MODIS Terrestrial Products at the Land Processes Data Center: A Decade of Archiving and Distributing Valuable Land Data to the Global Community

Bhaskar Ramachandran[1] and David Meyer[2]

[1]EOS Land Processes DAAC, ADNET Systems, Inc., contractor to the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD, USA. Work performed under USGS contract 08HQCN0005. [2]USGS/EROS, Sioux Falls, SD, USA.

I. Introduction

The Earth Resources Observation and Science (EROS) Center, established in the early 1970s, is a USGS data management and distribution, and research facility for remotely sensed data as well as other forms of geographic information. EROS holds the world's largest collection of civilian remotely sensed data covering the Earth's land surface. The EROS archive contains over 2 Petabytes of digital data, which serves to develop and promote remote sensing data-based science applications to identify, monitor, and understand varying spatial dimensions of change in the Earth's terrestrial landscape.

In 1990, the Land Processes Distributed Active Archive Center (LP DAAC) was established at EROS as part of NASA's EOSDIS initiative to process, archive, and distribute land-related data collected by the EOS MODIS and ASTER instruments. These activities help promote the inter-disciplinary study and understanding of the integrated Earth system. LP DAAC's role includes ingest, archival, and higher-level processing and distribution of ASTER data products, and ingest, archival, and distribution of MODIS land products derived from both the Terra and Aqua platforms. LP DAAC is also responsible for implementing MODIS product quality updates to the science metadata flags based on quality determinations made by the MODIS Land Data Operational Production Evaluation (LDOPE) facility.

II. History of MODIS land products at the Land Processes Data Center

Originally geared to handle all aspects of a MODIS data production system, the LP DAAC assumed MODIS data ingest, archiving and distribution functions after the EOSDIS plans to process such data at the DAACs were descoped in the late 1990s. This led to the creation of the Science Investigator-led Processing Systems (SIPS), which in MODIS' case is located at the MODIS Adaptive Processing System (MODAPS) facility at the Goddard Space Flight Center (GSFC).

MODIS data from the Terra and Aqua satellite platforms follow the same path today as they did in the beginning. Raw instrument data are downlinked to the White Sands ground segment of the Tracking Data Relay Satellite System (TDRSS), transferred through high-speed networks to the EOS Data Operations System (EDOS) at the GSFC for level-0 processing, and then transmitted to MODAPS/GSFC, where the higher-level land data products are produced and electronically transmitted to the LP DAAC (Figure 1). Following standard communication protocols ensure that the data are successfully transferred from MODAPS to the LP DAAC archive.
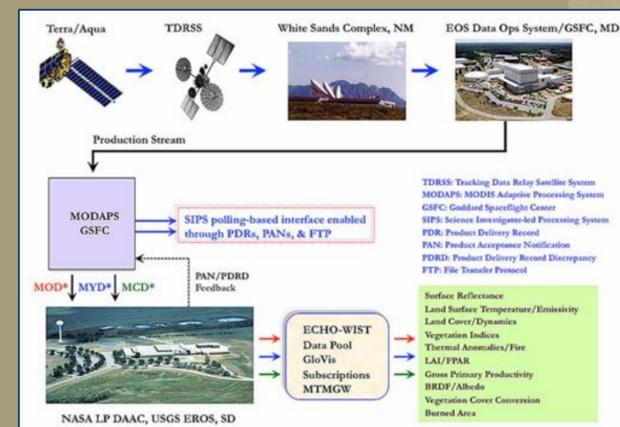

Figure 1.

III. Generations of MODIS land data products: 4 Public Versions and Continuing…

LP DAAC has hosted four public versions (V1, V3, V4, and V5) of MODIS land products in the last decade. Each preceding version is publicly available for a certain time period before it is decommissioned (Figure 2). Exceptions were made in V4 where, certain product collections (land surface temperature and emissivity, land cover, the Climate Modeling Grid, etc.) are still publicly maintained. Another development is the implementation of an interim version (V41) of land surface temperature, which is produced with V4 algorithms and V5 input data. This trend is likely to continue with certain products in the V5 era as well. All components of the MODIS mission (Science and instrument teams, algorithm developers, MODAPS, and the DAACs) continue to work towards the V6 era, slated to begin in 2010.
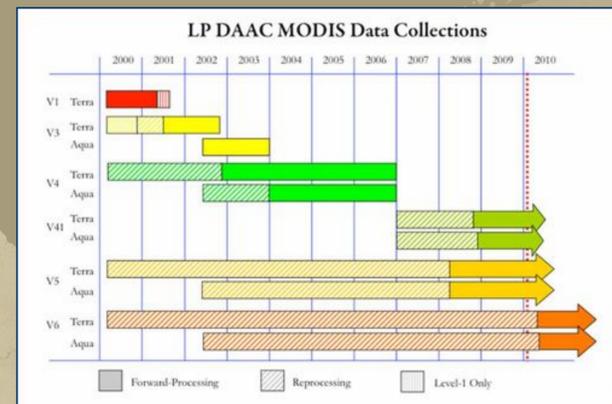

Figure 2.

IV. The distribution big picture: Granules and Gigabytes

In the last decade, in the Terra world, thermal anomalies and fire products were distributed the most (> 18 million), followed by the land surface reflectance products (~14 million), land surface temperature and emissivity (LSTE) products (~12 million), and equal amounts of vegetation and leaf-area index/fraction of photosynthetically active radiation (LAI/FPAR) products (~12 million each) (Figure 3). Between 2 and 2.5 million files were distributed each for the Gross Primary Productivity, and BRDF/Albedo suites. In terms of data volume, the land surface reflectance, and the vegetation indices suites command the bulk (~1.5 PB).

Among Aqua products, the patterns are similar albeit with fewer numbers due simply to a shorter mission history (Figure 4). Thermal anomalies and fire products account for > 11 million files, followed by LSTE products (< 5 million), land surface reflectance products (< 3 million), LAI/FPAR (1.6 million), and vegetation indices (1.3 million). The land surface reflectance and vegetation indices suites lead in terms of data volume. Among the combined products, which use Terra and Aqua data inputs, the BRDF/Albedo suite leads the way with nearly 4.5 million files followed by LAI/FPAR with under a million files. The burned area, and land cover suite products are relatively new (released in October and December 2008, respectively) to garner higher numbers.
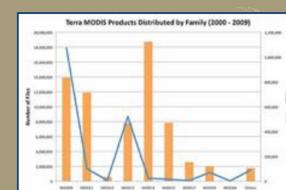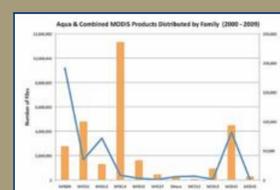

Figure 3.

Figure 4.

V. The Top-ten MODIS products distributed in the last decade:

Rank	Terra MODIS	Aqua MODIS	Combined MODIS
1	MOD14	MYD14	MCD43B3
2	MOD15A2	MYD11A1	MCD15A2
3	MOD11A2	MYD15A2	MCD43A4
4	MOD11A1	MYD11A2	MCD43A2
5	MOD13A2	MYD09GA	MCD43B2
6	MOD09A1	MYD11B1	MCD43B4
7	MOD17A2	MYD14A1	MCD45A1
8	MOD13Q1	MYD13A2	MCD43A1
9	MOD09GHK	MYD11_L2	MCD43A3
10	MOD09GA	MYD09GQ	MCD43B1
	>50 million files/1.11 PB vol.	<20 million files/141 TB vol.	<6 million files/70 TB vol.

VI. From Off-line to On-line archives

Two Petabytes of MODIS data were ingested over the last decade as the MODIS mission experienced four generations of data versions (Figure 5). The year 2008 saw the beginning of the evolution of the on-line and off-line archives at LP DAAC. The on-line archive is a complement of spinning-disk hardware known as the "Data Pool," which is universally accessible to all data users. The off-line archives are a collection of Linear Tape-Open (LTO) magnetic storage tapes that provide the data backup. What began as the limited Data Pool has evolved to become a direct line of access to all public collections of MODIS land products that users may download at any time of their choosing.

VII. Data discovery interfaces

MODIS land products from the LP DAAC are available for search and discovery through the following interfaces:
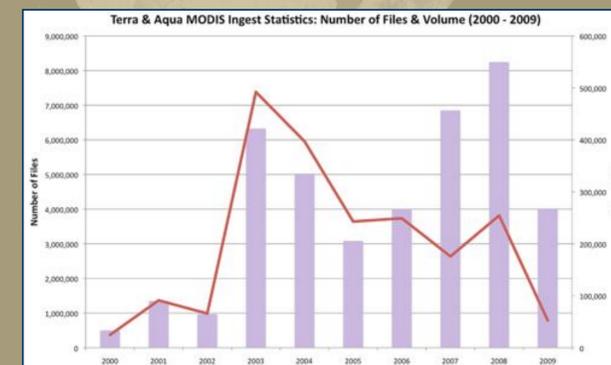

Figure 5.

- ECHO-WIST (https://wist.echo.nasa.gov/~wist/api/imswelcome/)

- Data Pool (http://lpdaac.usgs.gov/datapool/datapool.asp)

- Via subscriptions (contact lpdaac@eos.nasa.gov), and

- Machine-to-Machine Gateway (contact lpdaac@eos.nasa.gov)

Besides providing full access through standard EOS interfaces, LP DAAC has also developed site-specific tools to serve the land remote sensing community. They include the following interfaces:

USGS GloVis (http://glovis.usgs.gov/)

The U.S. Geological Survey (USGS) Global Visualization (GloVis) tool supports satellite-derived earth science data at EROS. It provides a browse image-based product selection process, whose simplicity makes it popular in the user community. The GloVis inventory includes MODIS land products, and provides an alternative search and order method (Figure 6).

MRTWeb (http://lpdaac.usgs.gov/datapool/datapool.asp)

The MRTWeb interface combines browse-based data discovery with the processing options offered in the MODIS Reprojection Tool (MRT). Using the GloVis model, users can select tiles of interest from a composite image of relevant browse files, and also request additional processing including the following functions: spatial and spectral subsetting, mosaicking, reprojection, and reformatting. All MODIS land products populated in the LP DAAC Data Pool are potentially available through this interface (Figure 7).
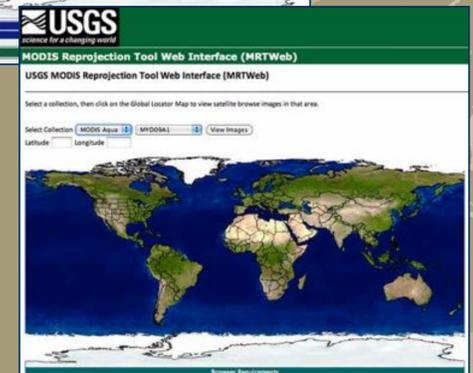

Figure 6.

Figure 7.

VIII. The future

Following 10 years of serving the EOS Terra and Aqua MODIS missions, the LP DAAC continues to evolve as it performs diverse data management activities. They include data ingest, archival, distribution, and QA metadata updates. Version management remains an active role as MODIS data itself continues to evolve, and a new data version is spawned roughly every two years. Advances in hardware and software engineering have also directed the larger evolution of the EOSDIS Core System that has helped LP DAAC become more efficient with a smaller footprint.

How users experience and interact with our data is a vital area that is constantly receiving more attention. Web-enabled technologies hold much promise in this area to facilitate both data discovery and on-demand processing tailored to user specifications. Developments in the eMODIS processing system provide a window into how such a capability can benefit applied science needs. Similarly, developments in service-oriented architecture provide packaged functionality as interoperable Web services across different organizations. Advances in Web technology provide an ideal framework to develop an e-Science infrastructure to support LP DAAC's varied science data management requirements in the future.