International EOS DB Meeting 2003

This meeting was held from 17-20 November 2003 and was the sixth in the series of EOS DB meetings. There were about 50 participants, mostly from the US and Pacific region, with only a handful from Europe. The sounding instruments on the EOS platforms, AIRS, AMSU and AMSR-E were prominent in the discussions though many users were presenting mostly MODIS applications and activities. The full set of presentations is on the web at: http://www.eos.hawaii.edu/eos_db.html.

The GSFC Direct Broadcast Group released a pair of DVDs to the community at the meeting containing a number of scripts and ancillary datasets and institutional code for use in the DB environment. I obtained copies and subsequently distributed these to the Bureau of Meteorology, CSIRO Marine Research (Hobart), Geoscience Australia (ACRES) and the Department of Land Information (WA Govt.). The CSIRO Earth Observation Centre also has a copy.

What follows is my summary of the subset of talks for which I took notes or had particular interest in. The complete set on the web at the above address should be used in conjunction with my notes.

Edward King, 2004.

Jim Dodge – Introduction (to EOS DB)

Goal of this meeting is to focus more help towards software and products (ie less of an emphasis on hardware systems for DB). It is becoming more and more clear that version control of software and products is of increasing importance to the user community, especially when it comes to longer term applications like change detection.

Electronic papers from this meeting will be collected by Dodge and Coronado for inclusion on the direct readout portal (http://directreadout.gsfc.nasa.gov/)

This is the sixth DB meeting (the previous being in 1997, 1998, 1999, 2000, and 2001)

DB is clearly popular and useful – as evidenced by the more than 70 attendees at this meeting. The widespread participation has made a great deal of difference to its support within NASA, especially in the current political climate (i.e. 9/11+) when there have been a number of attempts to propose turning it off. In future U.S. systems there will almost certainly be encryption included so that (selective) data denial can be implemented. For almost everyone (i.e. everyone who is not the bad guys) that will simply mean that they need to register in order to be kept in the DB loop with decryption keys.

NASA’s Earth Science Enterprise has five basic science questions:

1. How is the earth system changing
2. What are the forcings of these changes

3. How is the system responding
4. What are the consequences of this response
5. How can we predict these

These lead into a further 23 detailed questions, that can be found on a NASA web site somewhere, presumably not too far from the Vision and Mission statements. All of the activity is now grouped into 6 key focus areas, with a science basis, which would make it very difficult if not impossible to get a proposal to support a DB system up today. We are very fortunate to have it, shouldn't take it for granted, and need to work to maintain it.

The direction that the science is taking indicates that the interactions between different elements of the earth system are important and that it is a complicated, if not complex, system. There is a consequent need to bring both models and data together to address the issues of concern.

DB is providing an opportunity for people with a pressing need to study and characterise regional and local events to place them within a global context.

NASA has also recognised a need to train and encourage the next generation of scientists – a very substantial proportion of the current staff (at GSFC?) will be retireable within the next 5 years.

There are an enormous number of great web sites out there posting DB data sets. For example, U. Wisconsin with a cloud mask, cloud phase, cloud top pressure and precipitable water measurements. Elsewhere there are real time observations of:

- Fire
- Coastal conditions
- Tropical cyclones
- Flooding
- Thunderstorms
- Ice conditions
- Snow cover
- Ocean gradients
- Dust
- Air Pollution
- Vegetation condition

It is like a "web-cam in the sky". Louisiana State University is combining IR channels (hotspots) with aerosol detection (smoke) to provide enhanced fire products. There is a huge collection of ocean products coming out of the University of South Florida (USF).

A great deal of earlier material on DB is at the original direct readout portal (http://rsd.gsfc.nasa.gov/goes/eosdb). This will shortly be updated to bring it into line with the new site. It contains software to get from level 0 to level 1, for decoding CCSDS packets, fire hotspot software and, increasingly, software from the EOS global processing streams. The amount of material there is testament to what a lot has been done – we should not lose sight of the fact that the DB road has been a long slog and that it is too valuable too lose.

Arthur Frederick Hasler – Video/animations

Numerous jazzy videos and simulated views.
Relevant keywords for google – keyhole earth view
http://etheater.gsfc.nasa.gov/

Patrick Coronado – Part B of EOS DB Talk

Coronado is the project manager of the Direct Readout Lab at GSFC, which has the goal of community enablement. They have made available a pair of DVDs with data sets and software necessary for DB processing. We need to register for these when we collect them outside. There are some legal complications arising from NASA distributing this software:

1. anyone packaging it and ondistributing it becomes responsible for enforcing the ITARRS restrictions (no s/w export to the bad guys)
2. NASA needs to be notified annually of the algorithm status if you've made any changes (1 line email to Pat Coronado is sufficient)
3. NASA needs to be acknowledged

Angelita Kelly – EOS Mission Operations (ESMO)

ESMO provides the DB schedules and is responsible for managing the deep space network RFI mitigation. Following testing, it was discovered that line-of-sight to DSN stations (ie horizon to horizon) exclusion of DB was not necessary. A 20 degree cone around the DSN stations is sufficient – leading to about 20 minutes of DB data loss per day, rather than several hours. There have been no terra/DSN rfi incidents since launch. A DSN waiver was successfully applied for and granted over Australia during the recent bushfire 'events'.

Mission support can be accessed via http://directreadout.gsfc.nasa.gov/mission_support/aqua_mission_support.htm. Log in as anonymous to get access to things.

Wenjian Zhang – Wide Use of EOS DB Systems Across China

The first Chinese DB station was set up in November 2000. There are now at least 25, with something like 9 around Beijing, but only two in the westernmost areas. The Chinese Weather Bureau is focussing on 4 main stations (the two in the west, one in Beijing and one in the south) and is linking them by optical fibre to the Beijing weather head office. There the data is processed and retransmitted via a communications satellite for use throughout the country. It is expected that a total turn-around time of about 30-40 minutes will be achieved. This is set to be operational by the end of 2004.

Numerous impressive examples of the use of EOS DB data in the monitoring of flooding, dust storms (from Mongolia) and cloud properties were given. The first ever complete mapping of China at 1:1 million has recently been completed using MODIS data.

The next Chinese satellite FY-3 will have an MVIRS instrument with 5 bands at 250m resolution and 15 bands at 1km. Only one IR band though.

The next EOS DB meeting will be in Beijing later in 2004.

Alexander Shumilin – R&D Centre SCANEX

These people make
- Receiving stations (hardware and software)
- Image processing software
- Image archive and distribution systems
- Research in Remote Sensing and thematic applications

They have two different X-band systems available, and have a total of 17 already deployed in Russia and greater Asia. They have done a native port of IMAPP to Windows (!) which is available for download from http://eostation.scanex.ru/software.html. There appears to be an enormous amount of processing software of one sort or another.

Bob Bernstein & John Fahle – Seaspace

The major issues that users need to overcome are
1. barriers to data access
2. barriers to processing (ie tools)

Seaspace has now deployed 27 TeraScan stations worldwide. They are providing new tools with both MODIS and AIRS processing algorithms both from U. Wisconsin and the NASA DAACS.
Upcoming from Seaspace will be:
- New NASA and IMAPP algorithms
- Better integration with the TeraScan architecture
- Faster dater delivery/processing
- Support for both research and operational users

18/11

Liam Gumley – IMAPP for MODIS and AIRS

Developed in conjunction with the University of Wisconsin DB reception system. Seaspace antenna located on top of a building in Madison. Data is just archived to exabyte – the DAAC is the real backup system. Data from the DB site is available for free over the web for up to seven days after acquisition. Lots of material, images at http://eosdb.ssec.wisc.edu/modisdirect .

IMAPP is freely available software for EOSDB.

New features in 2003:
- AIRS/AMSU/HSB Level 1 processing first release (JPL s/w, binaries only)

- MODIS Level 1 update (Geolocation and calibration)
- Updated MODIS Level 2 science products

Coming in 2004:

- AMSR-E L1 first release (with Remote Sensing Systems http://www.remss.com)
- AIRS/AMSU/HSB L2 science products (with JPL)
- Additional L2 MODIS science products

IMAPP development is funded until 2006.

AIRS/AMSU/HSB package for processing AIRS L0 to L1B. Version 1 released 5 Nov 2003. Only executable code has been released so far, for RedHat Linux and Solaris sparc architectures. The L1B output is identical to that from the DAAC (6 minute granules). No dynamic ancillary data is required. The Level 2 product code will be released in Q1 2004. AIRS has 2000+ channels.

AMSR-E: status is that RSS delivered new code to SSEC on 10 Nov 2003. No dynamic ancillary data is required and the output is a flat binary file.

New MODIS Science Products:

- Aerosol Optical Depth (MOD04)
- Land Surface Temperature (MOD09) in conjunction with WASTAC
- Sea Surface Temperature (MOD28) currently in beta (11/2003), public release (01/2004)
- Cloud optical properties (MOD06_OD)
- Snow and Sea Ice (MOD10 and MOD29)
- Scene classification

Two new utilities:

1. Thermal IR destriping
2. Guide to converting IMAPP L1B format to DAAC L1B format

The striping arises in MODIS for a number of reasons, including the two sided mirror, multiple detectors, and has a noticeable affect on the L2 science data. The solution has been to apply the destriping algorithm developed for GOES imager data based on matching empirical distribution functions. Destriping is tailored for each granule. Currently it is implemented as IDL code running on either IMAPP or DAAC data. Note that some terra detectors are actually too noisy to correct.

Also Band 26 destriping has been updated for AQUA.

Note from question at the end: MODIS has never had a destriping L1B requirement written. That was seen as a secondary goal. Contrast that with VIIRS on NPP and NPOESS which will have – ie the striping in L1B data from these will have to be below some specified threshold level.

Liam has also created a tutorial on how to produce reprojected true colour L1B images. These are interesting/useful because they:

- Remove the bowtie effect
- Allow geographic overlay

- Are pretty pictures
- Allow for colocation with other data
- Prepare the data for ingest into GIS

The method maps channels (1,4,3) to (R,G,B) and uses interpolation to convert the 500m bands 3 and 4 to the 250m resolution of band 1. Good reprojection software can be hard to find – this tutorial helps. The details are on the WWW at ftp://origin.ssec.wisc.edu/pub/IMAPP/MODIS/TrueColor

Wentz & Gentemann – AMSR-E Web Products and Real-Time Potential

Part 1. Frank Wentz
AMSR-E is carried on AQUA and the software developed by Remote Sensing Systems (http://www.remsss.com) converts the science data stream to L1B product. It needs the GBAD data (Location, Velocity, Attitude). The AMSR-E has 14 channels, or 7 bands with dual polarisation. The L1B DB software is currently installed at 5 US sites. Products include:

- Wind
- Water vapour
- Cloud
- Rain
- SST

Because it is relatively low resolution data (around 40 km) it is possible to ftp an entire pass across the Internet in around 30 seconds. The microwave algorithms can be quite complex so it is possible to send the DB data to RSS to process it and get it back again – ie more efficient to move the data than the algorithms. NOTE** Australian users of AMSR-E may wish to use this facility.

RSS have a 24-year long archive of microwave retrievals on line at their web site.

Part 2. Chelle Gentemann
RSS will distribute DB data through their web site. Generally data take about 6 hours to arrive via the standard routes, but the DB data can be made available in less than 15 minutes. There is a great deal of historical and near real time data on the RSS web site.
A lovely example of the application of microwave data was its ability to perform SST retrieval through cloud. The example shown was a cold wake caused by ocean upwelling behind hurricane Danielle. This significantly affected the course of the immediately following hurricane Bonnie but was not detected by any of the other sensors.

Allen Lunsford – Converting MODIS Institutional Algorithms to DB Environment

Direct Readout Lab has been going since the 1980s, originally developing technology for NOAA/HRPT, GOES, GMS and DMSP. In the 1990s the commercial sector got going, a large DB community developed, and the ongoing continuous GOES/POES programs enabled continuous improvement.

However EOS DB systems struggled for acceptance, and DB community pressure was essential to get it up. A consequence was that no NASA group was tasked to provide higher than L0 support for EOS-DB. This is where the DRL stepped in.

The most recent release of the Standalone DB L1B code was V2.3. The DRL has now begun to release the institutional source code because DB users wanted to develop their own (compatible) products. The institutional code was designed to run under the MODAPS environment with a whole swag of toolkits and using the whole of an orbit (not just a relatively short swath). The first priority goal for the DRL was to release the institutional source code. The secondary goal is to assist DB users to make it go.

The institutional algorithms are provided as processing code by the science teams. They execute as part of the MODAPS environment and are grouped into PGEs (Product Generation Executables), the products themselves being identified by MOD#. There are a total of 49 PGEs, spanning disciplines, modes, satellites etc. They include:

- Terra, aqua
- Swath, grid, multi-day
- Land ocean atmosphere

The PGEs require MODAPS and have all sorts of external source code dependencies. They also expect certain specific external ancillary data. The goal of the standalone DB versions is to execute without MODAPS, targeting Linux OS with a simplified execution environment. Some validation of the science quality is also desired.

Issues to be addressed include:

1. Source code dependencies
   - Use free source HDF, HDF5
   - Shared source (from the science teams)
   - SDP (PGS) toolkits + SDST and MAPI
2. Work around non-DB requirements
   - Bracketing granules
   - Multiple granules
   - Real time ancillary data sets
3. Standalone support challenges
   - Documentation for these dependencies varies widely
   - There is no common source for support or assistance
4. Execution challenges
   - Requirement for MODAPs and all that comes with it (production rules, run-time parameters and ancillary data dependencies)
   - The dreaded PCF files

PGE -> Standalone Steps

- Convert L1 code to support full pass (more than a granule)
- Coordinate the support libraries
- Process test data and validate
- Create build scripts
- Create execution scripts (PCF)
- Release with data sets (via WWW portal and/or DVD)

- Support future algorithm changes.

Progress to date has resulted in the release of 5 algorithms:
- Cloud mask (MOD35)
- Fire detection and thermal anomalies (MOD14)
- Snow cover (MOD10)
- Sea Ice (MOD29)
- Surface reflectance (MOD09)

The 5 released algorithms have executables for RedHat 9 linux, support libraries, source code, build scripts and ancillary data retrieval scripts. The MOD35 output is "DAAC Compliant".

NOTE: The ATBDs for MODIS products are **NOT** maintained. However the software and user guides are. The latter are generally a better reference for would-be porters.

The standalone NDVI/EVI port is largely complete. Given what has happened with MODIS, the DRL is planning to be well ahead of the curve for future DB missions (ie NPP, NPOESS). Note however there are no plans to do the MODIS ocean codes, partly because it is a huge task already underway in the field, and partly because they want to get on with getting ready for the next missions.

<u>Patrick Coronado – NPP Overview</u>

The NPOESS Preparatory Project is due for launch in October 2006. The DRO is ramping up support for NPP. NPP is:
- is a joint NOAA+DOD+NASA mission,
- focussed effort, cost savings
- continuity
- 5 years lifetime
- No encryption capability (unlike future NPOESS)
- VIIRS imager compression factor of 1.4-1.8 (JPG2000)
- High rate data will be on all the time (No Deep Space Net conflicts)

Data will consist of:
- Raw Data Records (RDR), 150 GB/day (similar to EOS), "L0" data
- Sensor Data Records (SDR), "L1" data
- Environmental Data Records (EDR) – available within 180mins 95% of the time.

NPP carries a subset of the full NPOESS suite of instruments and will be able to produce 27 of the full 56 NPOESS EDRs.

<u>Kelvin Brentzel – The NPP In-Situ Ground System (NISGS)</u>

NISGS is part of the NPP. NISGS directly supports the Technology Insertion and NPOESS Risk Reduction objectives:
- Provides technology insertion for NPOESS, which will assure that the NPOESS High Rate Data (HRD) design will be readily usable for Field Terminal users.

- Enables the user community to readily transition from existing POES/Terra/Aqua DB data to NPP and NPOESS DB data.

There is an emphasis on modularity to enable freedom of user choice.
NISGS is also a bridge between the mission and the DB community.
The design criteria for the Direct Readout technology suite (that will be free to "anyone") include:

- Scalability
- Extensibility
- Portability
- Ease of use

More info available in detail at the DRO portal (http://directreadout.gsfc.nasa.gov/).

Allen Huang – IMAPP L2 Algorithm Overview

Description of plans for product generation within IMAPP environment beyond L1B. Including for AIRS, MODIS and AMSU-A/AMSR-E. Brief history of Uwisc. Heritage/track record in this area and summary of existing and intended capability as per table below. An important end goal is colocation of both AIRS and MODIS data to permit synergistic processing.

<table>
<tr><th></th><th>MODIS</th><th>AIRS/AMSU</th></tr>
<tr><td>Current</td><td>Geo-location/Navigation
Cloud mask
Cloud Phase
Cloud top Property
Clear T/Q Sounding
Total Precipitable Water</td><td>Geo-location/Navigation
(5 November, 2003; 1st Version released)</td></tr>
<tr><td rowspan="2">Planned</td><td>Cloud Particle Size
Cloud Optical Thickness
Aerosol Optical Thickness
Surface Reflectance
Sea Surface Temperature
Snow Detection
Sea Ice Detection
Scene Classification (Clouds and Land Surface)</td><td>Clear/Cloudy T/Q Sounding
Cloud Detection
Cloud Clearing
Cloud Height/Emissivity
Surface Skin Temperature
Cloud Liquid Water
AMSU Precipitation estimate</td></tr>
<tr><td colspan="2">MODIS / AIRS Collocation</td></tr>
</table>

Considerable discussion of various products: AIRS profiles, MODIS cloud mask cloud phase, cloud top pressure and total precipitable water. Also MODIS SST and preliminary AMSU and AMSR-E products.

Steve Friedman – Characteristics of AIRS data for DB Users

Description of mechanism for release of DB software from JPL via U.Wisc.
Overview of instrument properties and capabilities, scan geometry etc.
Goals for processing.

Judd Taylor - The Direct Broadcast Ocean Products System at USF

Sea Space receiving systems.

Description of processing system:

- Use DAAC software and MODAPS environment on Linux Cluster
- Cluster scheduling and automation with S4P
- 10 node cluster (mix of opteron and athlons)
- GigE interconnect & IDE RAID file servers
- Locally grown mapping/projecting s/w

Applications

- Coastal water quality (estuaries)
- Connection between terrestrial and coastal ecosystems (rivers, runoff, reefs)
- Coastal carbon dynamics
- Red tide detection
- Data validation and merging
- Algorithm development

Jim Simpson – Parallel Image Processing Environment

Description of a particular parallel processing system (Linux Cluster) including general design principles. Well worth looking at PPT file – much detail on what is and what is not important.

Danielle Forsyth – MODIS Data Classification for Collaboration, Tracking, Re-Processing and Search

Interesting system for management of metadata to support data discovery and comprehensive lineage tracking and search and collaboration ….

Based on emerging W3C standards RDF & OWL (re. OGSA).

Darryl Nickless – The EOSDIS System for Obtaining Large-Area and Long-Term EOS Data Sets to Aid DB Users

MODIS DAAC data sets comprise a uniquely comprehensive record of the environment. Ie Global, full mission duration and dense. The overall quality is largely unsurpassed.

The GSFC DAAC supports ftp pull (preferred), ftp push, DVD, 8mm & DLT tape.

The data pool, a 50TB online collection, supports a web hierarchical ordering mechanism. http://daac.gsfc.nasa.gov/data. It is at the Goddard Earth Sciences DAAC only. Is fast and easy and Googleable at the dataset level (add "Kempler" to your search to keep to the DAAC pages).

There are currently 2 Petabytes in the GES DAAC, and they distribute about 2 TB per day.

Jacques Descloitres – The MODIS Rapid Response NRT Processing for Fire Monitoring and other Applications

System arose because EOSDIS "suffered contingencies" and Fires in 2000 needed attention – urgently. The system is low cost and low maintenance (Linux/Raid, fully automated). It relies on the spacecraft-broadcast ephemeris (ok for terra, 50-400m for AQUA).

There is a corrected reflectance product:

- Simple atmos. corr. With vis. And NIR bands (1-7)
- Corr. for molecular (Rayleigh) scattering & gas absorption (Water Vapour, Ozone)
- No RT input or ancillary data
- Climatology for gases
- No aerosol correction.

The fire detection (MOD14) uses a contextual algorithm (Giglio et al 2003).

Vegetation indices:

- NDVI and EVI are computed from the corrected reflectance product
- The pros and cons of that product therefore propagate.
- Computed for swath data – not composited or mosaicked
- No correction for directional effects (ie BRDF)

The L0 data feed from terra and aqua brings about 150GB/day.