U.S. 16 day Vegetation Index Product Guide

* Projection: Alber's Equal Area Conic Conformal

- reprojected from native MODIS projection: Sinusoidal
- Central Meridian used was 96d W
- std Parallels of 29d 30' 0" and 45d 30' 0"
- Latitude of Origin is 0d 0' 0"
- Easting & Northing are 0.0000000000 m
- Datum: sphere (a sphere of radius 6370997 m)
- Resolution is 250 meters

NOTE - When you load the geotiff, many programs may not correctly recognize the datum and report a wrong datum (i.e., WGS 1984). THE DATUM IS THE SPHERE USED IN MODIS DATA, WITH A RADIUS OF 6370997 METERS!!!

- There are three data files: 1 contains the NDVI, the other two contain the MODIS bands 1 and 2 respectively
- In addition there is a file which contains cloud information. This information is not available for the 2000 data but is available from 2001 Jan. 1 forward.
- Values in the cloud file range from 0 to 3, use the legend below to interpret:
    - 0=Clear
    - 1=Cloud
    - 2=Mixed cloud
    - 3=Not set assumed clear
- Image Size: 21000 Pixels x 13000 Lines
- Geographic Upper Left Corner: 128d 58' 17.49" W 49d 59' 50.83" N
- Geographic Lower Right Corner: 68d 38' 9.69" W 20d 36' 24.18" N
- Geocoded Upper Left Corner (m): -2356875.000 E 5769625.000 N
- Geocoded Lower Right Corner (m): 2893125.000 E 2519625.000 N
- Data format for the NDVI file is geotiff with 8 bit Unsigned integer
    - NDVI has been scaled using the formula (NDVI * 200)+50, this yeilds values from 0-250 corresponding to NDVI range of -0.25 - 1
    - No Data value: 253
- Data format for the Band files is geotiff with 16 bit Signed integer
    - No Data value: -200
    - Fill value: -28672
- For more information on the Dn values see the Surface Reflectance.
- These data should be considered a provisional release. Expect only minor improvements to the compositing procedure in later versions.
- There are NDVI values greater than 0 apparent in the water. These are naturally occurring values caused by a variety of things including turbidity in the water. In order to maximize the usefulness of the data we did NOT apply a water mask. The user could approximate water bodies by using the band 2 data (NIR) and creating a mask of all values less than 500 Dn. The user could be more or less agressive as they see fit.

Known Issues with specific time periods

- 2000 049 - 081 No data were available East of the Bahamas
- 2000 145 No data were available for a small section of Baja California
- 2001 161 - 177 Instrument problems caused no data for part of these two time periods result is more than expected cloud cover in these two time periods
- 2002 161 Geolocation shift in Western US seen as double rivers

- 2002 209 Geolocation shift in Western US seen as disappearing lakes

* The geolocation problems could be resolved if upstream data are reprocessed

For questions comments or concerns contact the GLCF.

MODIS Data Compositing: Lessons Learned at the SCF

Mark Carroll, Charlene DiMiceli, John Townshend, Rob Sohlberg
Department of Geography, University of Maryland, College Park

Abstract

As a MODIS Science Team member it was necessary to develop products using a variety of MODIS data. In the course of developing two standard products, Vegetative Cover Conversion and Vegetation Continuous Fields, a time series of cloud free composites was required. Here we describe methods and lessons learned for compositing MODIS L2G daily data, and MOD09A1 8-day surface reflectance data into composites of 16 and 32-day intervals respectively.

Background

MODIS data is available in three spatial resolutions, 250m 500m and 1km. Bands 1 and 2 are natively 250m resolution. Bands 3 through 7 are natively 500m resolution. The remaining bands are 1 km resolution. Bands 1 through 7 contain the visible, NIR, and SWIR (Short Wave Infra-red) which are of greatest importance for monitoring vegetation dynamics. It is on these bands that our efforts have focused. There are a variety of different data levels that any satellite data goes through. Level 0 data is straight from the instrument, raw radiance values. This form is not friendly to most image processing systems and requires custom code to transform it into something more end user friendly. This product is not made available to the general public. Level 1 products are the lowest level of product available to the end user. These products are in 5 minutes chunks of sensor acquisition and are still in swath format. Swath format is not a projection it is the view as the sensor “sees” the Earth. Though this product is available to the end user special tools are still required to manipulate these data into a map projection and to correct for the MODIS “bow-tie” effect. The bow tie is a phenomena associated with sweeping instruments where the instantaneous field of view increases as the sensor sweeps away from nadir. As the instrument looks further out to the side it “sees” a larger area on the ground. A consequence of this is that subsequent scans will have overlapping observations on the outer edges of a swath. This needs to be accounted for when putting the data into a map projection. The next step in the processing chain is Level 2. For MODIS Level 2 products are corrected for atmospheric contamination (aerosols, and geo angles). Level 2 products are not archived, they are immediately mapped into the Sinusoidal projection and gridded into 10-degree units called tiles. These files are daily level 2 gridded (L2G) surface reflectance and associated products in a format, which is easily read by image processing software and manipulated by an end user. Level 3 and 4 products are end-user value added products nearly all of which were derived from composites of some kind. For purposes of data reduction and removal (or minimization) of bad data it is useful to take the best observations over a series of days and produce one output. This process is called compositing and has been done for years with a variety of different satellite data products. For the MODIS instrument the repeat cycle of nadir overpasses is 16 days. This means that every 16 days the instrument will be traveling on nearly the exact same path. For this reason standard composite periods for MODIS are multiples of 8 days, exactly the mid-point of the repeat cycle. Thus we see time steps in composited products of 8, 16 and 32 days of data. In this paper we will discuss methods and lessons learned on the path to creating MODIS data products at the University of Maryland. Methods presented are the culmination of several years of effort on the part of the staff of the UMD-SCF. Readers are encouraged to see the products for themselves at our website http://modis.umiacs.umd.edu .

MOD44C

Vegetative Cover Conversion (VCC) is designed to be a global alarm product for areas where land cover changes are occurring due to anthropogenic influences. This product is generated at 250m resolution from 16-day composites of MODIS data. To this end the University of Maryland Science Computing Facility (UMD SCF) developed a compositing algorithm that takes daily L2G MODIS data and produces 16-day composites. This algorithm is called MOD44C, or the intermediate data collector for VCC. The time step of 16 days was chosen due to the large size of daily 250m MODIS data. In early production, the 250m L2G data was only produced for 10% of the global land tiles. This presented a problem for developing the VCC product due to a lack of a global spectral signature for vegetation. It has only been in year 3 of the program that 250m data was available globally, and it has taken until the end of year 4 (and the completion of the Collection 4 reprocessing) to get the entire MODIS 250m data record processed globally. In the meantime, a considerable amount of effort was expended to develop and refine the compositing methods for MOD44C. Specific effort was directed at maximizing spatial resolution of the composite. The compositing algorithm for MOD44C is based on the concept of choosing the value, which is the most cloud-free and highest spatial resolution possible. Quality flags in the L2G data are read in order to eliminate pixels that have contamination from clouds, high aerosols, and cloud shadow. For the remaining high quality pixels, NDVI is calculated. The top ten percent of these values are retained as the “maximum” values. From this pool of values the candidate with the lowest sensor zenith angle (i.e. closest to nadir view) is chosen. In the event that none of the pixels meet the criteria that have been established a value is chosen if it has a minimum of cloud and aerosol contamination. The quality flags for the pixels that are chosen are retained in a separate layer for use by later processes. In this effort a considerable amount of programming has been involved in order to handle the data structures in the HDF-EOS file format of MODIS L2G data. Early efforts focused on the use of the first plane in the L2G files and did not take into consideration the quality flags. The additional layers in the HDF files are comprised of the overlapping observations from the off nadir pixels and from overlapping instrument overpasses. It has been found that by not using the additional layers of data in the HDF file and using the quality flags the output product will be of insufficient quality to perform land cover change analysis. It is necessary, in particular, to minimize the occurrence of pixels where the aerosol contamination was considered to be high. The corrected values could have abnormally low reflectance values particularly in band 1. In addition the user should be aware that the quality flags are determined based on a 1 km resolution pixel. Since there are 16 250m pixels in every 1 km pixel, using the quality flags can create a blocky appearance in the composite. While this may make the composites appear noisy, it is accounted for by the VCC algorithm as it reads the quality flags. MOD44C uses only the 250m bands 1 and 2 and the quality flags to decide which observations to retain. This information is then used to select appropriate observations from the 500m bands and populate them into a 250m data plane. This is accomplished using daily L2G observation pointer files, which describe the crosswalk between the 3 resolutions of MODIS data. By populating the new data layers based on 250m observations it creates an enhanced data layer. The resulting enhanced data layers are then available for use by other algorithms such as the VCC and possibly the VCF algorithms.Considerable advances have been made in the quality of the surface reflectance daily product through algorithm improvements in the Collection 4 data set. The three years of MODIS data processed through the MOD44C algorithm have been used in the change detection product with success. It is apparent that the improvements made to the daily product as well as advances in our understanding of the data have resulted in an improved composite product. Future advances in the L2G data set including a 250m resolution cloud flag would significantly benefit this product. This product will easily serve the needs of the change detection algorithm as originally planned.

VCF

Vegetation Continuous Fields is a global 500m product depicting proportion of a pixel, which belongs to a particular cover type. This product is not a land cover classification; rather it is a proportional membership to a particular life form (tree, shrub, bare, etc). To generate this product at least 1 full year of a global 500m data set was necessary on a 32-day time step. The standard MODIS surface reflectance data composite only to 8-day time step. In order to produce 32-day composites 4 of these 8- day products were combined to produce a cloud free monthly composite. The 8-day product, MOD09A1, is generated using the minimum blue compositing method and is freely available from the LP-DAAC (Land Processes Distributed Active Archive Center). Details on the minimum blue compositing method can be found at the Surface Reflectance SCF and will not be discussed here. (http://modis.gsfc.nasa.gov/mod09) Several methods of combining these 8-day products into a 32-day product were attempted before finally settling on what has been described as a second darkest albedo approach. The early attempts will be mentioned below but not discussed in detail. In early processing attempts were made to further composite the MOD09A1 data using the minimum blue approach. The results of this yielded an output contaminated with numerous large dark features. These features appeared to be related to an accumulation of shadowed (possibly cloud shadow) areas. This is a known side effect of the minimum blue compositing method. The next attempt was to apply a maximum NDVI approach. For this approach the NDVI is calculated for each of the 8-day composites. A single output is created by choosing the observations with the maximum value for NDVI. The results of this process yielded a product with an accumulation of overly bright areas. These results, while better than the results of the minimum blue compositing method, still yielded an output that was of insufficient quality for use in the Vegetation Continuous Fields product. With the previous 2 methods producing an insufficient data set a completely new approach was tried. It was observed from the early tests that there were 2 specific " anomalous" values that needed to be eliminated bright and dark areas. Albedo is a measure of brightness in an observation. It is possible to estimate the albedo by generating the sum of the reflectance values of the visible bands. Using the 4 8-day MOD09A1 products the albedo was calculated and then through trial and error it was determined that the second darkest albedo yielded results of sufficient quality to be used in the Vegetation Continuous Fields product. This method while slightly unconventional has proven to yield excellent results in the VCF product see results: http://modis.umiacs.umd.edu/vcf.htm