HDF Group Support for NPP/JPSS

Mike Folk, Elena Pourmal, Larry Knox, Albert Cheng
The HDF Group

DEWG Meeting
June 19, 2012

Goal

Provide HDF5 risk-reduction support for the distribution of NPP/JPSS VIIRS, OMPS, and other sensor and environmental data products

PROJECT INFORMATION

Project Information

- Project Web site

  http://www.hdfgroup.org/projects/npoess/

- Project background
  - Started in 2009
  - NASA POCs are Richard Ullman and Alfreda Hall
  - Areas of primary needs
    - Easy and intuitive NPP/JPSS data access and use by diverse communities
    - Allow producers and consumers to view content, manage metadata, and convert data to other formats.
    - High-quality, rapid-response for NPP/JPSS users at all levels.

Project Information

- 2011-2012 Priorities
  - Maintain HDF software and software developed for NPP/JPSS on platforms critical to NPP/JPSS project
  - Provide rapid and high priority support for data producers and users
  - Develop software to facilitate access and management of NPP/JPSS products

HDF5 SOFTWARE MAINTENANCE

HDF5 Software Maintenance

- HDF5 library and command-line utilities
- http://www.hdfgroup.org/HDF5/
  - C, C++ and Fortran90 APIs
  - Command-line utilities to view, compare and copy data stored in HDF5 files
    - h5dump, h5ls, h5diff, h5copy

HDF5 Software Maintenance

- Latest release HDF5 1.8.9 on May 15, 2012
- For platforms and compilers supported for each release see http://www.hdfgroup.org/HDF5/release/obtain5.html
- Tested daily with NASA HDF-EOS5 and NPP/JPSS files from GRAVITE system

HDF5 Software Maintenance

- Let us know your needs! We will help with:
  - Porting to new platforms and compilers
  - HDF software features
  - Performance tuning
- Report defects and suggest improvements to HDF products
- How can you reach us?

PRIORITY USER SUPPORT

Support for NPP/JPSS Users

- Priority support to NPP/JPSS users
  - Send email to help@hdfgroup.org
  - Responding to “NPOESS”, “NPP”, or “JPSS” in subject field
    - *Subject: [NPOESS] Where is the plug-in?*
    - *Subject: [NPP] Why my program doesn’t work?*
    - *Subject: [JPSS] Need a tool to edit files with references*
    - **This has worked well, though occasionally we have to guess**

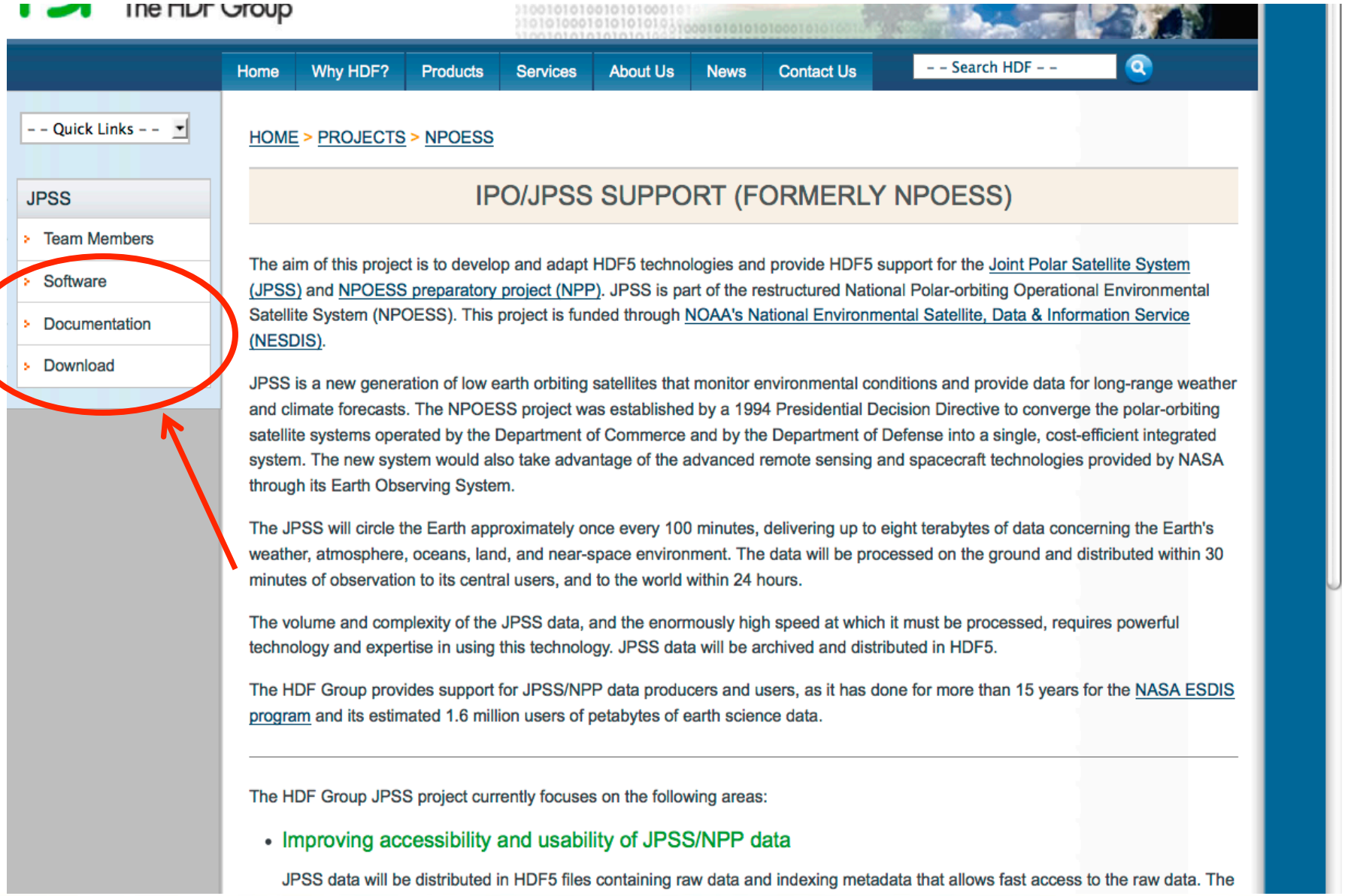
Support for NPP/JPSS Users

- Join us at the HDF and HDF-EOS Workshops
  - Tutorials
  - Consultations
  - Discussions of future directions
  - http://www.hdfeos.net/workshops/index.php

Software developed at The HDF Group for NPP/JPSS

SPECIAL PROJECTS

Software Developed for JPSS



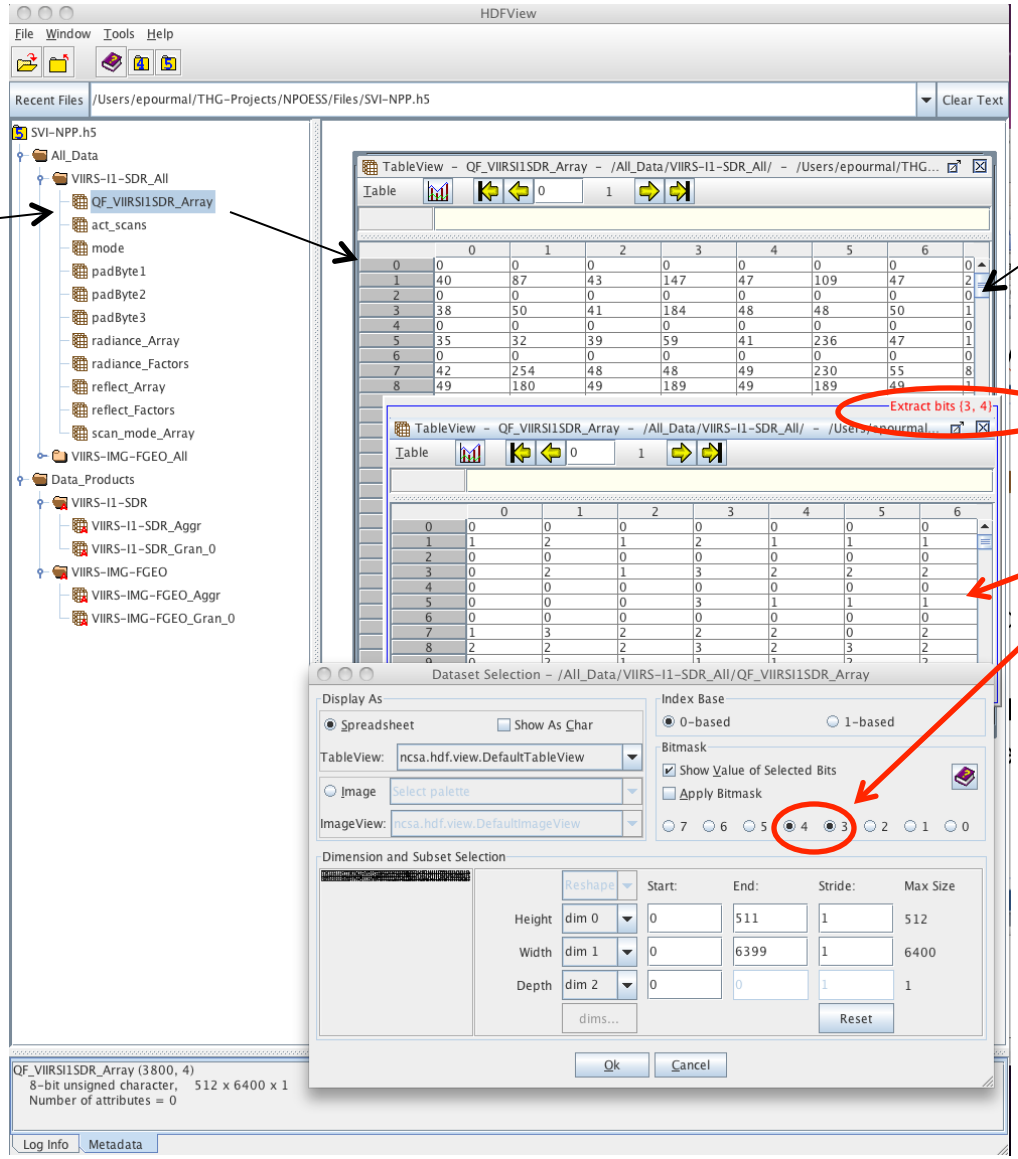
“JPSS FRIENDLY” HDF5 TOOLS

HDFView

- HDF Java Browser
- http://www.hdfgroup.org/hdf-java-html/hdfview/
- Version 2.8 released in December 2011
- Platforms supported
  - Linux, Mac OS X, Windows, Solaris
- JPSS specific features
  - Displays values of quality flags
  - Navigates through object and region references

JPSSS Quality Flags in HDFView

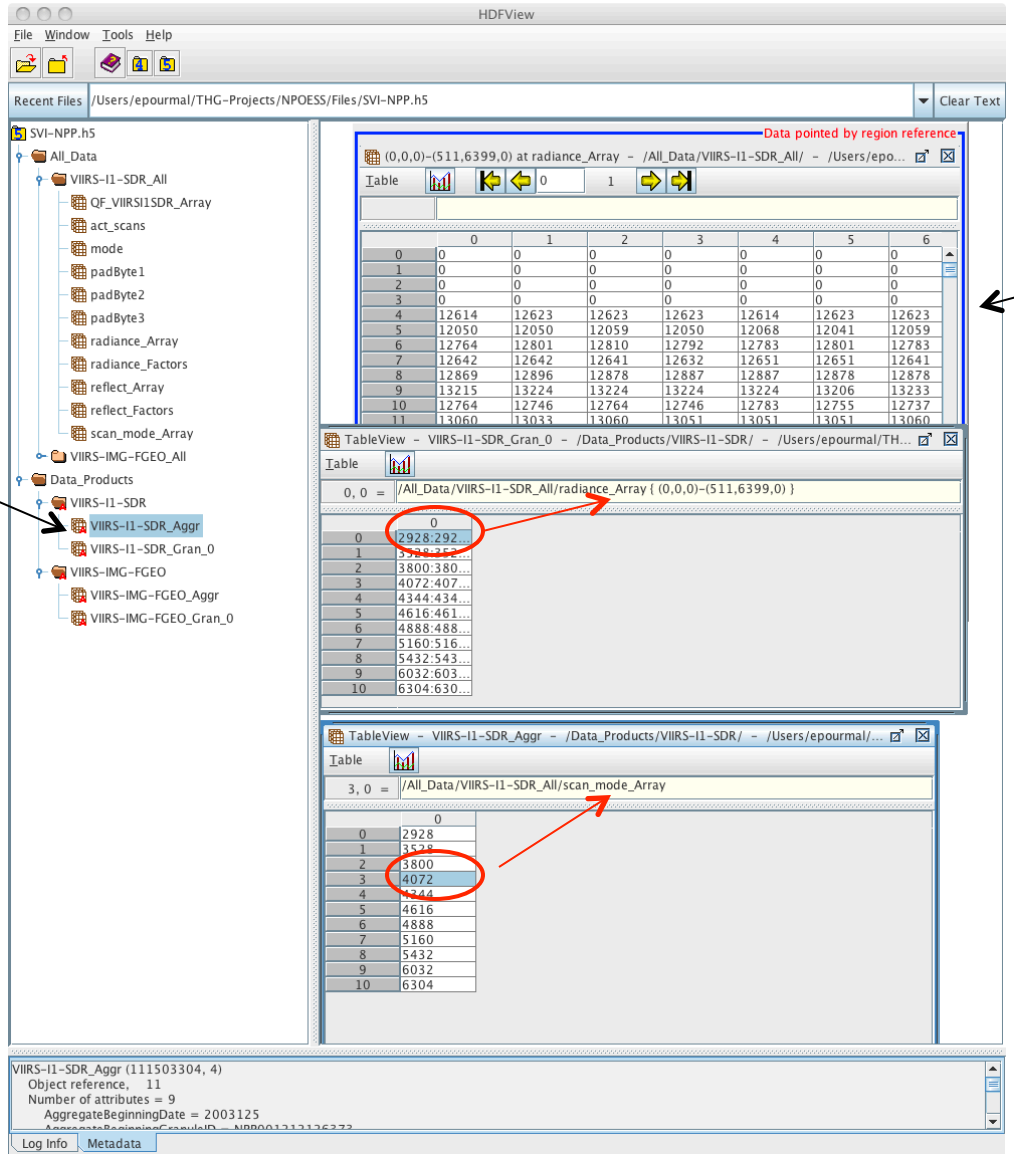
Dataset with quality flags



Flags packed in 8 bits

Extracted bits

HDF5 Object and Region References in HDFView



Data of the selected region

Dataset with HDF5 object references

h5dump

- h5dump is a command line tool to display the content of an HDF5 file
- Example: dump metadata information (no data displayed)

```
h5dump -H SVI-NPP.h5
```

- Example: display datasets pointed by object references

```
h5dump -d /Data_Products/VIIRS-I1-SDR/VIIRS-I1-
SDR_Aggr -R SVI-NPP.
```

- Example: display quality flags

```
h5dump -d /All_Data/VIIRS-I1-SDR_All/
QF_VIIRSI1SDR_Array -M 3,1 SVI-NPP.h5
```

h5edit

- h5edit is a command line tool to create and edit attributes
  - The tool is under development; more functionality is coming
- Example: add “scale_factor” attribute

```
h5edit -c "CREATE /Radiance scale_factor
{H5T_IEEE_F32LE SIMPLE(1) DATA{2.8339462E-4}};"
file.h5
```

- Example: add “units” attribute:

```
h5edit -c "CREATE /Longitude units {H5T_STRING
{ STRSIZE 12 } DATA {'degrees_east'}};” file.h5
```

High-Level Library

- Functions for conveniently handling HDF5 object and region references and packed bits (quality flags)
  - C and Fortran API to support NPP/JPSS data
  - Examples:
    - Create a region reference to a rectilinear region
    - Create and write a datasets with object references
    - Read quality flags
  - Fortran functions require HDF5 1.8.5 and later
  - Tested daily on Linux and AIX

JPSS SOFTWARE

JPSS Software

- Software developed at The HDF Group to support the JPSS project
  - High-level library for handling HDF5 object and region references
  - Augmentation tool h5augjpss
  - Aggregation tool nagg

http://www.hdfgroup.org/projects/npoess/jpss_software.html

Augmentation tool – h5augjpss

- h5augjpss is a command line tool to change an HDF5 JPSS file to make it accessible by netCDF-based applications
  - Hide HDF5 elements not supported by netCDF applications
  - Add metadata or data needed by netCDF or netCDF applications
- Attention!
  - h5augjpss modifies the file! Make a copy if you need to preserve original data

Clarification

- netCDF-3 files
  - Based on netCDF **classic** data model
- netCDF-4 files
  - Based on netCDF **enhanced** model
  - Uses HDF5 as a storage layer
  - Group hierarchy, user-defined data types, etc.
  - But can be restricted netCDF classic
- NPP files
  - HDF5 file with primary data
    - Incompatible with netCDF, unless modified
  - XML metadata file
    - Important information, including dimensions
  - Geo data in separate file, or group in primary file

Example: ncdump of an augmented file

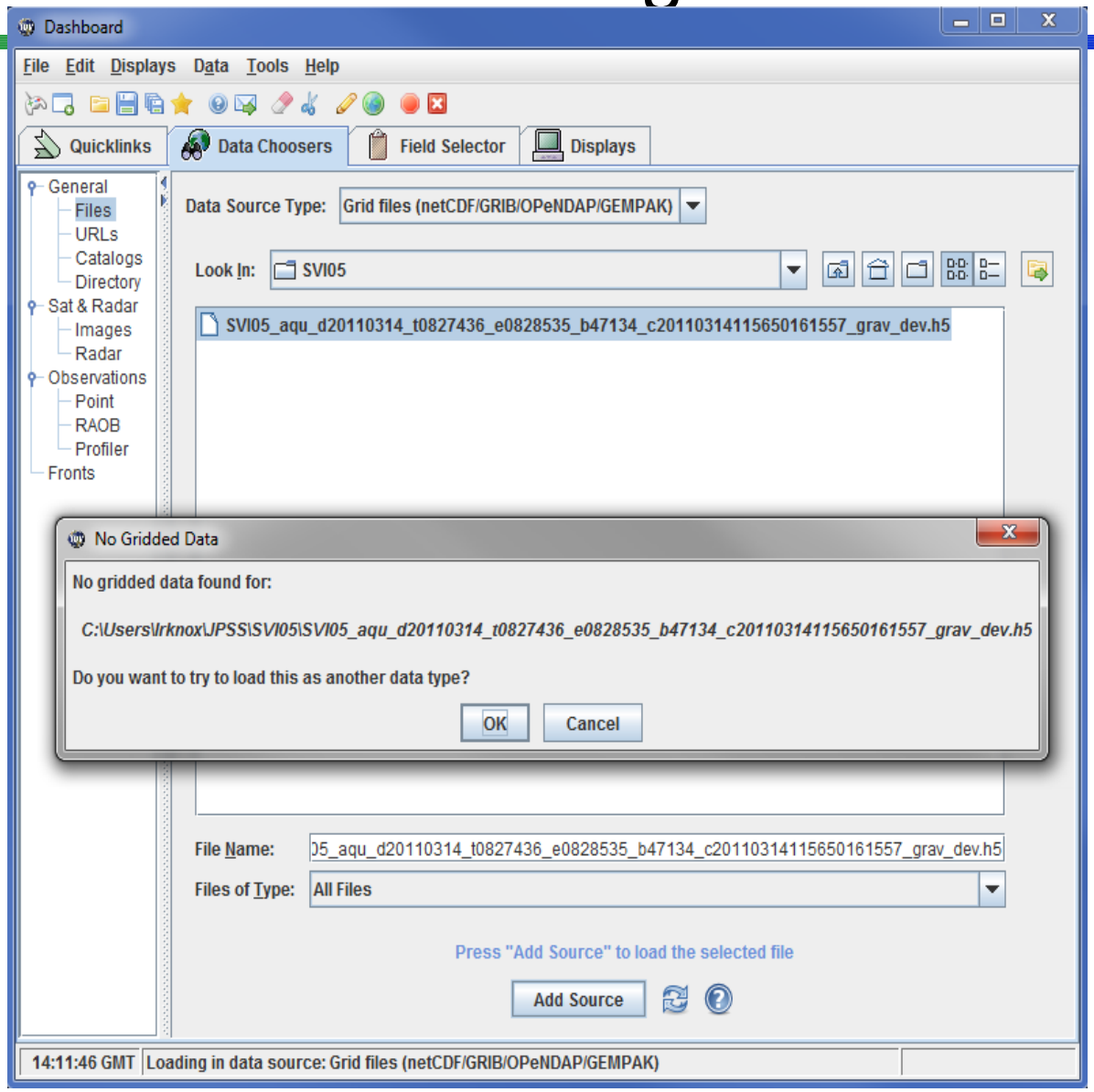
```
netcdf SVM07_ter_augmented-step1-step2 {
// global attributes:
                ……..
group: All_Data {
  group: VIIRS-M7-SDR_All {
    dimensions:
        AlongTrack = 768 ;
        CrossTrack = 3200 ;
        …
    variables:
        int AlongTrack(AlongTrack) ;
        ….
        float Radiance(AlongTrack, CrossTrack) ;
                Radiance:DatumOffset = 0 ;
                Radiance:Scaled = 1 ;
                string Radiance:MeasurementUnits = "W/(m^2 μm sr)”  Extra attributes
                ….
} // group VIIRS-M7-SDR_All
  } // group All_Data
}
```

Meaningful dimensions

Coordinate variables

Product attributes

IDV before augmentation



Two problems

- Has to be netCDF-3 conformant

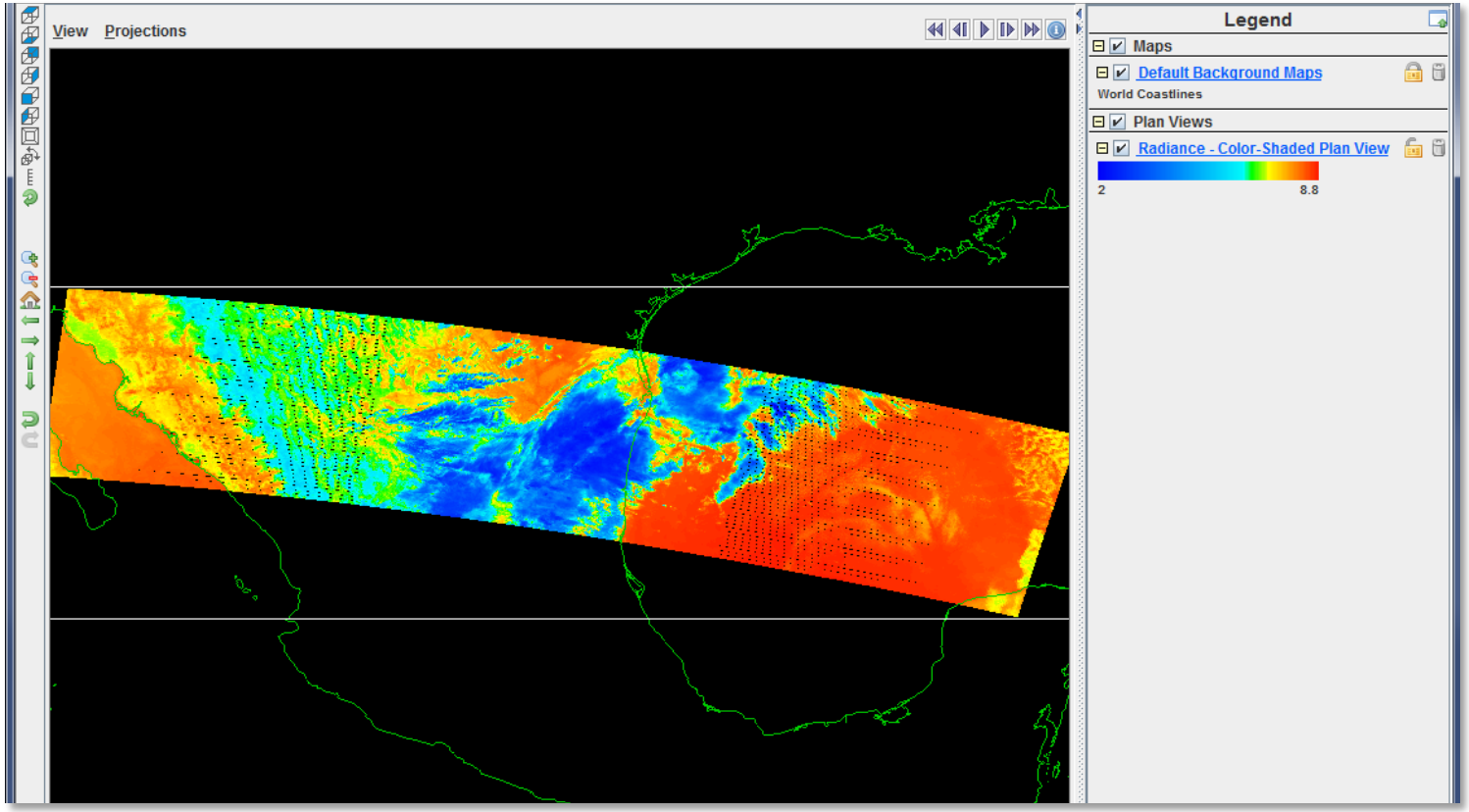
```
h5augjpss -o4 SVI05.h5
```

- Missing CF metadata

```
h5edit -c "CREATE /Radiance scale_factor
{H5T_IEEE_F32LE SIMPLE(1) DATA
{2.8339462E-4}};" SV105.h5

h5edit -c "CREATE /Longitude units {H5T_STRING
{ STRSIZE 12 } DATA {'degrees_east'}};"
SV105.h5
```

IDV after augmentation



IDV with augmented JPSS file

Aggregation tool – nagg

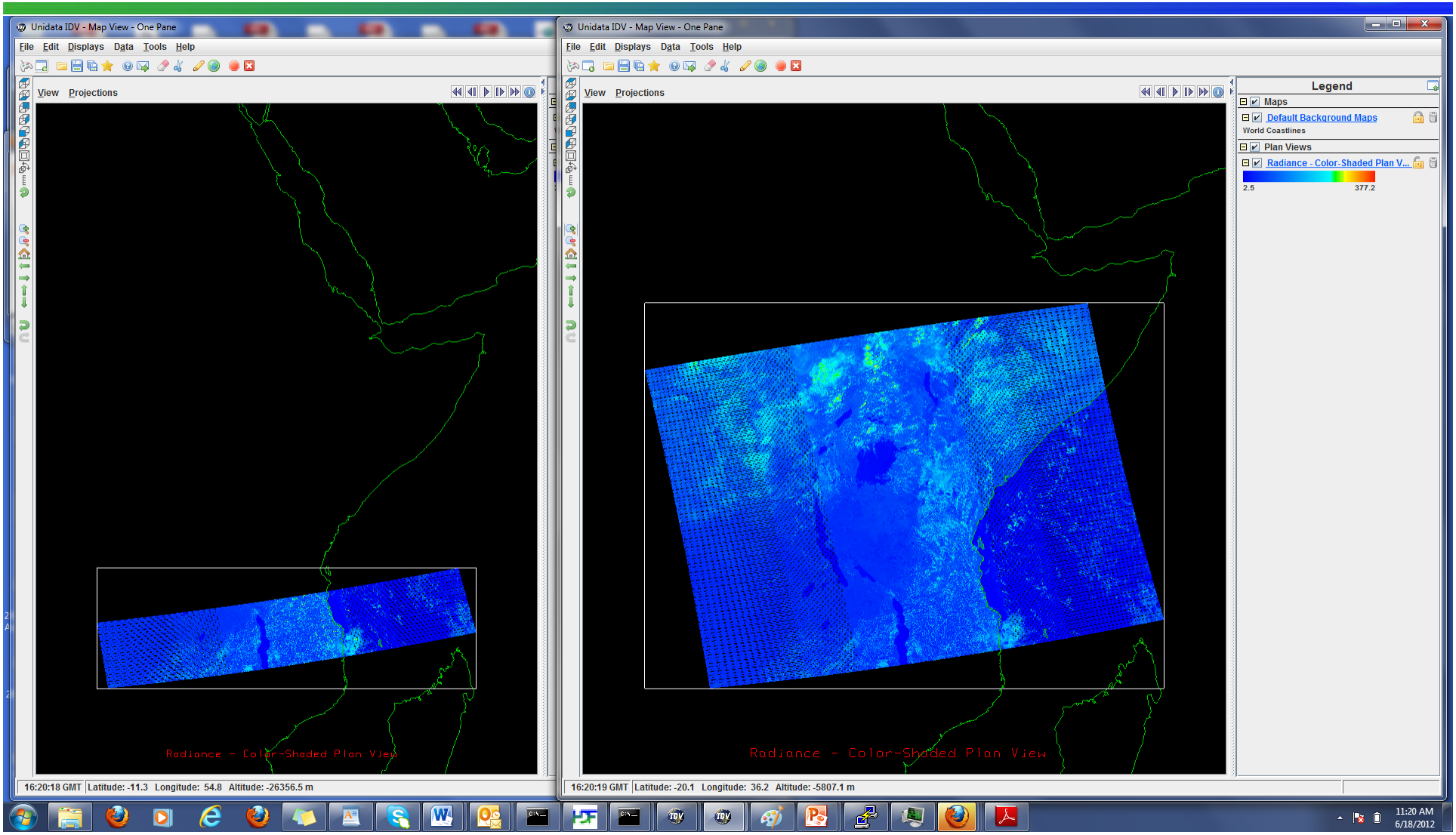
- nagg is a command line tool for aggregating JPSS data granules from existing files into new files with
  - Different number of granules per file or
  - Different combinations of compatible products than in the original files

Why nagg?

- NPP data products organized as “granules.”
- Granules are relatively small.
- Several granules may be packaged per file.
- Several products may be packaged per file.
- For convenience of a particular application, we may want to re-package them.
- May also want only a subset of them.

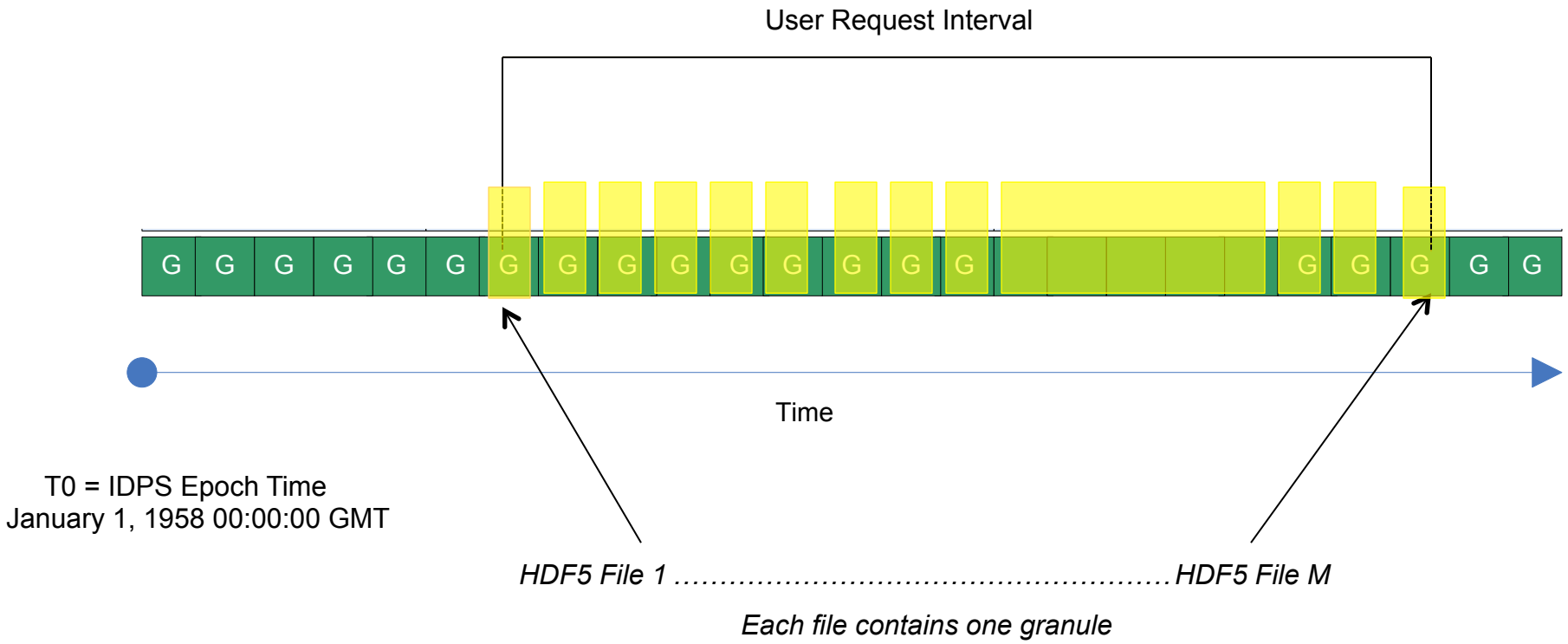
(Thanks to Richard Ullman)

Aggregated file in IDV



One granule is displayed on the left; four aggregated granules are displayed on the right

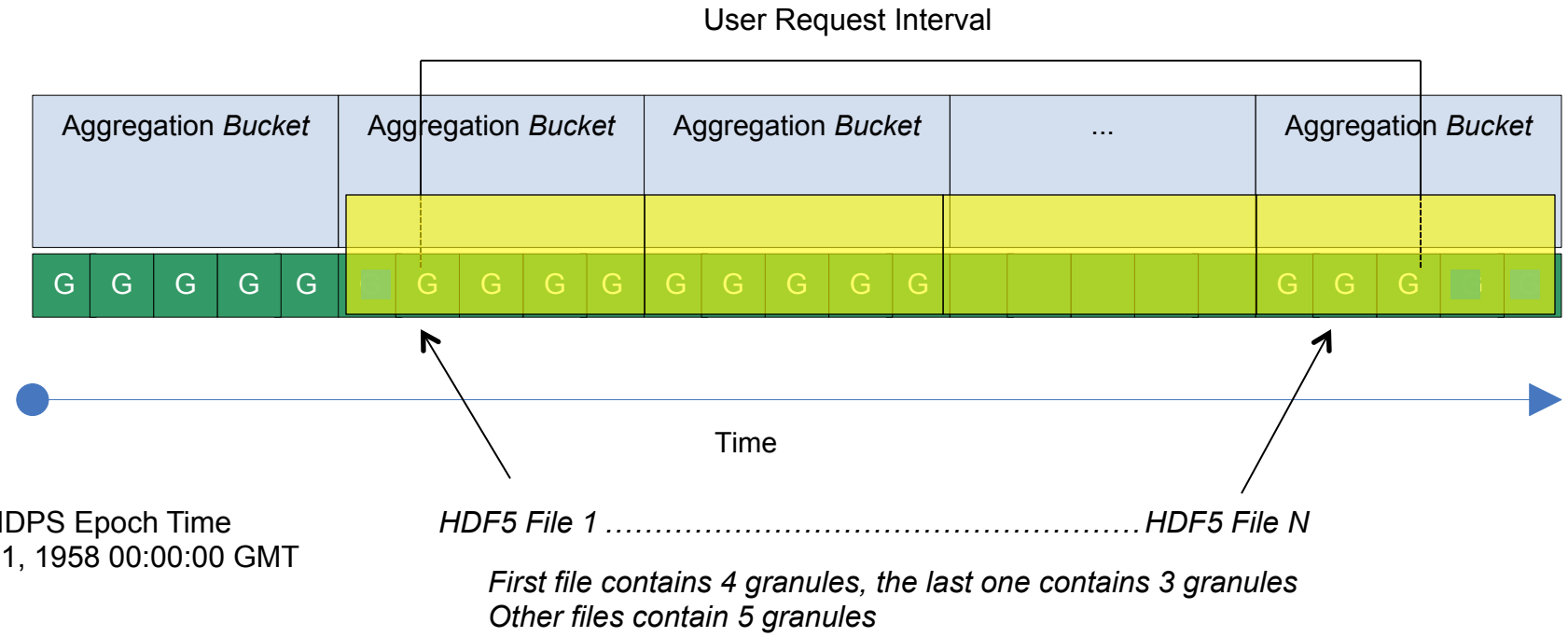
Aggregation Example



- User requests data from the IDPS system for a specific time interval
- Granules and products are packaged in the HDF5 files according to the request
- This example shows one granule per file for one product

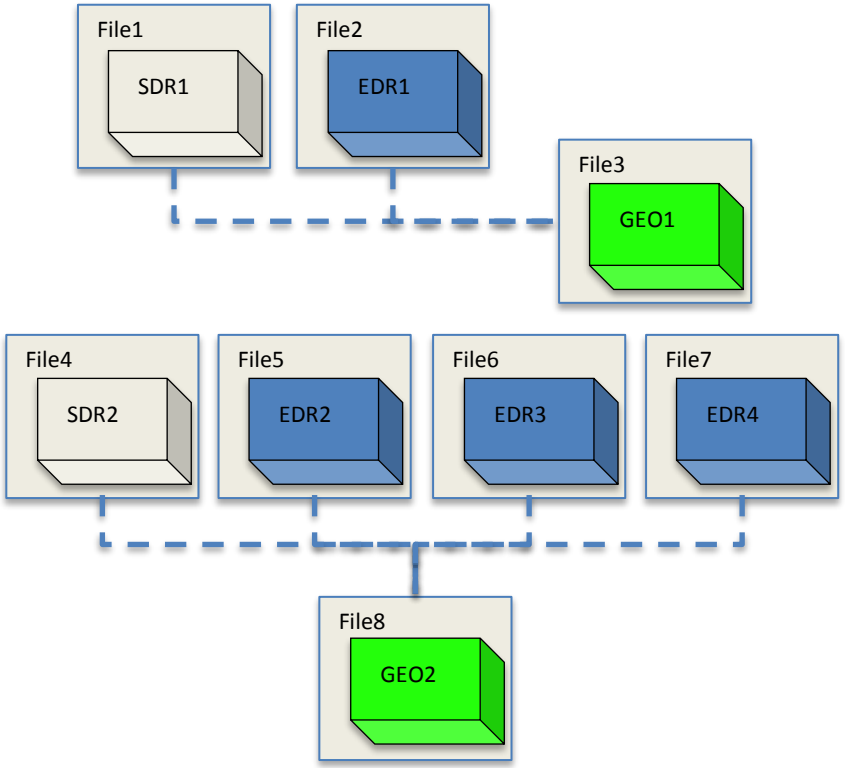
Aggregation Example

Example: nagg –n 5 –t SATMS SATMS_npp_d2012040*.h5



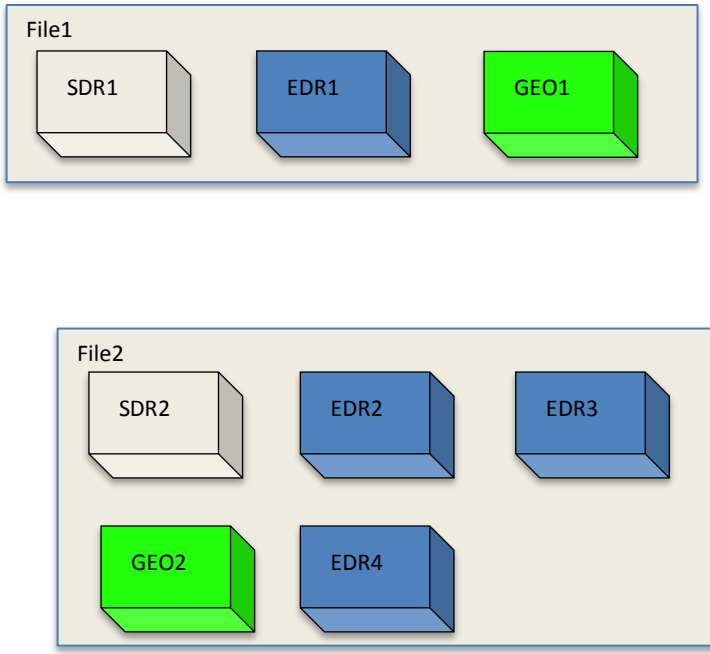
- Produced files co-align with the aggregation bucket start
- HDF5 files are 'full' aggregations (full, relative to the aggregation period)
- Geolocation granules are aggregated and packaged; see –g option for more control

IDPS Packaging Baseline
CDFCB-X Volume I



Packaging Off

Packaging only applies to products with geolocation data



Packaging On

Always makes a new copy. Doesn't destroy the original file.

OTHER SOFTWARE OF INTEREST

H5Py

- HDF5 Python APIs
- http://code.google.com/p/h5py
- Written and supported by Andrew Collette (not The HDF Group)
- Contains
  - High-level Python methods for HDF5
  - Low-level (follows C) Python APIs
- Easy to learn and use

H5Py Example

- Traverse the file, print object name and names and values of its attributes

```
import h5py
def print_info(name, obj):
    print name
    for name, value in obj.attrs.iteritems():
    print  "  "+name+":", value

f = h5py.File('GATMO-SATMS-npp.h5', 'r+')
f.visititems(print_info)
f.close()
```

H5Py Example Output

```
……
Data_Products/ATMS-SDR/ATMS-SDR_Gran_0
   Ascending/Descending_Indicator: [[1]]
   Beginning_Time: [['040230.416997Z']]
   N_LEOA_Flag: [['On']]
   Band_ID: [['N/A']]
   Beginning_Date: [['20030126']]
   East_Bounding_Coordinate: [[-19.96759987]]
   Ending_Date: [['20030126']]
   Ending_Time: [['040301.825163Z']]
   G-Ring_Latitude: [[ 25.41589928]
……..
```

FUTURE DEVELOPMENT EFFORTS

Future Efforts

- Finish nagg development
  - Add missing command line flags
  - Enhance test suite
  - Improve documentation
- Continue development of h5edit
- Continue with software maintenance and priority user support

- Questions?

- Requests?

Thank You!