MODIS Land Data Products: Generation, Quality Assurance and Validation

Edward Masuoka[1], David Roy[2], Robert Wolfe[1], Jeffery Morisette[1], Scott Sinno[1], Michael Teague[1], Nazmi Saleous[3], Sadashiva Devadiga[1], Christopher Justice[4] and Jaime Nickeson[1]

[1]Terrestrial Information Systems Branch at NASA's Goddard Space Flight Center
[2]Geographic Information Science Center of Excellence, South Dakota State University
[3]Department of Geography, United Arab Emirates University
[4]Department of Geography, University of Maryland, College Park

1. Introduction

The Moderate Resolution Imaging Spectroradiometer (MODIS) on-board NASA's Earth Observing System (EOS) Terra and Aqua satellites are key instruments for providing data on global land, atmosphere, and ocean dynamics (Salomonson et al. 1989). MODIS senses all of the earth's surface on a near daily basis in 36 spectral bands, spanning the visible (0.415µm) to infrared (14.235µm) spectrum with nadir pixel dimensions of 1km, 500m and 250m. Derived MODIS land, atmosphere and ocean products are central to NASA's mission to monitor and understand the Earth System. NASA has developed and generated on a systematic basis a suite of MODIS products starting with the first Terra MODIS data sensed February 22, 2000 and continuing with the first MODIS-Aqua data sensed July 2, 2002.

This chapter describes the production and distribution of the MODIS Land products, from initial software delivery by the MODIS Land Science Team, to operational product generation and quality assurance, delivery to EOS archival and distribution centers, and product accuracy assessment and validation. Progress and lessons learned since the first MODIS data were available in early 2000 are described.

The MODIS Land Science Team is funded by NASA to develop and maintain the science algorithms and processing software used to generate the MODIS products described in this book. The Science Team is responsible for coordinating, developing and undertaking protocols to evaluate product performance, both on a systematic basis through quality assessment activities, and on a periodic basis through validation campaigns. The Science Team was formed in 1989, 10 years before the launch of EOS/Terra. In this period the team provided feedback on the design and pre-launch testing of the MODIS instruments and the development of elements of the Earth Observing System Data and Information System (EOSDIS) and its infrastructure, the EOSDIS

Core System (ECS) (Asrar and Ramapriyan, 1995) which are still used to distribute the MODIS land products (Running et al., 1994). NASA continues to fund the science team to undertake product development but the emphasis has shifted to product maintenance and systematic quality assessment.

2. Land Products

The MODIS Land products are divided into three product suites: radiation budget products, ecosystem products, and land cover characterization products (Justice et al. 1998). The radiation budget products are used for developing a physical assessment of land surface processes through a better understanding of the surface/atmosphere energy exchange. They are important, for example, in studies of the hydrological cycle, biological productivity, and climate variability. The ecosystem products measure spatial and temporal dynamics of the Earth's terrestrial vegetation and are inputs into global productivity and biogeochemical modeling. They are important, for example, in studies of ecosystem functioning and characterization of seasonal vegetation productivity. The land cover characteristics variables measure both anthropogenic and natural changes in the terrestrial landscape and are used to help understand the causes and impacts of the change.

All of the land products are produced on a global basis. They have different spatial and temporal resolutions constrained by the resolution of the MODIS instrument and by data storage limitations. The MODIS land product format was developed in an attempt to satisfy the diverse needs of the user community and the MODIS land product generation algorithms. The format is needed to enable efficient processing and reprocessing of MODIS land products, and to support flexible subsequent application of the data. The MODIS products are produced in a hierarchy of four processing levels: Level 1 (L1) instrument data at full resolution, Level 2 (L2) derived geophysi-

cal parameters at the same resolution and location as the L1 products, Level 2G (L2G) and Level 3 (L3) earth-gridded geophysical parameters and Level 4 (L4) modeled outputs in an earth-grid.

All MODIS products are stored in an enhanced Hierarchical Data Format, known as HDF-EOS, composed of multidimensional data arrays and descriptive metadata. Special HDF structures for the L2 and L3 products were mandated by NASA and supported by NCSA to ensure a level of compatibility between products from different instruments on EOS satellites. The metadata conform to an ECS data model and are used in searching the production system archives, and may be used by the user community when they search EOS data center archives.

The smallest unit of MODIS data processed at any one time at L1 and L2 is a granule. A granule corresponds to 5 minutes of MODIS swath data that covers approximately 2340 km x 2030 km in the across- and along-track directions, respectively. The 5 minute granule size was chosen to limit the data volume per file to less than 1GB and to enable distributed processing across a large number of processors. For the gridded products at Levels 2G, 3, and 4, the smallest unit of processing is a tile. Each tile corresponds to a 1200 km x 1200 km area defined in a global non-overlapping grid in the Sinusoidal projection. The snow and sea ice products are defined in the Lambert Azimuth Equal Area projection. The L2 products sensed over each day or night are binned into an intermediate tiled data format referred to as L2G that contains the original sensed MODIS observations and their sub-pixel geolocation information (Wolfe et al. 1998). Subsequently, in producing the L3 and L4 products multiple observations for each grid cell are combined using criteria specific to each product.

The MODIS Land products are produced by the MODIS Adaptive Processing System (MODAPS) at NASA's Goddard Space Flight Center. The products are delivered via the internet to EOS data centers, called Distributed Active Archive Centers (DAACs), where the products are

archived and distributed to the user community. Cryospheric products are sent to the National Snow and Ice Data Center DAAC at the University of Colorado at Boulder, and all other land products are sent to the Land Processes DAAC at the USGS Center for Earth Resources Observation and Science (EROS).

Table 2.1 lists the MODIS Land products and gives some details on the products: the level(s), where the products are archived and distributed, spatial resolution(s) and temporal resolution(s).

3. MODIS Data Production

The initial design of the EOSDIS in 1995 envisioned a geographically distributed production of the MODIS Land products among three DAACs: GSFC Earth Sciences (GES) DAAC in Greenbelt, Maryland for L1 and L2 products, the National Snow and Ice Data Center (NSIDC) DAAC in Boulder, Colorado for the snow and sea ice products, and the Land Processes (LP) DAAC at the USGS Center for EROS in Sioux Falls, South Dakota for all the other Land products (Justice et al. 2002b). Delays in the development of the EOSDIS led to an alternate approach to science product generation in 1999. Responsibility for producing L2 and L3 science products was shifted from the DAACs to the MODIS science team. The science team worked closely with the Science Data Support Team (SDST) to develop MODAPS. Science team control of the processing system offered greater flexibility in terms of what new features would be incorporated into the system and reduced the time to get software changes into the production from 1-2 weeks to 1-2 days. Archiving and distribution of science products to the public has remained the responsibility of the three DAACs. However in 2006, the responsibility for archiving and distributing MODIS L1, Atmosphere and Oceans products were transferred from the GES DAAC to MODAPS and the Ocean Color Data Processing System to bring support for these products closer to their science communities. At present, the MODAPS system distributes land products

to the MODIS Land science team members, primarily for quality assessment and to a limited number of science and applications users while the LP and NSIDC DAACs distribute land products to the general community.

3.1 Data Flows

Figure 3.1 summarizes Terra MODIS Land data flows from the spacecraft to the end-users. The MODIS Aqua data flow is identical to the Terra flow illustrated in Figure 3.1 except that the instrument data are sent to ground receiving stations in Alaska and Norway rather then White Sands, New Mexico. The EOS Data and Operations System (EDOS) separates the MODIS instrument data from other spacecraft data streams, orders the data into chronological sequence, and checks for missing or bad data packets, before aggregating the data into 2 hour L0 files. Each 2 hour file is equivalent to the data sensed by MODIS orbiting the Earth approximately 1.2 times. Under normal circumstances, most of the MODIS instrument data are delivered to the MODAPS within 8 hours of acquisition; missing data may take several days to acquire.

Figure 3.2 illustrates the sequence of MODAPS processing required to produce the MODIS Land products. Ancillary data that include National Center for Environmental Prediction (NCEP) meteorological data and Global Modeling and Assimilation Office (GMAO) data are delivered to MODAPS from the GES DAAC as soon as they are acquired. These ancillary data characterize the atmosphere and are primarily used to generate the MODIS surface reflectance L2 product. Ancillary data delivery lags behind real-time by approximately 24 hours for NCEP predictions and by up to 2 weeks for certain GMAO data. For MODIS, spacecraft attitude data are measured in real-time using on-board gyros and star trackers and are available for immediate use in the data processing system. Spacecraft ephemeris data (orbit position and velocity) for MODIS on Terra are generated from an on-board navigation system that provides accurate real-time measurements of the ephemeris, which is also available for immediate use. However, for MODIS on Aqua, only predicted ephemeris with lower accuracy is available in real-time. Users must

wait up to 24 hours to receive ground-based definitive ephemeris to achieve sub-pixel geolocation accuracies.

Once L0 files arrive at the MODAPS, processing may be initiated for that current day. Algorithms that can be grouped together in a processing chain are collected into recipes denoted by a letter and number. The MODAPS scheduler launches an instance of a recipe as a job that runs on an individual processor and waits until the recipe completes and output products are written to disk before launching the next job on the processor. Over 900 processing jobs can be run concurrently.

The L1 products (Level 1B calibrated radiances, Geolocation fields, Cloud Mask and Atmospheric Profiles) are generated first and drive the higher order production streams. Due to differences in input requirements, the L2 processing is split among three different recipes. The first recipe (R1) creates the snow, sea ice and fire products. It runs independently on each 5 minute MODIS granule. The second recipe (R3) is used to create the L2 surface reflectance product. It requires one full orbit of data as input and generates 19 to 20 output granules. This recipe runs up to 15 times per day and requires all granules from one orbit to be available before starting. The third recipe (R4) is used to create the daily Land surface temperature products. This recipe divides the globe into six latitudinal zones and runs one processing stream per zone.

L2 data created by the R1 and R3 recipes are binned into the intermediate daily L2G data products. The L3 daily products are produced from these by selecting the best observation or combination of observations in each bin, based on algorithms unique to each L3 product. These operations are performed in the R5 and R5P recipes for equatorial Sinusoidal (most land products) and polar Lambert Azimuth Equal Area (LAEA) projections (snow and sea ice), respectively. The L2G and L3 daily products are temporally composited to create the 8-day products for Land surface reflectance, Land surface temperature, snow, fire, and LAI/FPAR (recipe R10),

and the 16-day products for BRDF/Albedo, VI and intermediate Land cover conversion products (recipe R12), and 32-day vegetation and intermediate Land cover products (recipe R14). Finally, the 32-day products are used to create the NPP/PSN and yearly Land cover products (recipe R18).

The processing of a current day for MODIS, proceeds at the rate of 2 data-days per day or 2x (one Aqua MODIS data-day and one Terra MODIS data-day) during which 34,000 files are staged as input to downstream algorithms or written as output to data archives at the DAACs or in MODAPS. MODAPS has additional capacity to reprocess L1 and Land products at the rate of 14x as new science algorithms become available and improvements are deemed to be significant enough to warrant a full reprocessing of the data record. Table 3.1 summarizes product volume shipped to the DAACs for each 1x of MODIS production. It illustrates the difference in volume of MODIS products in Collection 4, which were not compressed before distribution, and the new MODIS Collection 5 products, which use internal HDF compression to facilitate data transfer and to save storage on end-user systems.

3.2 The MODIS Adaptive Processing System (MODAPS)

The MODAPS was developed in 1999 to produce global science products from the MODIS instrument on EOS Terra and Aqua. Figure 3.3 illustrates the hardware components of the MODAPS production system.

The current MODAPS system runs on 3 processing systems which handle processing, reprocessing and large-scale testing of science software changes. Each system is based on a Silicon Graphics Incorporated (SGI) Origin 300 or 3000 with 100TB of disk and 60 to 100 Intel Xeon-based servers attached through a high performance network switch. The Intel Xeon-based servers

offer higher performance at a lower cost of ownership than SGI Origin servers though not all MODIS science software has been ported to run on them. Each production recipe has a dedicated database server which stores information about the processes running in the system, the location of all the files on the production disks and relevant metadata associated with each science data product file. Each of the three production strings is capable of generating in a single day the equivalent of 3-4 days of MODIS Land products.

MODAPS uses File Transfer Protocol (FTP) to ship land products to the DAACs using an automated process developed in concert with the DAACs to ensure that products are re-transmitted in the event of errors. In addition, the MODIS land products are made available to the science team and a limited number of science and applications users via FTP servers with a dedicated Web interface that supports searching and ordering against product temporal and spatial attributes and metadata values. The L0 and calibrated L1 MODIS Atmosphere products produced by the MODAPS are also made available to the public via a similar dedicated interface. These MODAPS distribution systems are efficient, as the distributed products are stored on online RAID disk, rather than on slower tape archives. They are also robust, as the web and FTP servers run on redundant Dell servers (currently 4 dual-processor servers) to ensure highly responsive performance without downtime. Products are stored in the archive in online Serial Advanced Technology Attachment (SATA) disk units. Archive storage includes 400TB for holding L0 products, and 300TB to hold L2 and L3 Atmosphere and Land products. L1 products are too large to store for the life of the mission and once produced the products are placed on a 100TB disk-based data pool, from which files are deleted after 30 days, with the exception of certain frequently ordered areas such as North America and Western Europe. If a L1 product is ordered after it has been deleted from the data pool, a processing job is automatically initiated to produce

it. During the ordering process, users can also request custom processing of MODIS standard products from the archive to better suit their needs. Processing options include subsetting and sub-sampling products, selecting parameters from a multi-parameter product, changing a product's map projection, changing a product's format (from HDF to GeoTIFF for instance), creating a mosaic of adjacent products and masking areas of a product based on political or geographic boundaries.

3.3 Software Integration and Testing

The MODIS Land Science Team develop and maintain the science algorithms and processing software used to generate the MODIS products described in this book. Science team developed software is delivered to the SDST at Goddard Space Flight Center. The SDST members are responsible for transforming the software into MODAPS production ready executables and for testing the production code in the MODAPS processing environment to ensure that it runs efficiently and correctly. The test process involves three levels of testing: unit tests, chain tests, and science tests.

The science team members provide input and output products with their software that SDST staff use to ensure that the software is producing expected results. When this occurs, the production-ready software passes unit testing. In the following chain test, the software replaces its predecessor in the production recipe, which is run to see if all processes in the recipe still produce correct output products and run without errors. Once the chain testing is completed, science tests are performed by the Science Team and by the Land Data Operational Product Evaluation (LDOPE) team to ensure that the software update has met its desired goals and that dependent product quality is maintained or improved by the code update.

3.4 Algorithm Improvements

The MODIS processing priorities have evolved since MODIS first started acquiring Earth view data. The early period of MODIS was devoted to addressing problems in the low level algorithms and ensuring proper production of higher order products. The extensive product quality assessment activities in the months that followed the initial 'first light' MODIS data led to a large number of algorithm fixes and code changes. A streamlined approach to promote code changes into operational production was adopted to accommodate frequent updates and to ensure a rapid transition to high quality products. As the algorithms matured, the processing priority shifted from rapid bug fixes and improvements to the creation of consistent data sets. As a result of this change, the approach to promote algorithm updates evolved to a more structured procedure. Figure 3.4 shows the sequence of events used in the current approach.

As the products became more mature, the extent of the science test process increased. These tests now are extensive, involving production of two 16-day periods globally and/or an entire year of MODIS products over 33 representative tiles. If no problems are identified in the science test, the science team members prepare Product Change Requests (PCRs) for their software that document the rationale for the change and the results from science testing. Each PCR must be approved by the Land science discipline leader and the MODIS science team leader before the software is promoted to operational status and used to generate science products.

4. MODIS Reprocessing - Collections

As with most science quality satellite data sets, the MODIS Land products will be reprocessed several times. Reprocessing involves applying the latest version of the science algorithm to the MODIS instrument data and using the best available calibration and geolocation information. A

collection number is included in the filename for all products to differentiate between different reprocessing runs.

The MODIS/Terra Land products generated in the months following first light were designated as "beta products" and not appropriate for scientific publication, as many software and algorithm issues were being resolved. These products were archived in the different DAACs as Collection 1 and were released to allow the users to gain familiarity with the new data formats and parameters (see Figure 4.1). Within a year after the MODIS/Terra launch most of the Land products matured to "provisional" status. These provisional products had a quality that was sufficient for use by the general research community, but users were urged to contact the science team before using the data in scientific publications. Although the quality of these products was not optimal, the improvement over the beta data set was substantial and a reprocessing of the MODIS/Terra record began in July 2001. The provisional products were archived and distributed from the DAACs as Collection 3. Further improvements in the algorithms and an on-going validation effort led many MODIS Land products to achieve a Validated Stage 1 maturity in November 2002. These science products had uncertainties defined for a few sites and a limited range of conditions and were ready to be used in scientific publications. The MODIS/Aqua and combined MODIS Terra and Aqua products were not released until they reached provisional status. To bring the whole record of MODIS data to the Validated Stage 1 level, a reprocessing of Terra and Aqua data started in December 2002 and July 2003, respectively. Products generated in this activity are archived and distributed from DAACs as Collection 4. More extensive validation activities have brought most of the MODIS Land products to Validated Stage 2, where the product accuracy has been assessed over a distributed set of locations representative of the range of conditions encountered in the data. An improved data set is being created as part of a Collection 5 forward and re-

processing effort that started in July 2006 and will include new products such as Burned Area (Justice et al. – in this volume).

5. Quality Assessment

Quality Assessment (QA) is a necessary part of any systematic production system and is integral to the MODIS Land production process (Roy *et al*. 2002). The purpose of MODIS Land QA is to provide the user community with an indication of product quality and to provide feedback to the science team developers. MODIS Land QA is undertaken by the science team and by members of the LDOPE facility. The success of the QA process is attested to by the incremental refinements to the science algorithms and production code over the four reprocessed MODIS Land product collections (Figure 4.1).

5.1 Rationale for Quality Assessment

It is generally neither desirable nor practical to delay satellite product distribution until products are proven error-free or until known errors have been removed by product reprocessing. This is because errors may be introduced at any time during the life of the sensing system and may not be identified for a considerable period, and because certain errors may be negated by appropriate user action and/or may have benign impacts for certain applications. In addition, the user community plays an important role in assessing product quality. For example, the NASA TOMS satellite data failed to reveal the Antarctic ozone hole because abnormally low ozone values were flagged as bad data. The ozone hole was instead discovered by the British Antarctic Survey using Dobson ozone spectrophotometer data. Subsequent reanalysis of 12 years of TOMS data verified the existence of the ozone hole and indicated that it had grown rapidly for that period over most of Antarctica (Farman *et al.* 1985).

The MODAPS production emphasis is on maximizing production efficiency while maintaining the integrity of the data production, transfer, archival and retrieval processes. Product errors may be introduced at any time, however, due to numerous, sometimes interrelated, causes that include: instrument errors; incomplete transmission of instrument and ephemeris data from the satellite to ground stations; incomplete instrument characterization and calibration knowledge; geolocation uncertainties; use of inaccurate ancillary data sets; software coding errors; software configuration failures (whereby inter-dependent products are made with mismatched data formats or scientific content); and algorithm sensitivity to these errors and to un-modeled surface, atmospheric and remote sensing variations.

It is critical that satellite products are distributed with supporting product performance information. Product performance information is required by users in order to consider products in their appropriate scientific context, and is required by algorithm developers to identify products that are performing poorly so that needed improvements may be implemented.

The MODIS Land science team has coordinated and developed protocols to evaluate the performance of their products through quality assessment and validation activities. MODIS Land validation is described in Section 6. Validation and quality assessment results fulfill quite different roles (Roy *et al*. 2002). Users may consider product validation results with respect to the general accuracy requirements of their application, but, unlike quality assessment results, validation results do not describe artifacts and issues that may reduce the accuracy of individual pixels or files.

5.2 MODIS Land Quality Assessment Roles

The MODIS Land QA process is complex, time consuming, and challenging to manage. This is because of the large number and volume of data products and the complex product inter-

dependencies (Figure 3.2). A centralized QA facility, the LDOPE, was formed in 1998 prior to MODIS Terra launch to coordinate the MODIS Land QA process. The LDOPE is staffed by a small group of scientific staff and is collocated with the MODAPS to enable efficient communication with the production managers and to ensure rapid data access. The LDOPE personnel undertake routine QA of all the MODIS Land products, track the quality of input products that are not the responsibility of the Land science team (e.g., the MODIS calibrated radiances and cloud mask products), check for the propagation of errors through interdependent products, develop and maintain QA tools, web sites, and procedures, disseminate QA results and information within the science team, and ensure that the QA results are available to the public.

The science team product developers undertake detailed QA of their own products and collaborate with the LDOPE and other members of the science team responsible for input (upstream) and output (downstream) products. As needed, the science team developers update production codes and/or science algorithms to rectify issues that have been found. Code updates frequently necessitate subsequent patches to fix bugs introduced by the update. This is a well known software life cycle phenomenon, but is complex to diagnose and rectify across the interdependent MODIS Land products. For example, bugs and QA issues found in the cloud mask product (Platnick *et al*. 2003) are passed through the surface reflectance products (Vermote *et al.* 2002) which in turn propagate into the vegetation index (Huete *et al.* 2002), LAI/FPAR (Myneni *et al*. 2002) and BRDF/Albedo (Schaaf *et al.* 2002) products. The number of code updates was highest in the first 12 months following launch, when for example, the MODIS Terra Land surface reflectance code was updated 14 times. The code updates are otherwise generally high in the first several months when the data collection reprocessing is initialized. These are periods of in-

tensive science team and LDOPE QA activities. Outside of these periods, systematic QA is undertaken primarily by the LDOPE personnel.

5.3 Product Quality Documentation

MODIS Land product quality information is generated automatically by the algorithm code and is stored in the products as binary encoded pixel data and as summary numerical and textual file metadata. This documentation enables users to consult QA results when ordering and using products. In many cases, MODIS Land products can only be used meaningfully after consideration of this information.

MODIS Land per-pixel QA results are generated by the production code for specific science data sets and are typically stored as separate datasets in each product file. The per-pixel QA information varies among the MODIS Land products. Products that can have meaningful error estimates assigned to them store per-pixel uncertainty estimates and/or ranges. Information on external factors known to affect product quality and consistency is also stored for each product. This information includes atmospheric conditions, surface type, viewing and solar geometry, and whether dynamic ancillary data or backup estimates have been used as input. Additional metadata fields that may also be stored include: the science code processing history (such as the logical criteria used by the algorithm), the results of different algorithm tests, and whether the input data were useful. To enable consistent interpretation across all the Collection 5 MODIS Land products, one generic QA bit is used to identify good MODIS Land quality pixels (Table 5.1). In previous collections, two generic QA bits were used, with the additional bit describing if a pixel was not produced due to cloud or other effects (Roy *et al.* 2002). This additional bit was dropped in Collection 5, reflecting the evolution of certain products to attempt to retrieve pixel values regardless of the cloud conditions.

File level QA metadata are stored in the MODIS Land products and include product-specific metadata in addition to generic metadata that describe temporal and geographic attributes, version information, filenames of the input data used to generate the product, and file level summaries of per-pixel QA information. All MODIS Land and NASA EOS products carry metadata summarizing the results of QA procedures performed by the production code and performed after product generation. These metadata were mandated by the ECS (Lutz *et al*., 2000). Of these, the Science Quality Flag and Science Quality Flag Explanation metadata may be set based on the outcome of QA performed at any time after product generation. The premise for these Science Quality metadata was that an ECS DAAC could update them at the behest of the science team and that users would be encouraged to inspect them when ordering products via ECS DAAC ordering systems. Because there can be up to three months lag between data production and the DAAC update to the Science Quality metadata, starting with Collection 3, all MODIS land products are generated with a default Science Quality Flag of "Not Investigated" including a Science Quality Flag Explanation text directing the interested user to a publicly accessible LDOPE web site for recent quality updates. The LDOPE web site documents product quality more comprehensively than is possible in the Science Quality metadata, although the LDOPE continues to update and make available the Science Quality Flag metadata on its web sites. The Science Quality Flag may be set to seven valid states ("Passed", "Failed", "Suspect", "Inferred Passed", "Inferred Failed", "Being Investigated", "Not Investigated"). Products that have been labeled as "Failed" should not be ordered. After the products are obtained users should examine the per-pixel QA data to filter pixels that are unsuitable for their applications.

5.4 LDOPE Web Site

The LDOPE web site at http://Landweb.nascom.nasa.gov/QA_WWW/ documents the quality of all the MODIS Land products, with the exception of the MODIS sea ice products, and provides a number of web enabled services that facilitate product visualization and quality assessment. The web enabled services evolved from the need to routinely sample the MODIS Land product stream. Standard sampling approaches used in manufacturing would not be efficient given the high average daily MODIS Land production volume and the non-trivial time required to quality assess a single product file. Instead, a number of sampling strategies and web enabled tools were developed to allow representative product samples to be selected for detailed QA and to enable synoptic quality assessment. In 2005 the web site was visited by an average of 150,000 users per month.

5.4.1 Known Issues

Issues that affect product quality are posted on a Known Issues Web site with example images, algorithm version, and occurrence information. Issues are categorized as ''pending,'' ''closed,'' ''reopened,'' or ''note,'' and are updated as software updates that fix the problem are promoted to the production system. This has proven to be an effective way to document the propagation of issues through the dependent hierarchy of MODIS Land products and to passively communicate product quality information within the science team and to the user community.

Figure 5.1 illustrates an issue posted to communicate the deleterious impact of a MODIS satellite maneuver on the gridded MODIS Land products over a forty day period. This issue occurred because the MODIS L1B and geolocation products did not carry the satellite maneuver in-

formation and so were incorrectly retained in the production system. As described, the impacted MODIS Land products were later reprocessed to remove this issue.

5.4.2 Global Browse

Coarse spatial resolution versions of the MODIS Land products are generated in the MODAPS by aggregation or sub-sampling to 5km. These are then projected into global images and posted, as they are generated, on a Global Browse web site in JPEG format with fixed contrast stretching and color look-up tables to enable consistent temporal comparison. Users can pan and zoom and interactively identify data granules by mouse clicks, allowing for synoptic visual inspection and identification of problematic data files. Figure 5.2 illustrates a Collection 4 Land surface reflectance (Vermote *et al.* 2002) global browse image.

5.4.3 Metadata Database

The metadata of every MODIS Land product generated in the MODAPS is archived in a database. The database has a web interface that allows the LDOPE personnel and science team to retrieve any product file metadata by querying against temporal, spatial, and metadata attributes of interest. In the early MODIS Land collections the potential of this system was reduced due to difficulties in maintaining consistent metadata definitions (format and meaning) while the product contents evolved. The database is most frequently used to track product inter-dependencies, both through the different product levels (2, 2G, 3 and 4) and between products. In this way, for example, all files affected by a known issue, or, the origin of known issues in input products, can be identified.

5.4.4 Time Series Analysis

A time series of summary statistics derived from all the gridded MODIS Land products at fixed globally distributed locations is maintained and monitored by the LDOPE personnel. Prod-

uct time series analyses are important because they capture algorithm sensitivity to surface (e.g., vegetation phenology), atmospheric (e.g., aerosol loading) and remote sensing (e.g., sun-surface-sensor geometry) conditions that change temporally, and because they allow changes in the MODIS characterization and calibration to be examined. Time series statistics are extracted at nine MODIS 1200 x 1200km areas that are expected to be representative of the variability of the Land products. These are referred to as the land "golden" tiles. A web interface is provided to enable visualization of time series plots generated with respect to different Land cover and biome types for each golden tile and provides access to browse images in JPEG format over each golden tile. Product files with large deviations from the surrounding time series values are examined in more detail as they indicate a potential quality related problem. Figure 5.3 shows time series extracted from the Land surface temperature product (Wan *et al.* 2002) for the first five years of MODIS Terra data. The plot immediately confirms two expected results: the seasonal variation in temperature and the substantial difference between the mean day and night temperatures.

6. Validation Approach

Building on lessons learned from the previous generation of global land imaging systems (Justice and Townshend, 1994; Cihlar et al., 1997), the MODIS Land team has committed to quantify the accuracy of the MODIS land product suite (Morisette, 2002). Determining the accuracy of the output products is referred to as validation. The validation activities have resulted in online accuracy statements for each of the land products. The MODIS Land team maintains and updates the accuracy statements and posts related supporting material on the MODLAND validation page http://landval.gsfc.nasa.gov, pull-down menu "Val Status". An example from this site for the LAI/FPAR product is illustrated below in Figure 6.1.

While MODIS Land products span a range of spatial scales, accurate validation data are typically derived from point measurements collected on the ground (Gower et al., 1999). Aggregating point data over larger areas (scaling up) is a critical component of validating relatively coarse-resolution global products (Cohen and Justice, 1999). In order to account for differences in scale, the MODLAND validation team (Cohen et al. 2006; Yang et al. 2006; Morisette et al. 2002) and the international community (Yang et al. 2006; Morisette et al., 2006) have emphasized the coupling of field data with fine-resolution airborne or satellite imagery. Establishing relationships between field data and the high-resolution imagery allows extrapolation of the point measurement to the continuous area covered by the imagery. The imagery can then be aggregated in a way that will allow for comparison with the coarse resolution MODIS data. The MODIS validation team has coordinated the acquisition of a multitude of medium- to fine-resolution data sets since 1999 to facilitate the up-scaling of ground measurements to digital data sets at resolutions used to validate of MODIS products as illustrated in Figure 6.2.

The scaling procedure is greatly enhanced by having the ASTER and MODIS sensors on board the same Terra platform. This allows for the validation activities to utilize the simultaneous, high spatial and radiometric data from ASTER to aggregate field measurement to the MODIS scale. Clearly the field, high resolution, and MODIS data should be as coincident in time as possible. How close they need to be in time will depend on the temporal dynamics of the product being validated. LAI, for example, will remain fairly stable over several days (Yang et al. 2006). Likewise, field-measured LAI from one day can be compared to imagery acquired within several days of the field data. In contrast, active fires are extremely dynamic. Field data and high-resolution imagery of fires acquired on a different days do not help validate a MODIS fire product. The ability to acquire coincident ASTER/MODIS data has provided the primary

validation data for the MODIS fire product (Csiszar et al. 2006; Morisette et al. 2005a; Morisette et al. 2005b). The Oak Ridge National Laboratory (ORNL) DAAC in Oak Ridge, Tennessee has archived combined ASTER and MODIS fire validation data for southern Africa (Morisette et al. 2003) and Brazil (Morisette et al. 2004).

Combined field measurements and ASTER high-resolution image data are currently being explored for validating other MODIS products. MODIS subsets and multi-sensor validation data acquired over the EOS Land Validation Core Sites are available through NASA's EOS Land Validation Core Site infrastructure on the MODLAND validation webpage. These Core Site pages provide links to ground, airborne, and satellite data sets (Nickeson et al., 2006) coordinated for MODIS validation. The Core Sites were established to provide the general community with some of the best and simplest opportunities for multi-sensor data comparisons and synergistic science. The Core Site philosophy has been to collect, archive and distribute as much EOS and other Earth observing satellite data as possible. Readers interested in exploring product and sensor inter-comparison are encouraged to utilize the data available through the EOS Land Validation Core Sites.

Some of this product inter-comparison has begun through cooperation within the Committee on Earth Observation Satellites (CEOS) Land Product Validation (LPV) subgroup (http://lpvs.gsfc.nasa.gov/). This is one of six subgroups of the CEOS Working Group on Calibration and Validation (WGCV). The LPV subgroup is leading international efforts on inter-comparison of global LAI and Albedo products currently available. As part of these inter-comparison efforts, an additional set of sites have been established as CEOS LPV Core Sites, some of which overlap with the EOS Core Sites. The web infrastructure currently supporting the

EOS Land Validation Core Sites serves as a model for this international effort and is being utilized to support the coordination of inter-comparison data sets.

7. Conclusion

This chapter describes the process by which the algorithms and the processing of the MODIS Land products have evolved since the start of the EOS Terra and Aqua missions. The MODAPS provided an environment that initially enabled rapid improvement in the quality of MODIS products and then later provided the stability needed to create a consistent time-series of validated products. Key to this effort was a committed science team with both the resources and the responsibility for algorithm development and refinement, as well as, product quality assurance and validation. The science team was complemented by a central support and data processing team that coordinated software deliveries, testing and software integration and assisted science team members by providing operational product quality assessment. A close working relationship between these two groups has accounted for the success of the MODIS land data production.

Lessons learned from this effort now are being applied to production of data sets from heritage instruments such as AVHRR and Landsat and to future missions such as the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP). In these efforts the experience of the MODIS science and data processing teams is invaluable in producing and assessing long-term land data records which will be used to study changes in the Earth's climate and ecosystems.

Acknowledgement

References

G. Asrar and H.K. Ramapriyan, "Data and information system for mission to planet earth," *Remote Sensing Reviews,* vol. *13*, pp. *1–25*, 1995.

J. Cihlar, J. Chen and Z. Li, "On the Validation of Satellite-derived Products for Land Applications," *Canadian Journal of Remote Sensing*, vol. *23*, no. *4*, pp. *381-389*, 1997.

W.B. Cohen, T.K Maiersperger, D.P. Turner, W.D. Ritts, D. Pflugmacher, R.E. Kennedy, A. Kirschbaum, S.W. Running, M.Costa, S.T. Gower, "MODIS Land Cover and LAI Collection 4 Product Quality across Nine Sites in the Western Hemisphere, *IEEE Transactions on Geoscience and Remote Sensing* (in press).

W.B. Cohen and C.O. Justice, "Validating MODIS Terrestrial Ecology Products: Linking *In situ* and Satellite Measurements," *Remote Sensing of Environment*, vol. *70*, no. *1*, pp. *1-3*, 1999.

I. A. Csiszar, J. T. Morisette, L. Giglio, "Validation of active fire detection from moderate resolution satellite sensors: the MODIS example," *IEEE Transactions on Geoscience and Remote Sensing* (in press).

J.C. Farman, B.G. Gardiner, J.D. Shanklin, "Large losses of total ozone in Antarctica reveal seasonal ClO$_x$/NO$_x$ interaction," *Nature,* vol. *315*, pp. *207-210*, May, doi:10.1038/315207a0, 1985.

M.A. Friedl, D.K. McIver, J.C.F. Hodges, X.Y. Zhang, D. Muchoney, A.H. Strahler, C.E. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, C. Schaaf, "Global Land cover mapping from MODIS: algorithm and early results," *Remote Sensing of the Environment*, vol. *83*, pp. *287-302*, 2002.
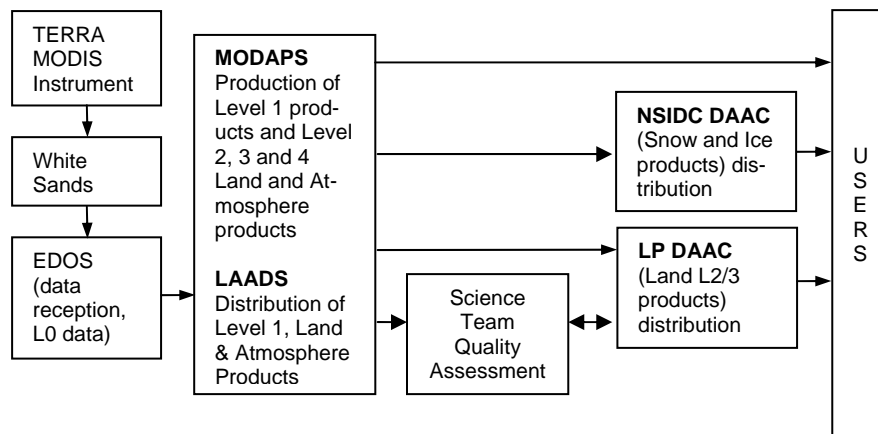
S.T. Gower, C.J. Kucharik, J.M. Norman, "Direct and Indirect Estimation of Leaf Area Index, f$_{APAR}$, and Net Primary Production of Terrestrial Ecosystems," *Remote Sensing of Environment*, vol. *70*, pp. *29-51*, 1999.

D.K. Hall, G. Riggs, V.V. Salomonson, N.E. DiGirolamo, K.J. Bayr, "MODIS snow-cover products," *Remote Sensing of Environment*, vol. *83*, pp. *181-194*, 2002.

M.C. Hansen, R.S. DeFries, J.R.G.Townshend, R. Sohlberg, C. Dimiceli, M. Carrol, "Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data," *Remote Sensing of Environment*, vol. *83*, pp. *303-319*, 2002.

A. Huete, K. Dida, T. Miura, E.P. Rodriguez, X. Gao, L.G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sensing of Environment*, vol. *83*, pp. *195-213*, 2002.

C.O. Justice, E. Vermote, J.R.G. Townshend, R. Defries, D.P. Roy, D.K. Hall, V.V. Salomonson, J.L. Privette, G. Riggs, A. Strahler, W. Lucht, R.B. Myneni, Y. Knyazikhin, S.W. Running, R.R. Nemani, Z. Wan, A.R. Huete, W. van Leeuwen, R.E. Wolfe, L. Giglio, J-P. Muller, P. Lewis, M.J. Barnsley, "The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research," *IEEE Transactions on Geoscience and Remote Sensing*, vol. *36*, pp. *1228-1249*, 1998.

C.O. Justice, L. Giglio, S. Korontzi, J. Owens, J.T. Morisette, D. Roy, J. Descloitres, S. Alleaume, F. Petticolin, Y. Kaufman, "The MODIS fire products," *Remote Sensing of Environment*, vol. *83*, pp. *244-262*, 2002a.

C.O. Justice, L. Giglio, D. Roy, L. Boschetti, I. Csiszar, D. Davies, S. Korontzi, W. Schroeder, K. O'Neal, "Global Fire Products from the MODIS instruments," In this volume.

C.O. Justice, J.R.G Townshend, E.F. Vermote, E. Masuoka, R.E. Wolfe, N. Saleous, D.P. Roy, J.T. Morisette, "An overview of MODIS Land data processing and product status" *Remote Sensing of Environment*, vol. *83*, pp. *3-15*, 2002b.

C.O. Justice and J.R.G. Townshend, "Data sets for global remote sensing: lessons learnt," *International Journal of Remote Sensing*, vol. *15*, no. *17*, pp. *3621-3639*, 1994.

B. Lutz , D. Roy, C. Leff, S. Lewicki, E. Geir, D. Ziskin, K. Kilpatrick, A. Chu, "A review of EOS Terra quality assessment (QA)," Proceedings of IEEE Geoscience and Remote Sensing Symposium (IGARSS), Honolulu, HI, 24 – 28 July. (CD-ROM INT_27_06.pdf, 0-7803-6362-0/00 (c) 2000 IEEE), 2000.

J.T. Morisette, F. Baret, J. L. Privette, R. B. Myneni, J. E. Nickeson, et al., "Validation of Global moderate resolution LAI Products: a framework proposed within the CEOS Land Product Validation subgroup," *IEEE Transactions on Geoscience and Remote Sensing*, 2006.

J.T. Morisette, L.Giglio, I. Csiszar, C.O. Justice, "Validation of the MODIS Active fire product over Southern Africa with ASTER data," *International Journal of Remote Sensing*, vol. *26*, no. *19*, pp. *4239–4264*, 2005a.

J.T. Morisette, L. Giglio, I. Csiszar, A. Setzer, W. Schroeder, D. Morton, C. O. Justice, "Validation of MODIS active fire detection products derived from two algorithms," *Earth Interaction*, vol. *9*, paper *9*, 2005b.

J.T. Morisette, L. Giglio, I. Csiszar, A. Setzer, W. Schroeder, D. Morton, C. O. Justice, LBA ECO LC-23 ASTER - MODIS Fire Data Comparison - Brazil 2003 & 2004. Data set. Available on-line [http://www.daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A, 2004.
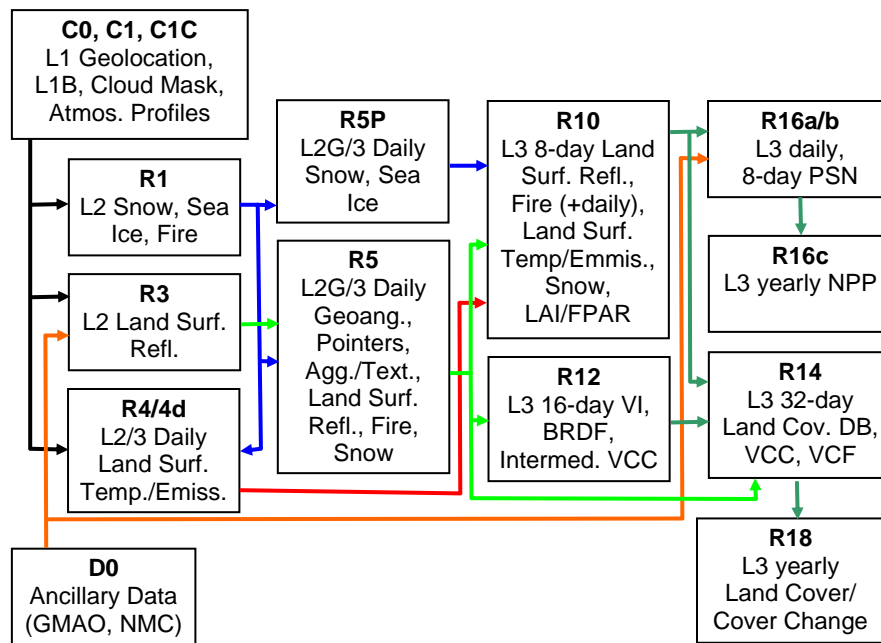
J.T. Morisette, L. Giglio, I. Csiszar, C. O. Justice, SAFARI 2000 ASTER AND MODIS FIRE DATA COMPARISON, DRY SEASON 2001. Data set. Available on-line [http://www.daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. Conference Proceeding, 2003.

J.T. Morisette, J.L. Privette, C.O. Justice, "A framework for the validation of MODIS Land products," Remote Sensing of Environment vol. *83*, pp. *77-96*, 2002.

J.E. Nickeson, J. T. Morisette, J. L.Privette, C. O. Justice, D. Wickland, "Coordinating EOS Land Validation: Core Site Status in 2006" (submitted).

R.B. Myneni, S. Hoffman, Y. Knyazikhin, J.L. Privette, J. Glassy, Y. Tian, Y. Wang, X. Song, Y. Zhang, G.R. Smith, A. Lotsch, M. Friedl, J.T. Morisette, P. Votava, R.R. Nemani, S.W. Running, "Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data" *Remote Sensing of Environment*, vol. *83*, pp. *214-231*, 2002.

S. Platnick, M.D. King, S.A. Ackerman, W.P. Menzel, B.A. Baum, J.C. Riédi, R.A. Frey, "The MODIS cloud products: Algorithms and examples from Terra," *IEEE Transactions on Geoscience and Remote Sensing*, vol. *41*, pp. *459–473*, 2003.

D.P. Roy, J.S. Borak, S. Devadiga, R.E. Wolfe, J. Descloitres, "The MODIS Land product quality assessment approach," *Remote Sensing of Environment*, vol. *83*, pp. *62-76*, 2002a.

S.W. Running, C.O. Justice, V.V. Salomonson, D. Hall, J. Barker, Y.J. Kaufman, A.R. Strahler, J-P. Muller, V. Vanderbilt, Z.M. Wan, P. Teillet, D. Carneggie, "Terrestrial remote sensing science and algorithms planned for the MODIS-EOS," *International Journal of Remote Sensing*, vol. *15*, no. *17*, pp. *3587-3620*, 1994.

S.W. Running, R.N. Nemani, F.A. Heinsch, M. Zhao, M.C. Reeves, H. Hashimoto, "A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production," *BioScience*, vol. *54*, no. *6*, pp. *547-560*, 2004.

V.V. Salomonson, W.L. Barnes, P.W. Maymon, H.E. Montgomery, H. Ostrow, "MODIS: Advanced facility instrument for studies of the Earth as a system" *IEEE Transactions on Geoscience and Remote* Sensing, vol. *27*, pp. *145-153*, 1989.

C.B Schaaf, F. Gao, A.H. Strahler, W. Lucht, X. Li, T. Tsang, N.C. Strugnell, X. Zhang, Y. Jin, J-P. Muller, P. Lewis, M. Barnsley, P. Hobson, M. Disney, G. Roberts, M. Dunderdale, C. Doll, R. d'Entremount, B. Hu, S. Liang, J. Privette, D. Roy, "First operational BRDF, albedo nadir reflectance products from MODIS," *Remote Sensing of Environment*, vol. *83*, pp. *135-148*, 2002.

J.P. Snyder, "Map projections – a working manual," *U.S. geological survey professional paper 1395*, Washington, DC: United States Government Printing Office, 1987.

E.F. Vermote, N.Z. El Saleous, C.O. Justice, "Atmospheric correction of MODIS data in the visible to middle infrared: first results," *Remote Sensing of Environment*, vol. *83*, pp. *97-111*, 2002.

R.E. Wolfe, D.P. Roy, E. Vermote, "MODIS Land data storage, gridding and compositing methodology: Level 2 Grid," IEEE *Transactions on Geoscience and Remote* Sensing, vol. *36*, pp. *1324-1338*, 1998.

Z. Wan, Y. Zhang, Q. Zhang, Z. Li, "Validation of the Land-surface temperature products retrieved from Terra Moderate Resolution Image Spectroradiometer data," *Remote Sensing of Environment*, vol. *83*, pp. *163-180*, 2002.
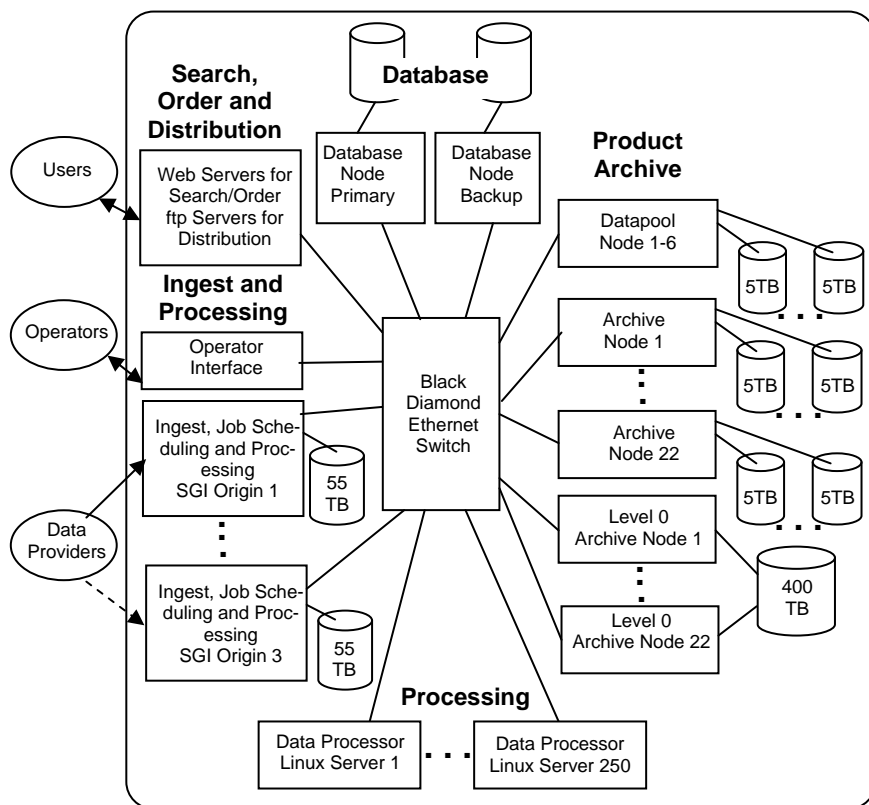
W. Yang, B. Tan, D. Huang, M. Rautiainien, N.V Shabanov, Y. Wang, J. L Privette, K. F Huemmrich, R. Fensholt, I. Sandholt, M. Weiss, R.R Nemani, Y. Knyazikhin, R. B Myneni, "Validation of Collections 3 and 4 Terra MODIS Leaf Area Index and Fraction Vegetation Absorbed Photosynthetically Active Radiation Products," *IEEE Transactions on Geoscience and Remote* Sensing. (in press).

X. Zhang, R.A. Sohlberg, J.R.G. Townshend, C. DiMiceli, M.L. Carroll, J.C. Eastman, M.C. Hansen, R.S. DeFries, "Detection of Land cover changes using MODIS 250 m data," *Remote Sensing of Environment*, vol. *83*, pp. *320-335*, 2002.
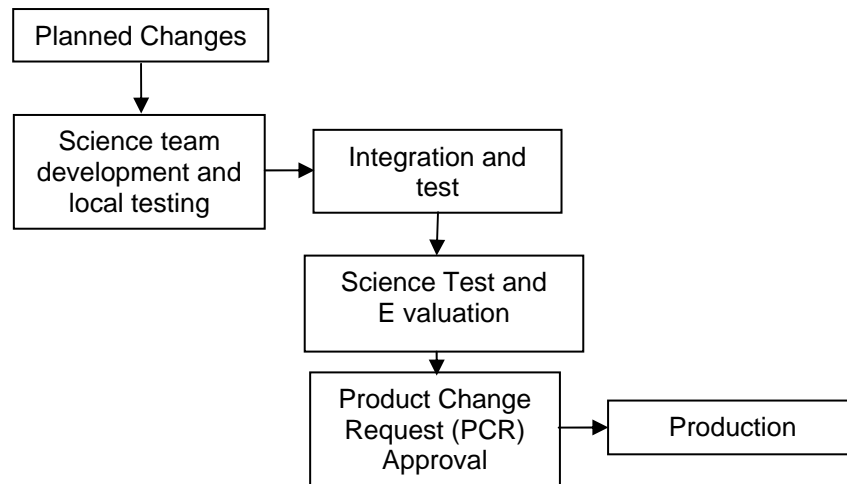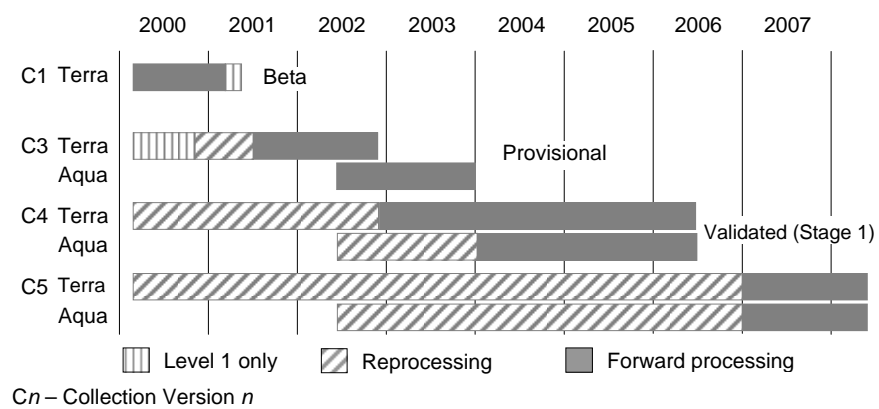
Fig. 3.1

Fig. 3.2

Search, Order and Distribution

Users

Web Servers for Search/Order ftp Servers for Distribution

Database

Database Node Primary

Database Node Backup

Product Archive

Datapool Node 1-6

5TB 5TB

Ingest and Processing

Operators

Operator Interface

Black Diamond Ethernet Switch

Archive Node 1

5TB 5TB

Ingest, Job Scheduling and Processing SGI Origin 1

55 TB

Data Providers

Archive Node 22

5TB 5TB

Ingest, Job Scheduling and Processing SGI Origin 3

55 TB

Level 0 Archive Node 1

Level 0 Archive Node 22

400 TB

Processing

Data Processor Linux Server 1

Data Processor Linux Server 250

Fig. 3.3

Fig. 3.4

Fig. 4.1

Figure 5.1

Figure 5.2

MOD11A2 h09v05 (Southwestern U.S.) Grassland_Biome
Terra, Collection 4, Year 2000-2005

Tue Jan 17 15:09:29 2006

Figure 5.3

Fig. 6.1

Combine field and tower data with high resolution imagery to produce high resolution products.

Aggregate Data Products

MODIS Scene

Correlate
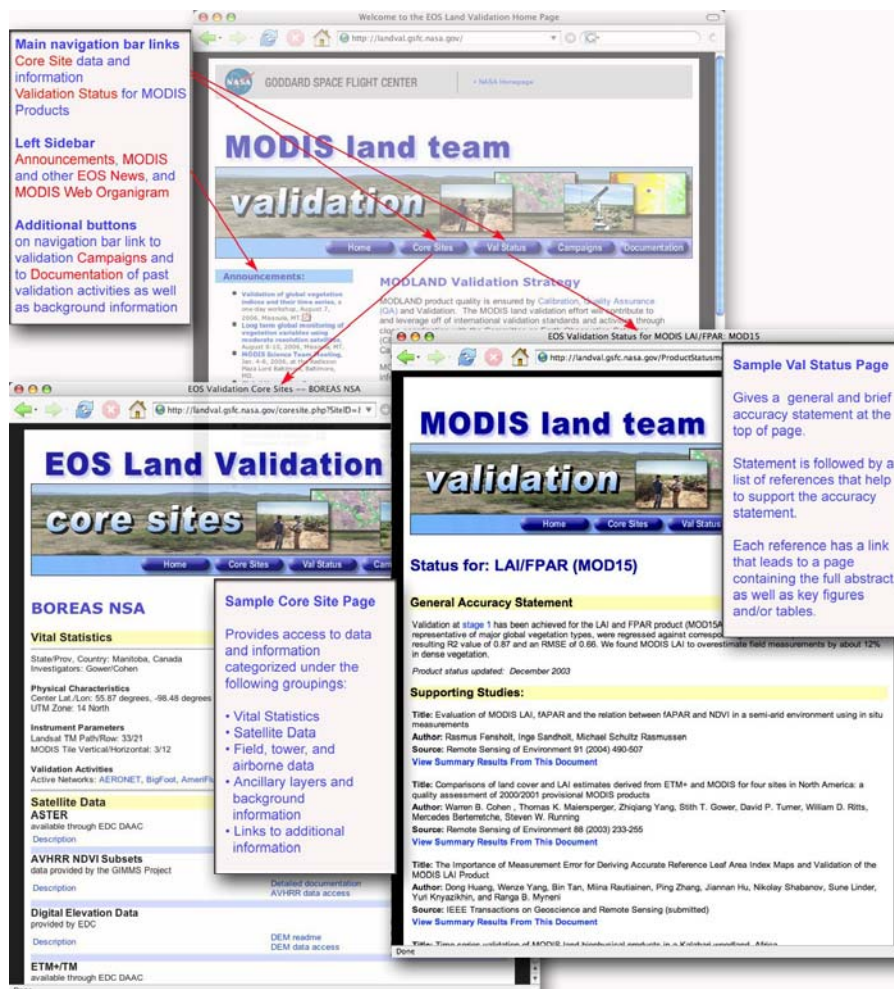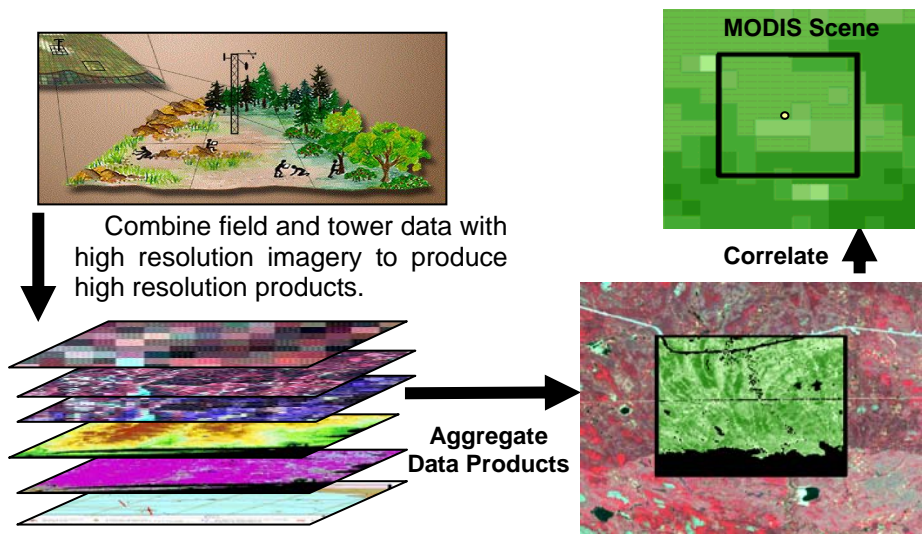
Fig. 6.2

Product	Name	DAAC	Level(s)	Resolution(s)	
				Spatial	Temporal

Radiation balance product suite

Product	Name	DAAC	Level(s)	Spatial	Temporal
MOD09	Surface reflectance	LP	L2G, L3	250 m, 500 m, 1 km, 0.05 deg.	1 day, 8 days
MOD11	Surface temperature and emissivity	LP	L2, L3	1km, 5 km, 0.05 deg.	5 min., 1 day, 8 days, monthly
MOD43	BRDF/Albedo	LP	L3	500 m, 0.05 deg.	16 days
MOD10	Snow cover	NSIDC	L2, L3	500 m	5 min., 1 day, 8 days
MOD29	Sea ice extent	NSIDC	L2, L3	500 m	5 min., 1 day

Ecosystem product suite

Product	Name	DAAC	Level(s)	Spatial	Temporal
MOD13	Vegetation indices	LP	L3	1 km, 500 m, 250 m	16 days
MOD15	LAI and FPAR	LP	L4	1 km	8 days
MOD17	PSN and GPP	LP	L4	1 km	8 days, 1 year

Land cover characterization product suite

Product	Name	DAAC	Level(s)	Spatial	Temporal
MOD12	Land cover and vegetation dynamics	LP	L3	1 km, 0.05 deg.	1 year
MOD14	Thermal	LP	L2, L3	1 km	5 min., 1 day,

	anomalies and				8 days
	fire				
MOD44	Vegetation	LP	L3	500 m, 250 m	96 day,
	Cover Change				1 year
	and VCF				

Table 2.1 MODIS Land data products and characteristics

LP = Land Processes DAAC at the EROS Data Center

NSIDC = National Snow and Ice Data Center

BRDF = Bidirectional Reflectance Distribution Function

LAI = Leaf Area Index

FPAR = Fractional Photosynthetically Active Radiation

PSN = Net Photosynthesis

GPP = Gross Primary Production

VCF = Vegetation Continuous Fields

	MODAPS Production (GB/day)		Export Volume (GB/day)				
			LP DAAC		NSIDC DAAC		MODAPS
	C4	C5	C4	C5	C4	C5	C5
Level 1	324	227					*227*
L2 – L3 Daily	456	140	**265**	**31**	**7**	**1**	*140*
Level 3 8-day >	23	26	**70**	**19**	**1**		**19**
Total	803	393	**335**	**50**	**8**	**1**	524

Key:

Italicized = produced on demand,

Bold = Archived on disk or tape

C4 = Collection 4 products,   C5 = Collection 5 products

Table 3.1 Average daily volume of products produced and archived for each 1x of processing for products from Collection 4 and Collection 5.

Per-Pixel QA code	Meaning
0	Pixel produced, good quality, not necessary to examine more detailed QA
1	Other quality (produced or not produced, if produced unreliable or unquantifiable quality, examination of more detailed QA is recommended.

Table 5.1 Generic per-pixel QA stored in all Collection 5 Land products

Fig. 3.1 MODIS-Terra land product data flows, from spacecraft to end-users.

Fig. 3.2 MODIS Land production sequence showing the data flow between algorithms running in MODAPS. Processes recipes are denoted by a letter and number.

Fig. 3.3 Elements of the MODAPS system.

Fig. 3.4 The MODIS Land science algorithm update and science test process flow.

Fig. 4.1 Summary of the different collections of MODIS data sets produced or planned, the range of data included in each collection and its production time.

Figure 5.1 Example of a known issue posted on the LDOPE Web Site (see text for details).

Figure 5.2 Example of a MODIS surface reflectance global browse image posted on the on the LDOPE Web Site. The MODIS 0.645, 0.555 and 0.469 um bands are displayed as red, green and blue respectively for MODIS Terra L2 Land surface reflectance sensed on Julian day 2005/150.

Figure 5.3 Time series plot posted on the LDOPE Web Site. Time series of mean 8-day 1km day and night Land surface temperature (LST) computed over approximately 120,000 grass Land biome pixels in a southwestern US tile are shown for 2000 to 2005.

Fig. 6.1 MODIS Land validation web site with accuracy statement and supporting material for LAI/FPAR (MOD15) product.

Fig. 6.2 Multi-tiered validation scheme involves scaling from field measurements and high spatial resolution Earth imaging sensors to MODIS resolution.