Product User's Guide
Vegetation Cover Conversion
University of Maryland, Department of Geography
July 2006

1. Introduction

The Vegetative Cover Conversion (VCC) method has evolved over time. The original method employed 5 different change detection methods and then combined them to yield a single end result (Zhan 2002). Also different types of change were determined simultaneously by the algorithm as it processed the data. In contrast the current method uses separate methods for determining different types of change. The method for deforestation is derived from the original space partitioning method (Zhan 2002) and relies on decision tree classification (Breiman, 1993) to determine antecedent vegetation condition and compares that to current vegetation condition. Change due to burning is derived using the difference Normalized Burn Ratio (dNBR) (van Wagtendonk, 2004) from two scenes a year apart. Flooding is determined similarly to deforestation using decision tree classification to determine where water exists in any given scene. This is then compared to a static map of water extent to determine areas that may be inundated. The differences in approach are necessitated by the very different space-time characteristics of the changes. For example extensive flooding typically lasts for days and weeks, whereas the signal of other types of change persists for months to years.

2. Inputs

In order to generate a global data product it is first necessary to assemble a coherent global data set. MODIS data are available in three spatial resolutions, 250m 500m and 1km (Barnes et al, 1998). Band 1 (red 620-670 nm) and band 2 (near-infrared (NIR), 841-876 nm) have 250m spatial resolutions and are also averaged to create 500m resolution products. Bands 3 - 7 (459-479 nm, 545-565 nm, 1230-1250 nm, 1628-1652 nm, 2105-2155 nm respectively) have a 500m spatial resolution. Bands 1 - 7 are designed primarily for remote sensing of the land surface.

For purposes of data reduction and removal or minimization of inferior data it is frequently useful to select the best observations from a series of days and produce a single output product. This process is called compositing and has been done for years for a variety of satellite data products (Holben, 1986). 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 MODIS composited products have been generated for 8, 16 and 32 day periods. Longer time periods lead to a reduction in problems such as clouds but have the potential disadvantage of including in the same image scene components with different phenologies.

In the early days of MODIS production, the 250m daily data was produced for only 10% of the global land tiles. Only in year 3 of the program has 250m data been available globally, and it has taken until the end of year 4 (and the completion of the Collection 4 reprocessing) for the entire MODIS 250m data record to be processed globally.

2.1. 250 meter Composite

The 250m compositing algorithm, known as MOD44C, uses only bands 1 and 2 at 250m resolution along with the quality flags to decide which observations to retain. The rules for the compositing algorithm for MOD44C are described in Table (1). The general concept is to choose the observation that is the most cloud-free and is closest to nadir so that it has the highest spatial resolution possible to represent the compositing period. Quality flags in the L2G data are used to eliminate pixels that are contaminated by clouds, high aerosols, or cloud shadow, and then the pixel with the lowest sensor zenith angle and highest overlap of the pixel with the MODIS grid cell is chosen. The "value" in column 2 of Table (1) refers to the value in the quality flags from the L2G data. A description of the quality flags and the values associated with a given phenomena is available in the user's guide for the MODIS Surface Reflectance Product (Vermote, 1999). For the selected pixels, the NDVI is then calculated. From this pool of pixels the single pixel with the least contamination from cirrus and aerosol is chosen to represent the given 16-day period. In the event that none of the pixels meet the criteria that have been established, the pixel with the minimum cloud and aerosol contamination is selected. The quality flags for the pixel that is chosen are retained in a separate layer for use by later processes.

The 250m pixels selected are used to identify the appropriate observations from bands 3 - 7 and then they are populated into a 250m data plane. This is accomplished using daily level 2 gridded (L2G) observation pointer files, which store the addresses of the L2 observations that intersect each L2G grid cell. This allows the appropriate coarse resolution observation to be populated into a finer resolution grid (i.e. 500m observations placed into the 250m grid) creating an enhanced resolution output. The resulting enhanced data layers are then available for use by other algorithms such as the VCC and possibly the VCF algorithms.

3. Change Products

3.1. Anthropogenic Deforestation Method

Decision trees are derived using a training data set that relates spectral phenomena to land cover. The decision tree can only be as good as the training data that is used to generate it, so special care must be taken to ensure the training data is of high quality. For the deforestation method, MODIS data from 2 different years are classified using the decision tree and then compared to identify differences. The training data set is generated from a spectral sample of MODIS data from the 2 years to be analyzed. To minimize potential impacts of spatially variable phenology the study region (30°N to 30°S) was stratified by month and by region. Three regions based on the continents were created Central and South America, Africa, and S.E. Asia/Oceania. For each month and region a spectral sample was taken from the input data using the MODIS Vegetation Continuous Fields (VCF) product to identify classes based on percent tree cover. Two sets of forest cover classes were derived from the VCF product and are shown in Table (2).

A land/water mask is employed to identify an additional class for water; this helps limit some confusion in the classifier between water and certain types of forest. A random sample of 20,000 observations per class is then chosen using the data quality flags (Roy, 2002) to avoid cloudy observations. This data sample then serves as the training data set and is used to generate a decision tree. The input data are processed on a quarterly basis yielding 4 data products per

data year. In each quarter there are six 16day composites except the fourth quarter which contains only five 16-day composites.
Based on date and location the algorithm chooses which decision tree is most appropriate to classify each image. Each of the input data sets in the quarter is classified under both classification schemes. Input data that have poor quality, such as cloudy, high view zenith angle, or heavy aerosols are not processed and a fill value is assigned in the classification. The classifications for year 1 and year 2 are then compared as in Table (3).

A change is determined to have occurred for each of the above comparisons:

- Scheme 1
    - Observation changed from class 4 to class 2 or 1
    - Observation changed from class 3 to class 1
- Scheme 2
    - Observation changed from class 5 to class 3, 2, or 1
    - Observation changed from class 4 to class 2, or 1
    - Observation changed from class 3 to class 1

In order to be represented in this layer a change must have been recorded in either scheme in at least 2 of the 6 time periods in the quarter. By using both schemes together up to 30% more actual changes were detected without any substantial increase in errors of commission.

3.1.1. Anthropogenic Deforestation Product

The output product is a single 8-bit data plane with the following values

0	No Change
1	Change
255	Undetermined (due to insufficient input data)

The undetermined class occurs only when all of the input data was cloudy or otherwise of poor quality. This occurs in less than 1% of all cases in the global transect from 30N to 30S. This product is a global alarm for land cover change. It is designed to be a conservative product that identifies locations where change is occurring but does not necessarily quantify that change.

Data year 2001 was compared to data year 2005 for this product. The outputs have been verified in specific locations in South America and Africa with independent sources of information both data sets and expert knowledge.

3.1.2. Known Issues for Anthropogenic Deforestation Product

- Wetlands can occasionally be labeled as change if they were particularly wet in 2001 and particularly dry in 2005
    - We have attempted to address this by using a water mask and DEM data from the Shuttle Radar Topography Mission (SRTM)
- Spurious un-identified clouds exist in the input data set and can cause false detections of change

    - We have attempted to address this by inserting a spectral test for clouds
- Terrain shadows in rugged mountains can cause false detections
    - Again we used DEM data and instrument and solar geometry to identify and remove likely areas of false change

3.2. Change Due to Burning Method

Burning may convert wooded areas to non-wooded areas. The label is restricted to wooded areas, where changes in land cover occur. Fires in grasslands and agricultural areas generally do not cause long term changes in land cover or land use.

The Normalized Burn Ratio (NBR) has been used to calculate burn extent and severity using Landsat TM and ETM+ data (van Wagtendonk, 2004). By calculating the NBR using a time series of data and calculating the difference NBR (dNBR) it is possible to locate areas where fires have occurred. Using composites that are 1 year apart, the dNBR is initially calculated only for those locations where fires were recorded in the MODIS active fire location product and the dNBR value is greater than a threshold set at 0.2. This threshold was derived by trial and error using a test data set. The active fire locations are pre-screened to select only those in or immediately adjacent to a wooded area. Wooded areas are determined using a combination of VCF layers (percent tree cover, percent herbaceous and percent bare) to determine likely ground cover. The observations showing dNBR values greater than the threshold are determined to be burns. This threshold is somewhat conservative, to reduce the errors of commission; other studies have used a threshold of 0.1 dNBR (van Wagtendonk, 2004). These areas are allowed to grow through an iterative process of extending to adjacent pixels with dNBR values higher than the 0.2 threshold. The number of iterations varies based on initial fire size as determined by number of MODIS active fire location points. Fires smaller than 1500 acres (6 fire location points), estimated by calculating the area based of 1 kilometer pixels from the MODIS active fire location product, were ignored.

Three sizes of fires were identified and different numbers of iterations were used with each. Allowing a large number of iterations on smaller fires often results in an inaccurate spreading of the changed area.

- Small fires: 6 - 18 fire location points (5 iterations)
- Medium fires: 19 - 62 fire location points (20 iterations)
- Large fires: > 62 fire location points (40 iterations)

Results are then summarized for the year to yield one product per year. Processing occurs on a per-fire basis and the composite is elected, which is closest to the start time of the fire, as determined by the MODIS active fire location points derived from the MODIS Web Fire-Mapper.

3.2.1: Change Due to Burning Product

The output product is a single 8-bit data plane with the following values

0	No Change
1	Change
251	Water

255 Undetermined (due to insufficient input data)

The undetermined class occurs only when all of the input data was cloudy or otherwise of poor quality. In addition, the dNBR values for the areas labeled as change are available in a separate layer. The dNBR values can be used as a confidence level where increasing dNBR value equals increasing confidence in the change. The dNBR has been scaled by 100, to return to the original difference value simply use the following equation: Dn/100.

The change due to burning product is produced for each data year starting with 2002 and proceeding through 2004 and represents change that occurred from the preceding year to the present year. (example the 2002 product shows change that occurred in 2002 by comparing to 2001 as a baseline) Processing is ongoing for additional years.

3.2.2. Known Issues for Change Due to Burning Product

- The NBR is susceptible to inter-annual variability in vegetation condition which can result in false detections.
- The best available global water mask is from the MODIS land cover data set which is at 1km, this leaves square, blocky, discontinuous looking water bodies at 250m resolution.
- The MODIS Vegetation Continuous Fields product was used to determine forest state. It is at 500m resolution and leaves some blocky artifacts where burns occurred in low density forest.

4. Outputs

These VCC products are distributed by the GLCF (Global Land Cover Facility) in GeoTIFF format using an in-house tile scheme based on a modified UTM grid, in geographic coordinates and a WGS84 datum. VCC products in native HDF format and Sinusoidal projection may be obtained from the LPDAAC (Land Process Distributed Active Archive Center).

5. References

Breiman, et al. 1993. Classification and Regression Tree Chapman and Hall 1993

Holben, B.N. 1986. Characteristics of maximum-value composite images for temporal AVHRR data. *International Journal of Remote Sensing*, 7, 1435 – 1445.

PRODES Digital 2002, http://www.obt.inpe.br/prodes/

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

Van Wagtendonk, J.W., Root, R.R., Key, C.H. 2004. Comparison of AVIRIS and Landsat ETM+ detection capabilities for burn severity. *Remote Sensing of Environment*, 92 397-408.

Vermote, E., Ray, J. 1999. MODIS Surface Reflectance User's Guide. http://modis land.gsfc.nasa.gov/MOD09/MOD09ProductInfo/MOD09Level2G250m.htm

Zhan, X., Defries, R.S., Townshend, J.R.G., DiMiceli, C.M., Hansen, M.C., Huang, C., Sohlberg, R. 2000. The 250 m global land cover change product from the Moderate Resolution Imaging Spectroradiometer of NASA's Earth Observing System. *International Journal of Remote Sensing*, 21 1433 – 1460.

6. Supporting Tables

Rule	**Value**
Exclude observations with negative reflectance values	
Exclude cloudy observations	1,2
Exclude observations with cloud shadow	1
Exclude observations with high solar zenith angle	>85 degrees
Exclude observations with highest aerosol correction	3
Choose observations with low sensor zenith angle	< 45 degrees
Choose observations with high observation coverage	> 25%
Choose observations with lowest aerosol correction	1
Calculate NDVI for all remaining observations	
Rank observations from high to low NDVI	
Exclude observations with low NDVI	>10% below the max NDVI
Choose observations with lowest cirrus detected	1
Choose observation with the lowest sensor zenith angle	minimum

Table 1 (MOD44C compositing rules)

Scheme 1	Percent Tree Cover	Scheme 2	Percent Tree Cover
Class 1	0 - 19	Class 1	0 - 18
Class 2	20 - 39	Class 2	19 - 36
Class 3	40 - 59	Class 3	37 - 54
Class 4	60 - 100	Class 4	55 - 72
		Class 5	73 - 100

Table 2 (VCC Deforestation Class Schemes)

Year 1 Time 1 compared to Year 2 Time 1
Year 1 Time 2 compared to Year 2 Time 2
Year 1 Time 3 compared to Year 2 Time 3
Year 1 Time 4 compared to Year 2 Time 4
Year 1 Time 5 compared to Year 2 Time 5
Year 1 Time 6 compared to Year 2 Time 6

Table 3 (comparison schemes for VCC Deforestation)