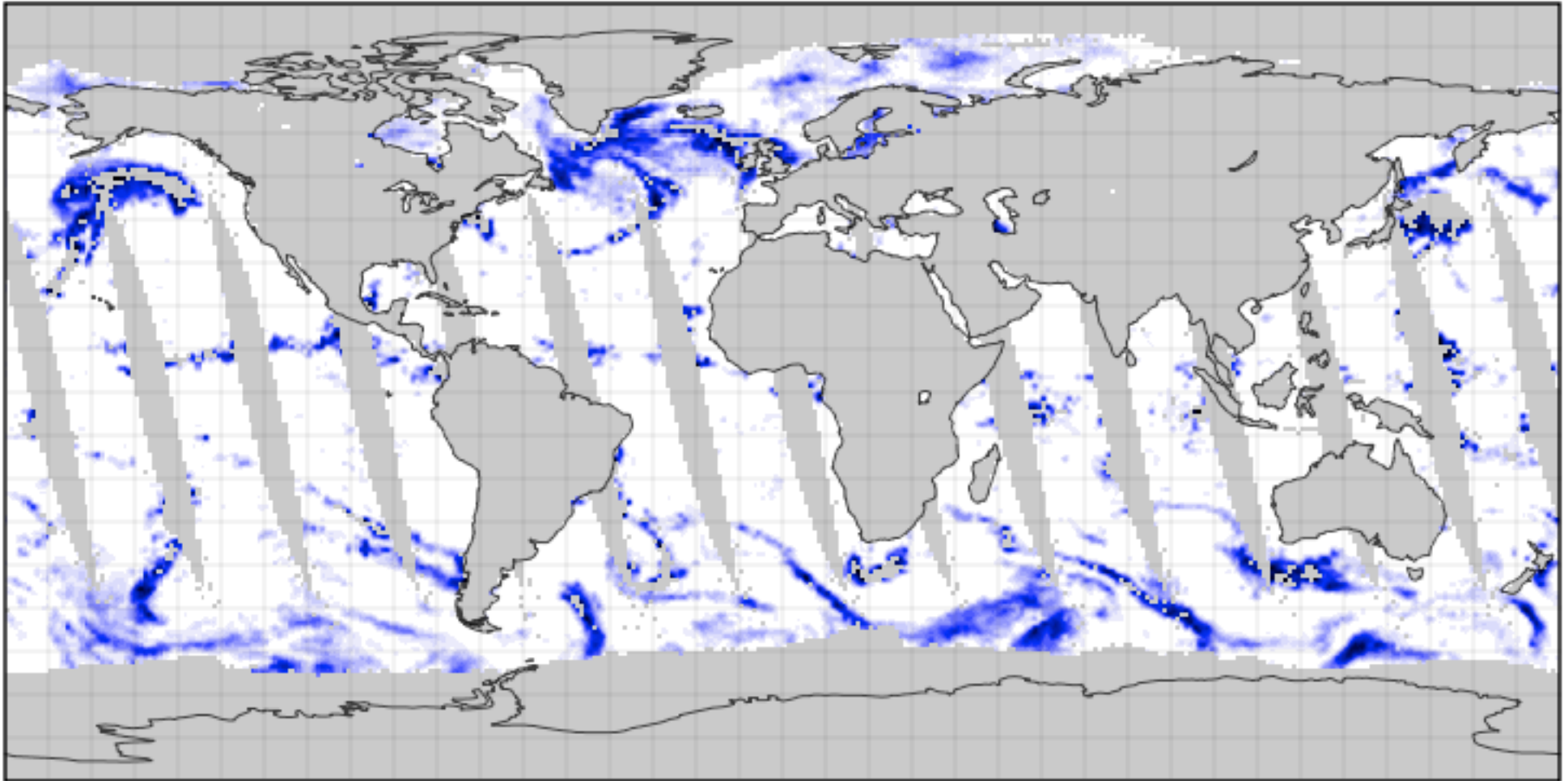
- On-the-fly conversion to netCDF-3
- Custom C++ code
- Acquiring the URL executes the conversion and downloads the file

<input checked="" type="checkbox"/>	AIRS.2009.10.27.L3.RetStd001.v5.2.2.0.G09302133019.hdf (70.24 MB) Download: HDF (FTP) HDF (HTTP) NetCDF KMZ OPeNDAP
<input checked="" type="checkbox"/>	AIRS.2009.10.26.L3.RetStd001.v5.2.2.0.G09301132521.hdf (70.64 MB) Download: HDF (FTP) HDF (HTTP) NetCDF KMZ OPeNDAP



Panoply w/converted AIRS L3

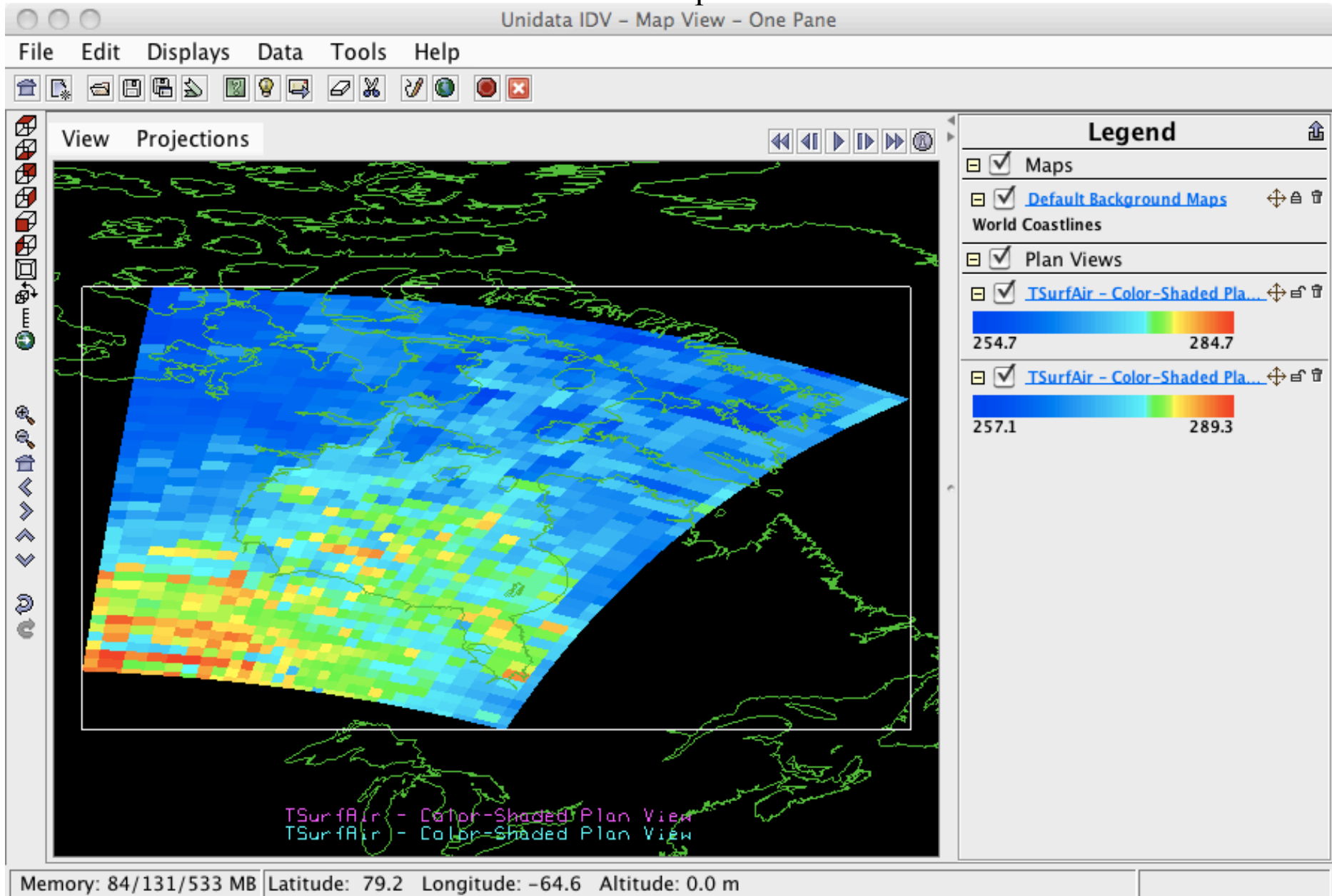
TotCldLiqH2O_A





IDV w/Converted AIRS L2

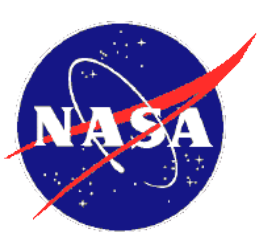
Surface Air Temperature





But...

- Custom C++ code is expensive to maintain
- Z dimension scale values are hard to find in AIRS data
 - vdatas w/non-standard names
 - external documents
 - therefore: hard-coded in C++ class
 - *if only we could attach standard Z dimscales to the SDS*
- New class for each dataset



Solution #2: Lats4D Conversion

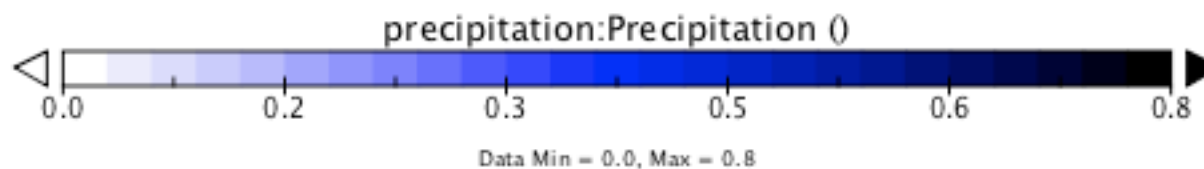
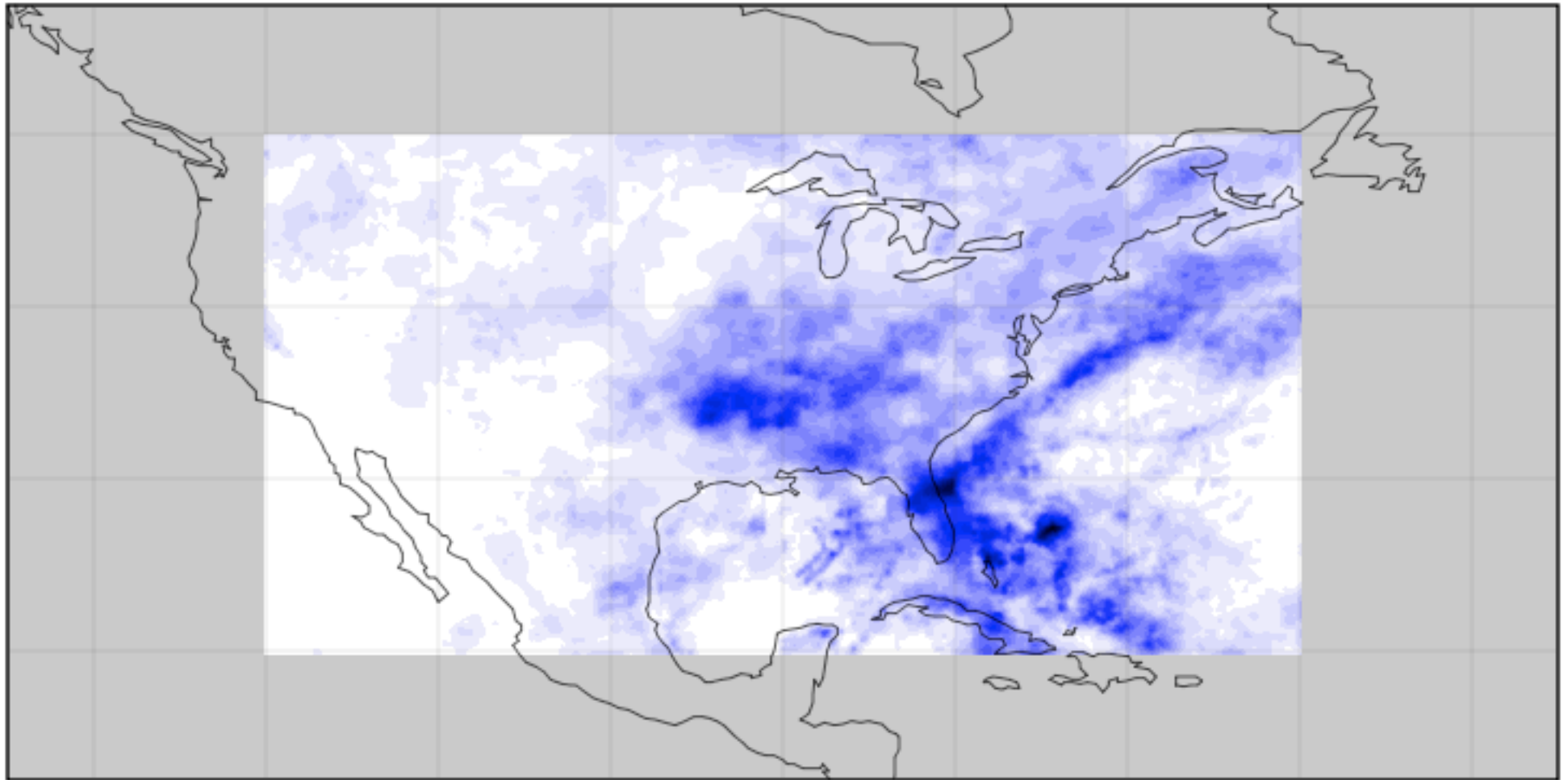
- Essentially a GrADS script from Arlindo DaSilva, NASA/GSFC/GMAO
- Same on-the-fly mechanism as NetCDF converter
- If you can read it into GrADS, you can convert it (usually)
 - Dimensions can be defined in GrADS control file
 - Limited to L3 and L4 gridded products
 - Handles both HDF-EOS and HDF
- Also supports subsetting

Select Subsetting Criteria (Spatial and/or Parameter) for the data set.			
<input type="button" value="Submit"/>			
South	<input type="text" value="20"/>	West	<input type="text" value="-120"/>
North	<input type="text" value="50"/>	East	<input type="text" value="-60"/>
TRMM_3B43.006			
Parameter Names		<input type="button" value="Select All"/>	<input type="button" value="Reset"/>
<input checked="" type="checkbox"/>	precipitation		
<input type="checkbox"/>	relativeError		
Select Subsetting Criteria (Spatial and/or Parameter) for the data set.			
<input type="button" value="Submit"/>			



TRMM Monthly (May 2009) in Panoply Lats4D Conversion + Subset

precipitation:Precipitation





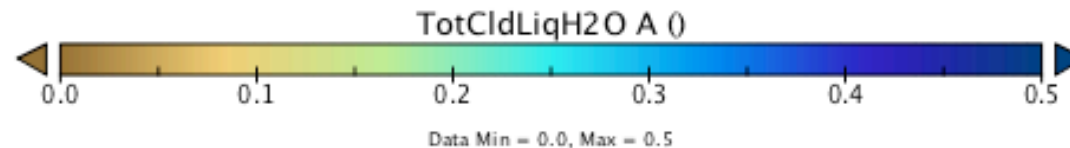
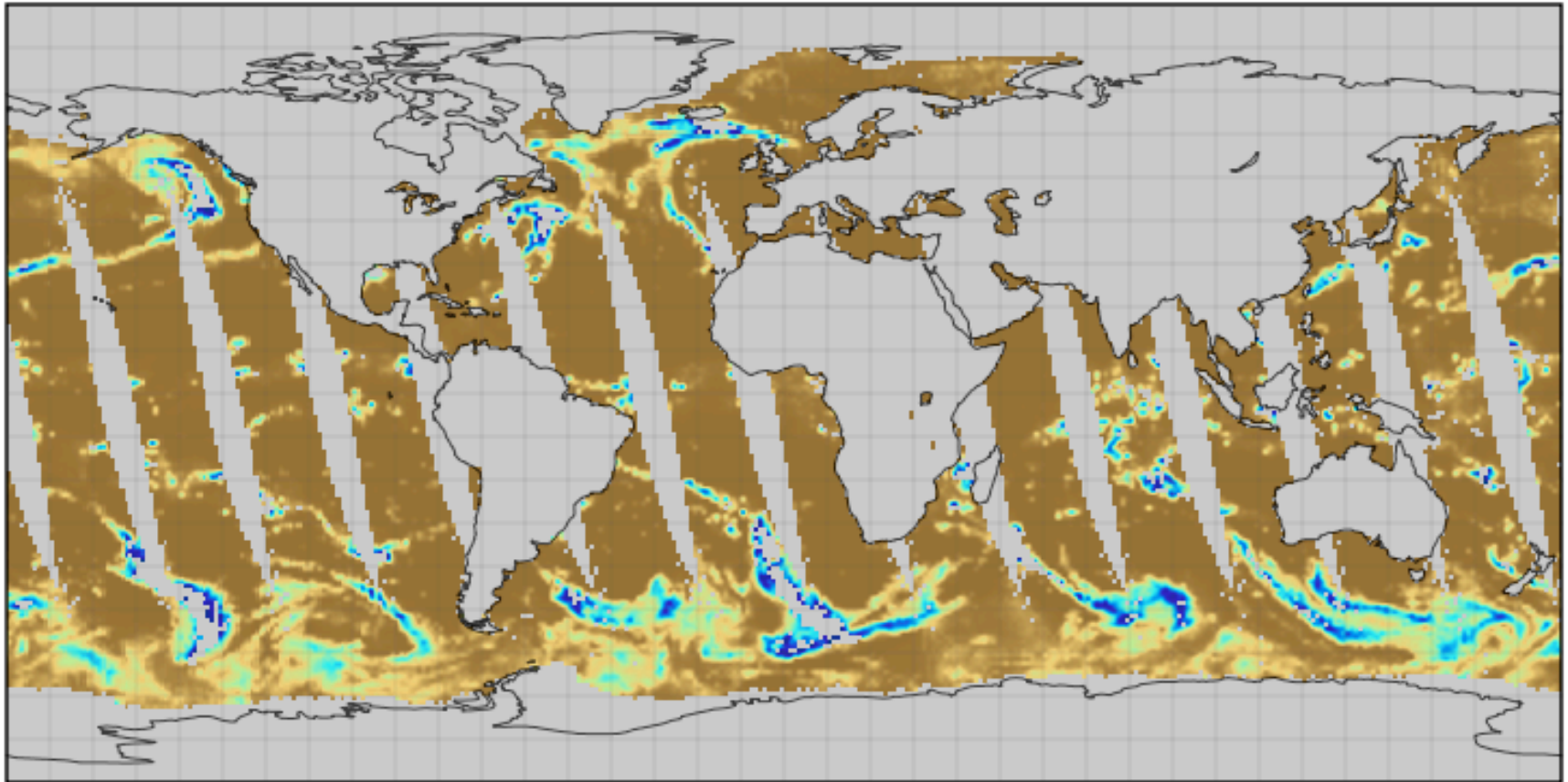
Solution #3: OGC Web Coverage Service

- OGC WCS with netCDF/CF-1 profile
 - Supports subsetting
 - But: “prefers” interpolation to user’s resolution
 - Requests are difficult for users to submit
 - Works better for machine-to-machine interfaces
 - DataFed
 - DLR
 - Using GMU-developed WCS Server



AIRS L3 via WCS in Panoply

TotCldLiqH2O A

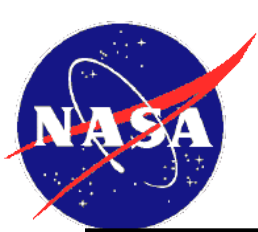


```
wget -O airs_wcs.nc 'http://acdisc.sci.gsfc.nasa.gov/daac-bin/wcsAIRSL3?VERSION=1.0.0&SERVICE=WCS  
&REQUEST=getCoverage&FORMAT=NetCDF&COVERAGE=AIRX3STD:ascending:TotCldLiqH2O_A  
&BBOX=-179.5,-89.5,179.5,89.5&TIME=2007-04-07/2007-04-07/P1D&CRS=CRS:84&RESX=1.0&RESY=1.0'
```



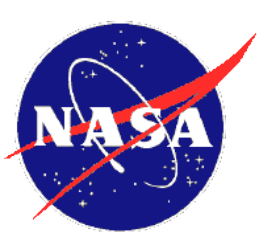

Solution #4: OPeNDAP

- HDF handlers
 - Released HDF5 version supports L3 grid
 - New HDF4 and HDF5 version supports
 - HDF4
 - L2 swath

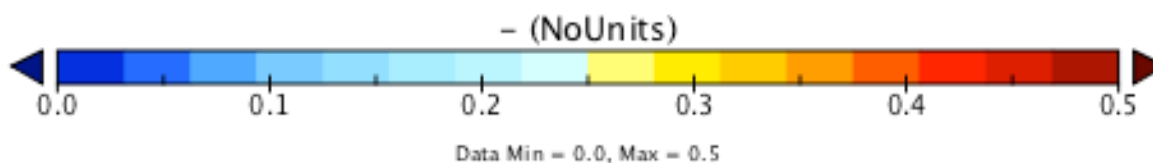
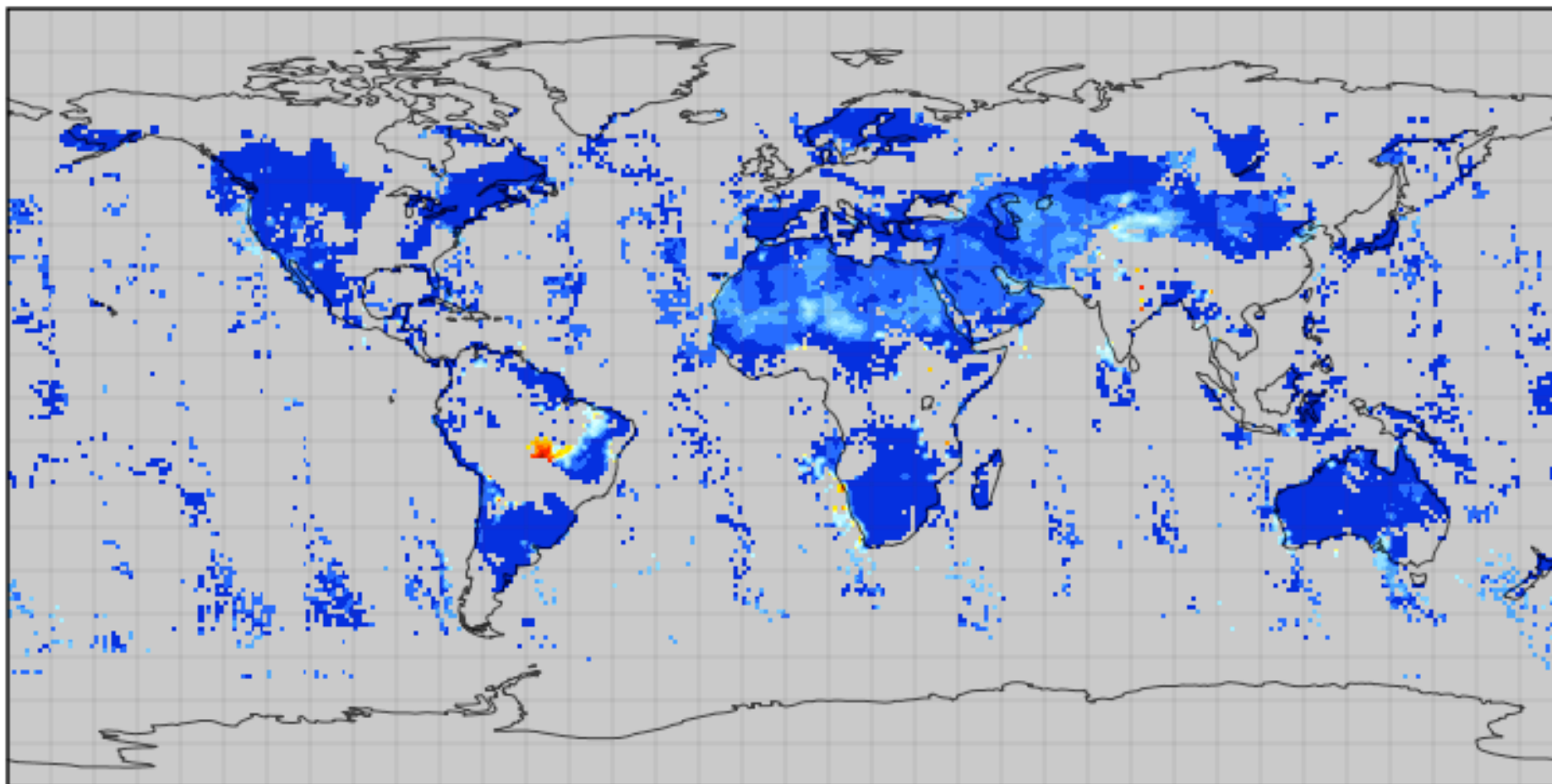


OPeNDAP Usability Results

	IDV	Panoply	Ferret	GrADS	Comments
AIRS L3	✓	✓	✓	✓	exc. AIRX3C2M
MODIS L3	✓	✓		✓	latest version only
OMI L3	✓	✓	✓	✓	
TOMS L3	✓	✓	✓	✓	
AIRS L2	✓	✓	✓	x	
HIRDLS L2				x	
MLS L2				x	
MODIS L2	✓	✓	✓	x	so far: MOD04, MOD07
OMI L2	✓		✓	x	exc. OMIBRO, OMHCO, OMOCLO, OMUVB

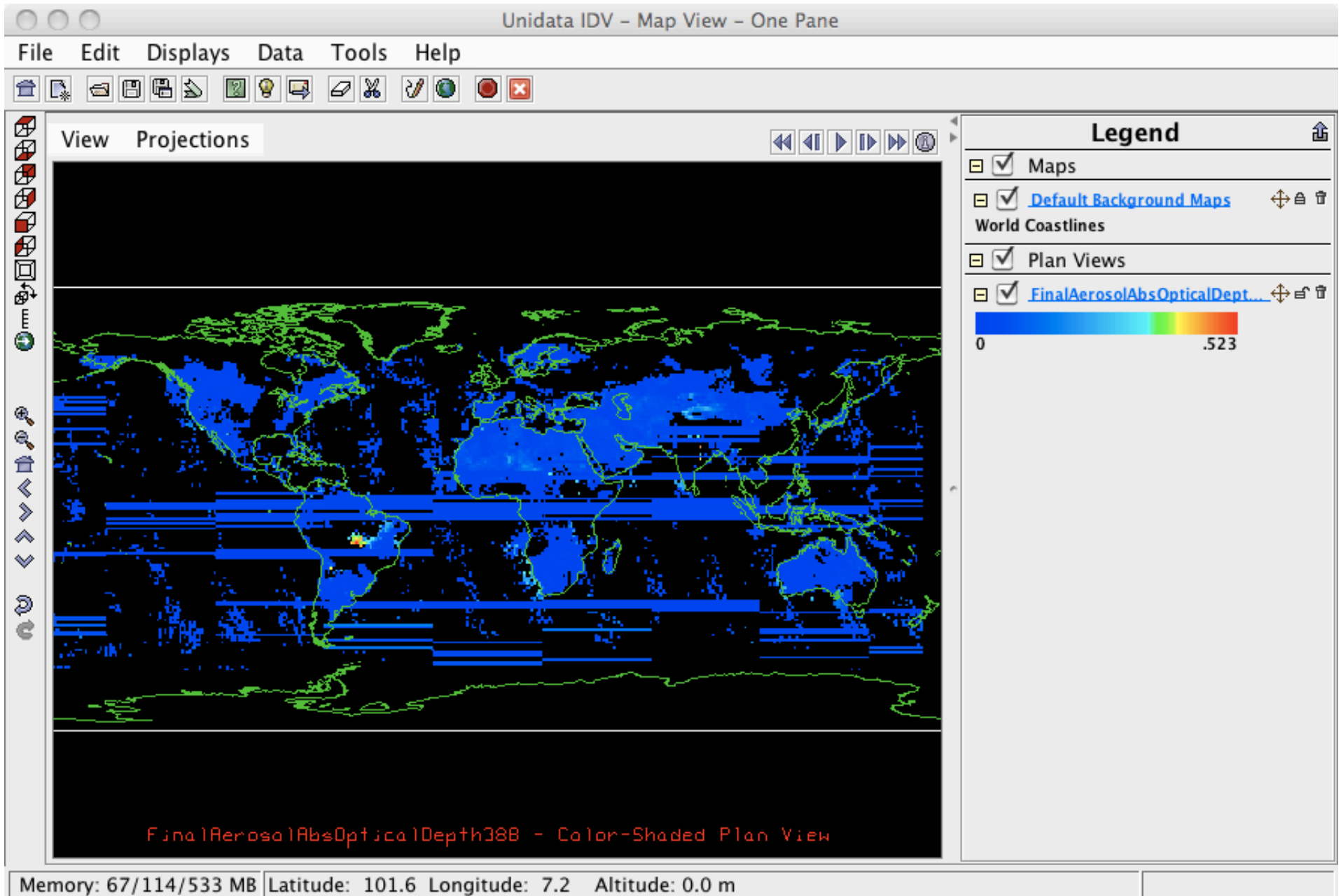


OMI L3 NearUV AAOD via OPeNDAP in Panoply



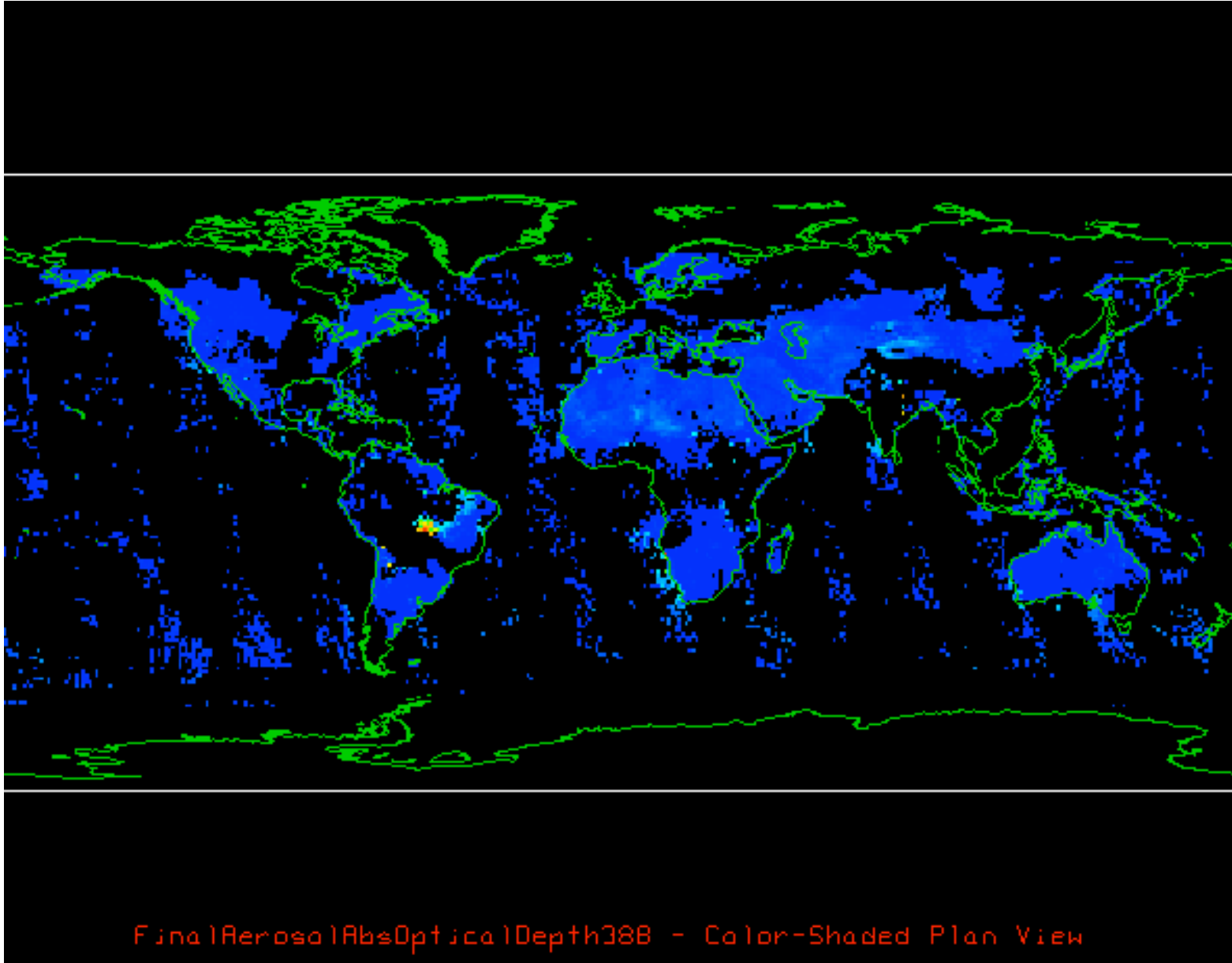


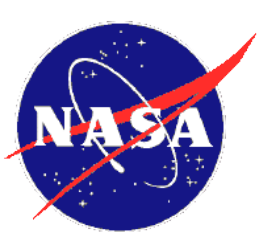
(but watch out for fast rendering in IDV)



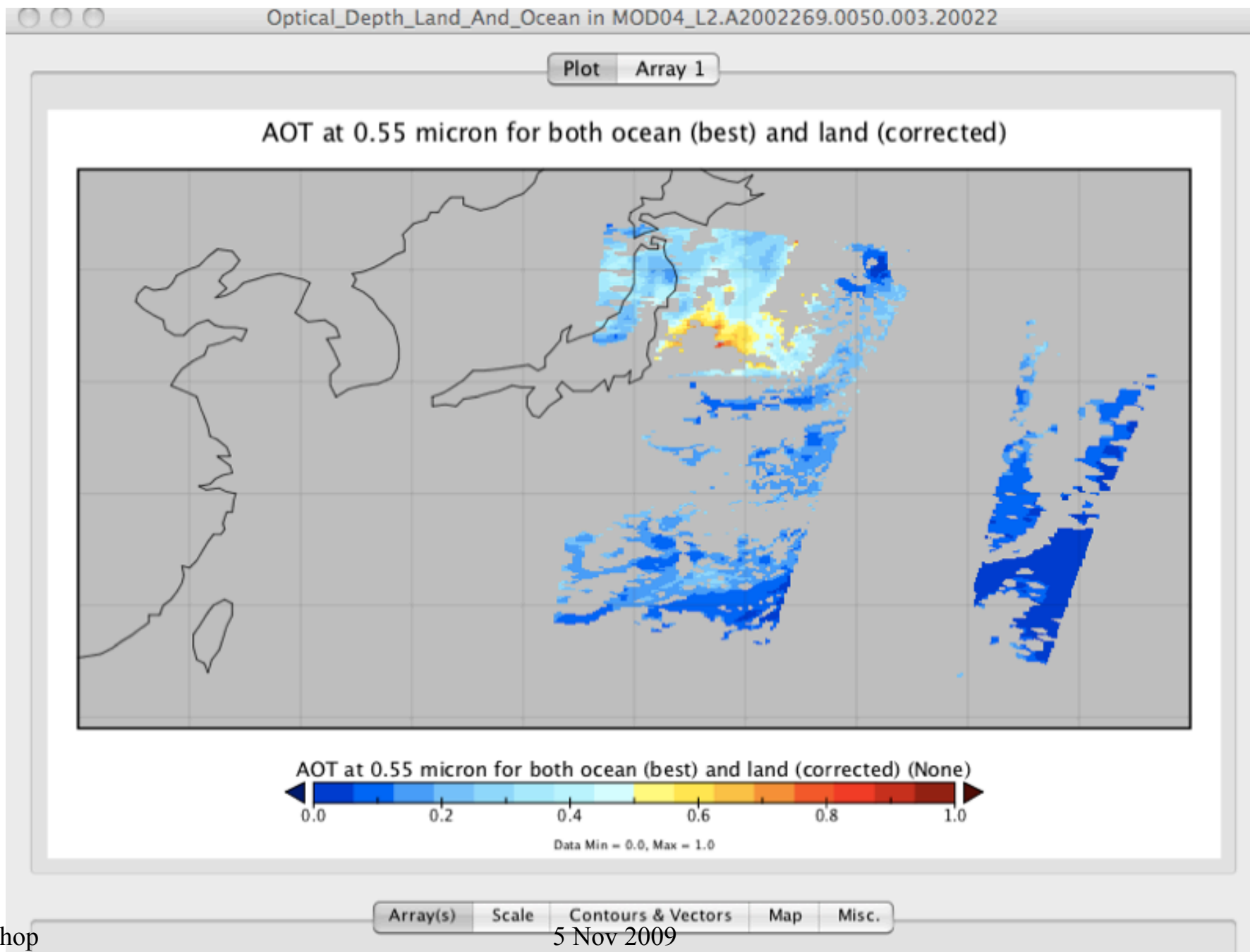


IDV w/fast rendering off



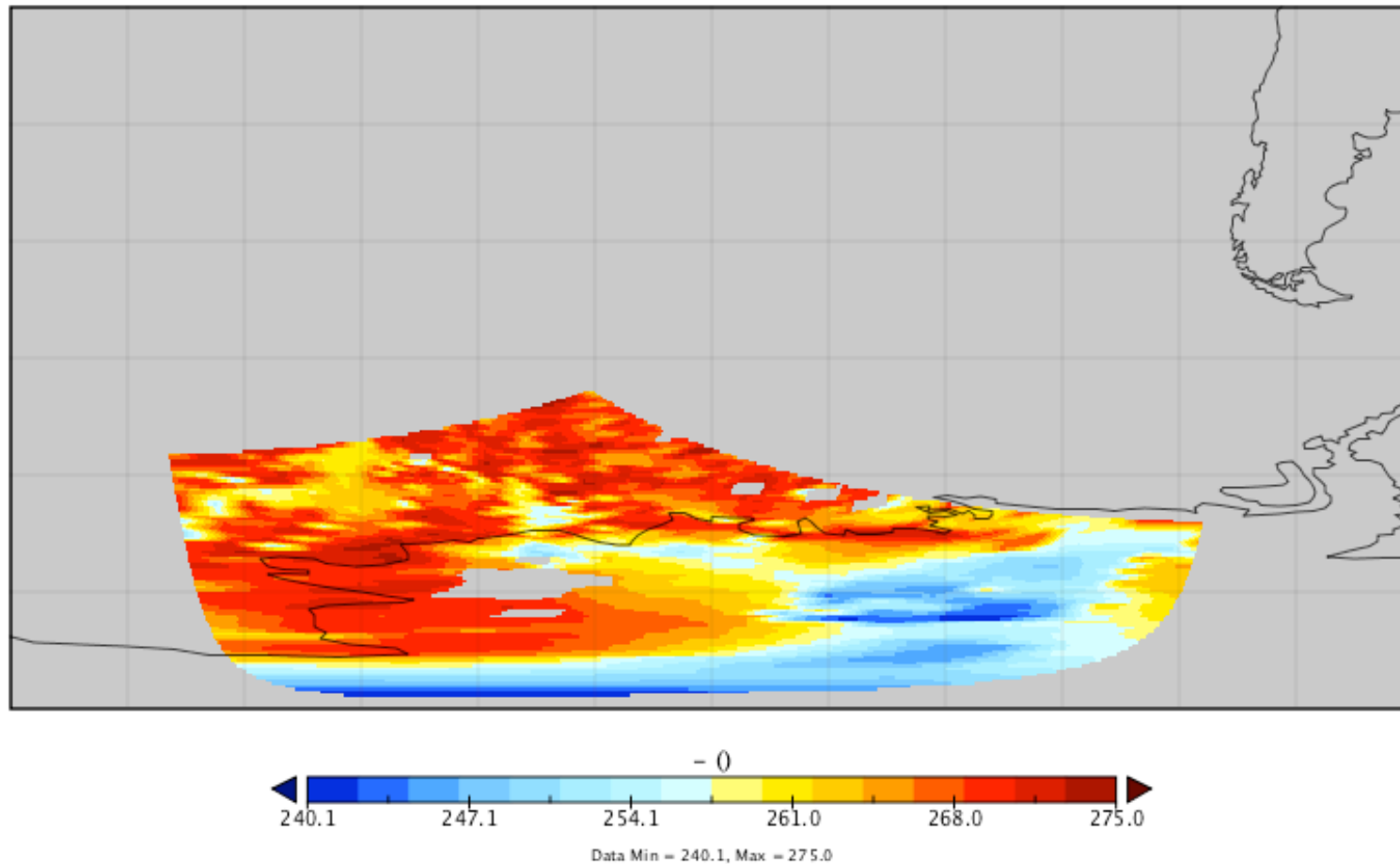


Panoply w/MODIS L2 via new OPeNDAP HDF4 handler





Panoply w/AIRS L2 via new OPeNDAP HDF4 Handler



- Cool..but who projected the data and how?



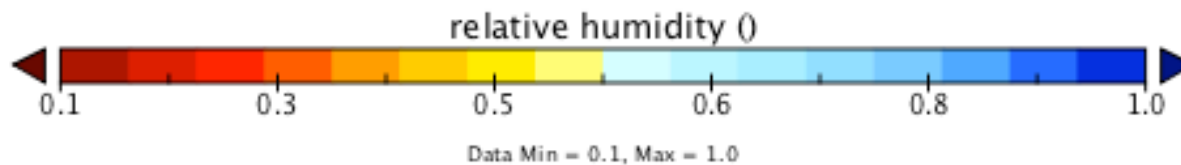
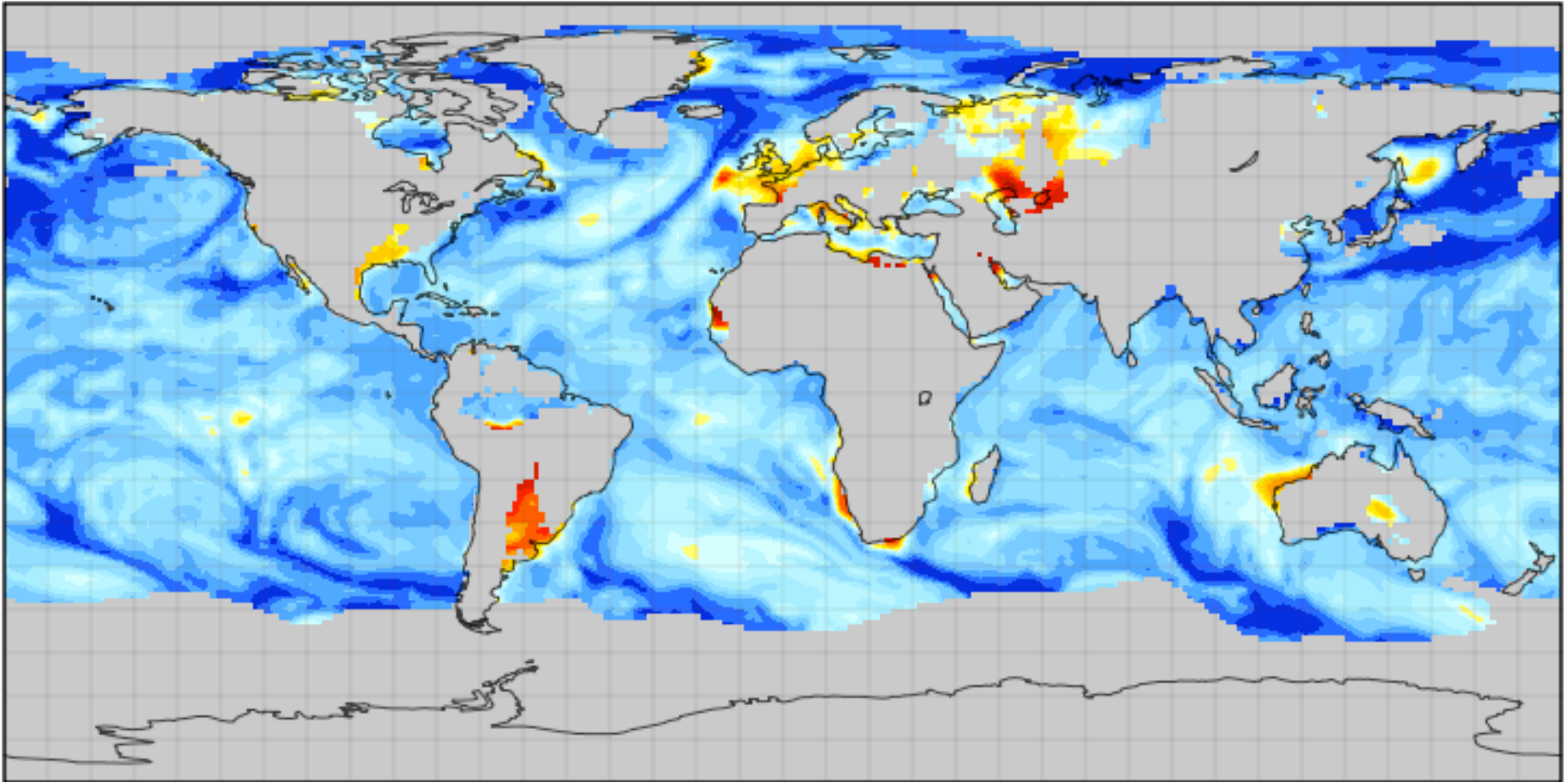
Solution #5: GrADS Data Server

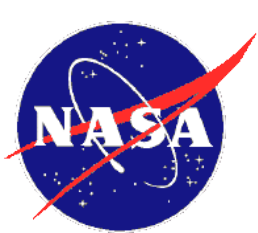
- DAP protocol, but data read by GrADS
 - Sort of like an OPeNDAP with a GrADS handler
 - If you can read it in in GrADS...
- Includes time aggregation
 - (Not *always* a good thing)



MERRA via GDS in Panoply

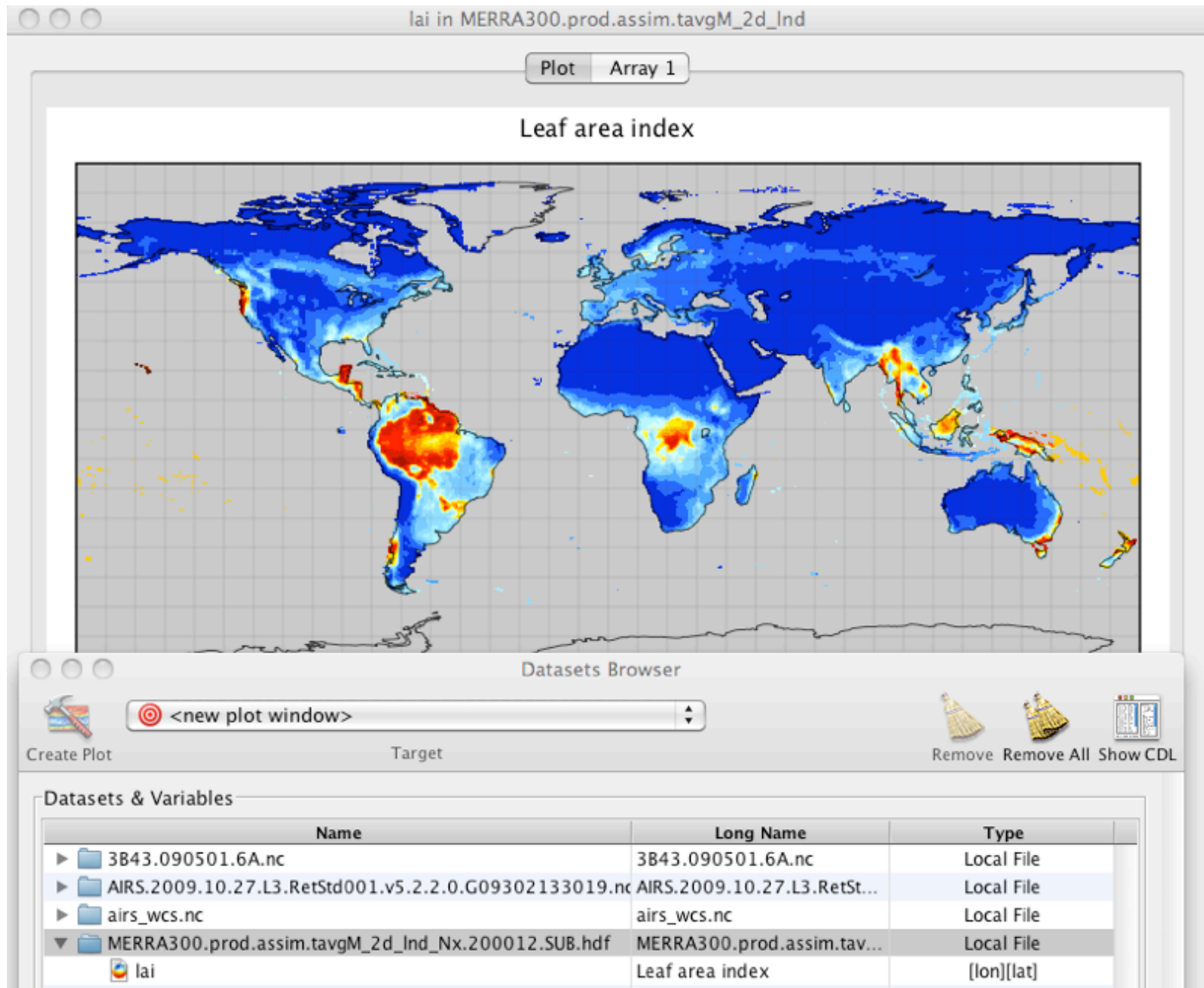
relative humidity





Solution #6?: netcdf library

MERRA *HDF* Subset from Lats4D in Panoply



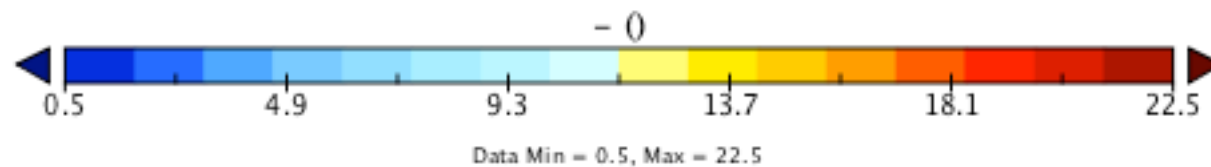
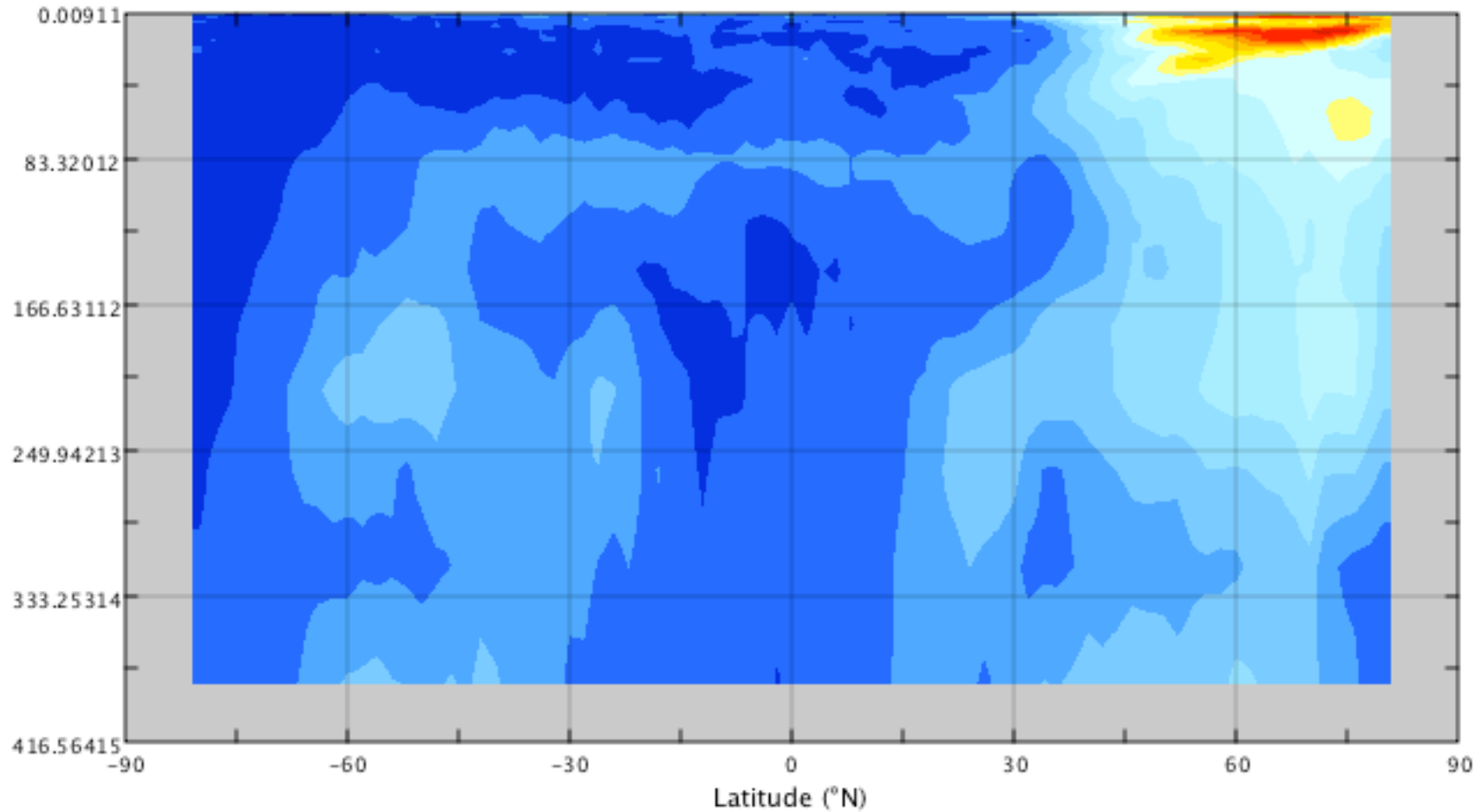


netcdf libraries

- netcdf-java library can read some HDF files
 - But not HDF-EOS...
 - Tried adding lat/long dimscales to an HDF-EOS2 grid, to no avail
 - Except for an MLS L3 HDF-EOS5
 - (not a standard product)



Panoply, using netcdf-java w/ MLS L3, HDF-EOS5





Can we make netcdf libraries read HDF-EOS?

- Might require joint effort between UCAR, THG and tool developers to develop and debug...
- OR, solution might emerge from HDF-EOS5 / netCDF-4 Augmentation effort
- ...BUT
 - could make a *lot* of conversions unnecessary
 - would make data usage much easier for users



Trade Space

	Pros	Cons
Custom C++ Conversion	L2 swath + L3/L4 grid Fast	Costly Data size expansion
Lats4D Conversion	Easy to add HDF + HDF-EOS	L3/L4 grid only Slow
WCS	Subsetting	Interpolation Costly
OPeNDAP	L2 swath + L3/L4 grid Efficient Inexpensive	Reprojects L2 swath
GrADS Data Server	HDF + HDF-EOS	L3/L4 only Slow
netcdf library	Enables many tools Obsoletes conversion	HDF-EOS



Future of HDF-netCDF Tool

Interoperability: Key Challenges

- Level 1b data (calibrated radiance)
 - How should we fit the channel/wavelength dimension into CF1?
- Level 2 in native swath coordinates
- netCDF library direct access to HDF-EOS
- Non-self-describing implementations of HDF
 - e.g., NPP