Global land cover classification at 1 km spatial resolution using a classification tree approach

M. C. HANSEN*, R. S. DEFRIES†, J. R. G. TOWNSHEND‡ and R. SOHLBERG

Department of Geography, University of Maryland, College Park, MD 20742, USA

Abstract. This paper reports the production of a 1 km spatial resolution land cover classification using data for 1992–1993 from the Advanced Very High Resolution Radiometer (AVHRR). This map will be included as an at-launch product of the Moderate Resolution Imaging Spectroradiometer (MODIS) to serve as an input for several algorithms requiring knowledge of land cover type. The methodology was derived from a similar effort to create a product at 8 km spatial resolution, where high resolution data sets were interpreted in order to derive a coarse-resolution training data set. A set of 37294 × 1 km pixels was used within a hierarchical tree structure to classify the AVHRR data into 12 classes. The approach taken involved a hierarchy of pair-wise class trees where a logic based on vegetation form was applied until all classes were depicted. Multi-temporal AVHRR metrics were used to predict class memberships. Minimum annual red reflectance, peak annual Normalized Difference Vegetation Index (NDVI), and minimum channel three brightness temperature were among the most used metrics. Depictions of forests and woodlands, and areas of mechanized agriculture are in general agreement with other sources of information, while classes such as low biomass agriculture and high-latitude broadleaf forest are not. Comparisons of the final product with regional digital land cover maps derived from high-resolution remotely sensed data reveal general agreement, except for apparently poor depictions of temperate pastures within areas of agriculture. Distinguishing between forest and non-forest was achieved with agreements ranging from 81 to 92% for these regional subsets. The agreements for all classes varied from an average of 65% when viewing all pixels to an average of 82% when viewing only those 1 km pixels consisting of greater than 90% one class within the high-resolution data sets.

1. Introduction

Vegetative land cover is an important variable in many Earth system processes. Many general circulation and carbon exchange models require vegetative cover as a boundary layer necessary to run the model (Sellers et al. 1997). Vegetation also represents an important natural resource for humans and other species, and

*E-mail address: mhansen@GLUE.umd.edu
†Also at Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD 20742, USA.
‡Also at Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA.
quantifying the types and extent of vegetation is important to resource management and issues regarding land cover change (Townshend 1992).

With increasing frequency, remotely sensed data sets have been used to classify global vegetative land cover. The primary goals in developing these products are to meet the needs of the modelling community and to attempt to better understand the role of human impacts on Earth systems through land cover conversions. Recent work in classifying regional, continental and global land cover has seen the application of multi-temporal remotely sensed data sets, which describe vegetation dynamics by viewing their phenological variation throughout the course of a year (Verhoef et al. 1996). Tucker et al. (1985), Townshend et al. (1987) and Stone et al. (1990) have produced continental-scale classifications of land cover using this approach. For global land cover products, DeFries and Townshend (1994) derived a one-by-one degree map and more recently an 8 km map (DeFries et al. 1998) using AVHRR data. The current global land cover products are much finer in resolution than traditional climate modellers require, although there are some who have begun to take advantage of the additional information in the depiction of landscape heterogeneity provided by finer resolutions (Dickinson 1995). As the resolutions of global data sets become finer, the ability to monitor short-term anthropogenically induced land cover changes has increased. Sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) have resolutions sufficient for global depictions of land cover change. Establishing a global baseline for land cover at 1 km is an important step in understanding how change can be depicted with newer sensors such as MODIS.

A 1 km data set employing AVHRR data has been developed based on the recommendations from the International Geosphere–Biosphere Programme (IGBP) for use in global change research (Townshend 1992). Loveland et al. (1999) have produced a 1 km resolution global land cover layer, named the IGBP DISCover product, in which each continent is classified separately and then stitched together. They used 12-monthly Normalized Difference Vegetation Index (NDVI) values in an unsupervised clustering algorithm that was supplemented with ancillary data analysis. The DISCover product has also been included as an at-launch land cover product for the MODIS sensor.

This paper describes the development of another 1 km data product to be included as another layer within the MODIS at-launch product. Building on the 8 km map and methodology of DeFries et al. (1998), this product provides an alternative 1 km land cover data set based on the individual spectral bands as well as NDVI values. The approach involves a supervised method where the entire globe is classified using a classification tree algorithm. The tree predicts class memberships from metrics derived from the same AVHRR data employed by Loveland et al. (2000), except here all five spectral bands as well as NDVI values are used. The application of the tree classifier utilizes an imposed hierarchy of vegetation form similar to that proposed and implemented by Running et al. (1994, 1995), except that the relationships between multi-spectral data and vegetation type are empirically derived.

Maps produced using satellite data have advantages over traditional ground-based maps due to the continuous coverage and internal consistency of remotely sensed data sets. A primary reason for attempting to create maps from these data sets is the potential for creating more accurate products, where the areas of disagreement between products are less than past efforts compiled from ground-based maps (DeFries and Townshend 1994a). Classifying the entire globe at once allows for the
consistent extrapolation of spectral signatures in order to improve the consistency of global land cover characterization. In the end, a limited amount of regional relabelling of pixels is required and this reflects the limitations of the present method while pointing the way for improved iterations. In the absence of independent validation data, a comparison between the University of Maryland (UMd) land cover layer with other digital land cover maps is made at the end of this paper, as well as with forest statistics produced by the United Nations Food and Agriculture Organization. These comparisons help identify potential strengths and weaknesses of the UMd product and also raise a number of issues which may help future efforts in creating improved products.

2. Data

2.1. Training data

The majority of the training data were derived via the method described by DeFries et al. (1998), using an overlay of co-registered coarse-resolution and interpreted high-resolution data sets. Previous work for the 1984 8 km product consisted of interpreting 156 images, the great majority of which were Landsat Multi-spectral Scanner System (MSS) data sets. Interpretations were aided using ancillary data sets, a list of which can be obtained, along with the 8 km data plane, from the University of Maryland website (http://www.geog.umd.edu/landcover/global-cover.html).

After overlaying the 1 km global data grid with high-resolution data, only those 1 km grid cells which were interpreted from the MSS as consisting of 100% of the cover type of interest were included in this training set. However, manipulating and analysing the 1 km data set at full resolution proved to be beyond the available computing resources, so a subset of the entire data set, including the training data, was then derived. Roughly every fifth pixel was sampled across each row and line of the data set in order to create a much reduced, but viewable and usable subset of 7205 pixels by 3122 lines. From this subset, 27031 pixels were taken from the high-resolution scenes as training sites.

Additional MSS scenes outside of the 156 original images were needed to address some shortcomings seen in the 8 km land cover product. The original scenes were selected from areas where three largely ground-based global land cover characterizations agreed (Matthews 1983, Olson et al. 1983, Wilson and Henderson-Sellers 1985). For some classes, this greatly limited the successful depiction of land cover across all latitudes. For example, the wooded grassland class, better described by its definition of 10–40% tall woody canopy cover, had training sites only within the tropical regions of the globe. Thus, the depiction of areas with 10–40% tall woody canopy within temperate and boreal zones was limited. Clearly, there are large areas outside of the tropics which have land cover fitting the description of this class. To address this issue and others like it, MSS thumbnail images were downloaded using the Global Land Information System (GLIS) from the EROS Data Center (http://edcwww.cr.usgs.gov/webglis/) and physical features were identified using the same ancillary data as were used to create the original training sites. This procedure added 10218 pixels to the database, yielding a total of 37249 training pixels.

The entire exercise of augmenting the training data was based on interpretative analysis of the 8 km product as well as some preliminary work with 1 km data. The goal of adding to the training sites was to improve upon the limitations of the 8 km map. No quantitative sampling procedure was available in order to guide the
acquisition of the training data, as no reliable *a priori* knowledge of their global distributions exists.

2.2. AVHRR data

The data used in this classification were from the AVHRR 1 km data set processed at the EROS Data Center under the guidance of the IGBP (Eidenshink and Faudeen 1994, Townshend *et al.* 1994). For this project, data were radiometrically calibrated, geo-registered to the 1 km Goode’s Interrupted Homolosine equal area projection, composited over a 10-day period using maximum NDVI values, and then atmospherically corrected for ozone and Rayleigh scattering and solar zenith angle to yield surface reflectances. From a set of twelve data layers, the following were included in this study: Channel 1 (visible red reflectance, 0.58–0.68 μm), Channel 2 (near-infrared reflectance, 0.725–1.1 μm), Channel 3 (thermal infrared, 3.55–3.93 μm), Channel 4 (thermal, 10.3–11.3 μm), Channel 5 (thermal, 11.5–12.5 μm) and the NDVI (Channel 2 − Channel 1)/(Channel 2 + Channel 1). The first twelve months produced for this data set, beginning 1 April 1992 and ending 31 March 1993, were used in this classification.

To reduce data volumes and cloud contamination, a maximum NDVI composite was created for every month, along with all five associated channel values. This, however, did not remove all noise from the data set, and a filtering of the data was performed in the context of a time series by identifying and removing data spikes (DeFries *et al.* 1998). Since there is no quality control flag for the 1 km data, each monthly value of each band was viewed in isolation and compared to the standard deviation of the remaining monthly values. Those monthly values which were greater than seven standard deviations away from the mean of the remaining eleven months were removed. The value of seven standard deviations was chosen through visual inspection of the data and found to be a conservative level which removed the most obvious spikes. The metrics calculated from the AVHRR time series are very sensitive to noisy data, including maximum and minimum annual metrics (see § 3.3), and the removal of inordinately extreme values preserves their utility. Other AVHRR processing techniques note the presence of digital counts of extreme low and high values which are not readily detectable, even in a data set production mode (Agbuj and James 1994). The attempt here was to find a simple remedy which reduced the problem of noise without removing useful data.

Preliminary work on the 1 km data revealed a number of characteristics which dictated some modifications to the classification methodology used for the 8 km product. In general, the 1 km data set appears to have more artifacts and cloud contamination than the 8 km Pathfinder data set. The 8 km data set is derived from Global Area Coverage (GAC) 4 km data (James and Kalluri 1994), which are continuously recorded onboard the National Oceanic and Atmospheric Administration (NOAA) platforms, while the 1 km data set uses Local Area Coverage (LAC) 1 km data, which must be recorded by regional receiving stations. The result for the 1 km data set is a less continuous product temporally, as receiving stations are not always operating. This yields a composite more heavily contaminated by cloud. Also, the geo-referencing for the 8 km uses onboard navigation to bin pixels, while the 1 km uses ground control points (GCPs). For any data scan at 1 km, the swath was divided into areas corresponding to the sections in the Goode projection. Any section not having a minimum number of GCPs was excluded, unlike the 8 km product, which retained all data. This also results in a less clean composite as the
number of samples entered into the compositing scheme is substantially reduced. Both compositing schemes used maximum NDVI, and then applied an atmospheric correction. Binning on the basis of NDVI results in composites being biased in the forward scatter direction due to bidirectional reflectance distribution function (BRDF) effects (Holben 1986, Cihlar et al. 1994), resulting in less clean time series, particularly for Channels 1 and 2. This also increases the presence of pixels with distorted view geometries. Misplaced scans are also present due to poor header information on the input data (Eidenshink 1998 Personal communication). While both the 8 km and 1 km data sets have many common problems, such as no BRDF correction, the aforementioned differences can have a significant effect on the production of like products, and lead to a need for somewhat different treatments of the data in producing land cover maps.

3. Classification scheme
3.1. Algorithm

A decision tree was used to classify the dependent variable of class membership using the independent variables of AVHRR metrics. Decision tree theory (Breiman et al. 1984, Quinlan 1993, Venables and Ripley 1994) has previously been used to classify remotely sensed data sets (Hansen et al. 1996, Freidl and Brodley 1997, DeFries et al. 1998), and offers some advantages over other classification methods. Trees are a non-parametric, hierarchical classifier which predicts class membership by recursively partitioning a data set into more homogeneous subsets. This procedure is followed until a perfect tree (one in which every pixel is discriminated from pixels of other classes, if possible) is created with all pure terminal nodes or until preset conditions are met for terminating the tree’s growth. The method used here is that of the S-plus statistical package (Clark and Pergibon 1992), which employs a deviance measure to split data into nodes which are more homogeneous with respect to class membership than the parent node. The reduction in deviance \( D \) is calculated as:

\[
D = D_s - D_t - D_u
\]

where \( s \) represents the parent node, and \( t \) and \( u \) are the splits from \( s \). Right and left splits along the digital counts for all metrics are examined. When \( D \) is maximized, the best split is found, and the data are divided at that digital count and the process is repeated on the two new nodes of the tree. The deviance for nodes is calculated from the equation

\[
D_i = -2 \sum n_{ik} \log p_{ik}
\]

where \( n \) is the number of pixels in class \( k \) in node \( i \) and \( p \) is the probability distribution of class \( k \) in node \( i \).

Trees usually overfit the training data and a pruning procedure is needed to better generalize the relationships between the dependent and independent variables. In other words, the tree can fit the training data too well by growing on noise and errors. To avoid this, a pruning procedure is employed to better generalize the predictive ability of the tree. Pruning is performed by splitting the data into two sets and using one to grow the tree and the other to prune it by eliminating nodes which increase errors within the pruning data set. In this study, pruning was performed by visual interpretation due to problems inherent in both the training and the AVHRR data, as will be discussed shortly.

Because trees are non-parametric and nonlinear, multiple terminal nodes are
created for classes which have multi-modal distributions in spectral space. This allows for the clearer depiction of the intraclass variability which exists at the global scale. Trees are also useful for identifying classes which represent subsets of a continuous parameter, such as tree canopy for wooded grasslands, woodlands and forests. Trees operate not on statistics of central tendency, but along the thresholds in multi-spectral space which best characterize boundaries between classes. Hansen et al. (1996) found that a single classification tree threshold was superior to a maximum likelihood classifier in identifying tall from short global vegetation. However, since the training pixel counts are used to estimate probabilities, larger classes can be overemphasized in optimizing the splits and in the assignments of terminal nodes. Smaller classes are easily identified within trees if they are dominant in any part of the multi-spectral space. But, if smaller classes are mixed with larger classes, the smaller class can be lost via the assignment of terminal nodes to the classes with a dominant proportional representation.

The hierarchical nature of trees yields explicit relationships between the dependent variable, class membership, and the independent variables of multi-temporal metrics. In so doing, it allows for a reader biophysical interpretation through the description of vegetation characteristics such as height of vegetation and canopy closure. This ease of interpretation is unique among popular remote sensing classifiers and allows for the input of an expert analyst in correcting splits associated with faulty or contradictory training data.

3.2. Classes

The IGBP has developed a list of classes for use within global change research and to which the 1 km MODIS at-launch product and post-launch products will conform (Rasool 1992). The UMd class definitions closely fit this scheme and are listed along with corresponding IGBP classes in table 1. The 1 km training areas were derived from 156 high-resolution scenes. These were originally interpreted for the 8 km map which employed a classification scheme for use with the Simple Biosphere (SiB) general circulation model (Sellers et al. 1997). The SiB scheme does not have agricultural mosaic, wetlands or snow and ice classes. As a result, the mosaic and wetlands classes are absent from this classification, while the IGBP snow and ice cover class is included in the bare ground class. The urban and built-up class was taken directly from the EROS Data Center (EDC) 1 km IGBP classification by Loveland et al. (2000), which was in turn obtained from the Digital Chart of the World (Danko 1992). The water layer was taken from a preliminary water mask made for the MODIS sensor in a sinusoidal projection and reprojected into the Interrupted Goode Homolosine projection for use with this project. The SiB mosses and lichens class does not exist within the IGBP scheme, and scenes from this class were reinterpreted to extract other covers where possible. More subtle differences between the UMd and the IGBP schemes, such as height of trees, are irreconcilable and differ largely because of the definitions used by the ancillary sources in interpreting the high-resolution data.

3.3. AVHRR metrics

A set of 41 metrics was created for input into the decision tree. The first 29 metrics were created from values associated with the eight greenest months of the year. These metrics differ from those used to derive the 8 km map. The 8 km metrics used all twelve months of 1984 Pathfinder 8 km data in the classification, and this
Table 1. Comparison of University of Maryland class definitions to the IGBP-DIS definitions.

<table>
<thead>
<tr>
<th>University of Maryland vegetation classes</th>
<th>IGBP-DIS Land Cover Working Group vegetation classes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cover types in common with IGBP</strong></td>
<td><strong>Cover types in common with UmD</strong></td>
</tr>
<tr>
<td><strong>Evergreen Needleleaf Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 5 m. Almost all trees remain green all year. Canopy is never without green foliage.</td>
<td><strong>Evergreen Needleleaf Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 2 m. Almost all trees remain green all year. Canopy is never without green foliage.</td>
</tr>
<tr>
<td><strong>Evergreen Broadleaf Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 5 m. Almost all trees remain green all year. Canopy is never without green foliage.</td>
<td><strong>Evergreen Broadleaf Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 2 m. Almost all trees remain green all year. Canopy is never without green foliage.</td>
</tr>
<tr>
<td><strong>Deciduous Needleleaf Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 5 m. Trees shed their leaves simultaneously in response to cold seasons.</td>
<td><strong>Deciduous Needleleaf Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 2 m. Consists of seasonal needleleaf tree communities with an annual cycle of leaf-on and leaf-off periods.</td>
</tr>
<tr>
<td><strong>Deciduous Broadleaf Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 5 m. Trees shed their leaves simultaneously in response to dry or cold seasons.</td>
<td><strong>Deciduous Broadleaf Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 2 m. Consists of seasonal broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods.</td>
</tr>
<tr>
<td><strong>Mixed Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 5 m. Consists of tree communities with interspersed mixtures or mosaics of needleleaf and broadleaf forest types. Neither type has &lt; 25% or &gt; 75% landscape coverage.</td>
<td><strong>Mixed Forests:</strong> lands dominated by trees with a per cent canopy cover &gt; 60% and height exceeding 2 m. Consists of seasonal broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods.</td>
</tr>
<tr>
<td><strong>Woodlands:</strong> lands with herbaceous or woody understories and tree canopy cover of &gt; 40% and &lt; 60%. Trees exceed 5 m in height and can be either evergreen or deciduous.</td>
<td><strong>Woodly savannas:</strong> lands with herbaceous and other understory systems, and with forest canopy between 30–60%. The forest cover height exceeds 2 m.</td>
</tr>
<tr>
<td><strong>Wooded Grasslands/Shrublands:</strong> lands with herbaceous or woody understories and tree canopy cover of &gt; 10% and &lt; 40%. Trees exceed 5 m in height and can be either evergreen or deciduous.</td>
<td><strong>Savannas:</strong> lands with herbaceous and other understory systems, and with forest canopy between 10–30%. The forest cover height exceeds 2 m.</td>
</tr>
<tr>
<td>University of Maryland vegetation classes</td>
<td>IGBP-DIS Land Cover Working Group vegetation classes</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>Cover types in common with IGBP</td>
<td>Cover types in common with UMd</td>
</tr>
<tr>
<td><strong>Closed Bushlands or Shrublands</strong>: lands dominated by bushes or shrubs. Bush and shrub per cent canopy cover is &gt;40%. Bushes do not exceed 5 m in height. Shrubs or bushes can be either evergreen or deciduous. Tree canopy cover is &lt;10%. The remaining cover is either barren or herbaceous.</td>
<td><strong>Closed Shrublands</strong>: lands with woody vegetation less than 2 m tall and with shrub-canopy cover &gt;60%. The shrub foliage can be either evergreen or deciduous.</td>
</tr>
<tr>
<td><strong>Open Shrublands</strong>: lands dominated by shrubs. Shrub canopy cover is &gt;10% and &lt;40%. Shrubs do not exceed 2 m in height and can be either evergreen or deciduous. The remaining cover is either barren or of annual herbaceous type.</td>
<td><strong>Open Shrublands</strong>: lands with woody vegetation less than 2 m tall and with shrub canopy cover between 10–60%. The shrub foliage can be either evergreen or deciduous.</td>
</tr>
<tr>
<td><strong>Grasslands</strong>: lands with continuous herbaceous cover and &lt;10% tree or shrub canopy cover.</td>
<td><strong>Grasslands</strong>: lands with herbaceous types of cover. Tree and shrub cover is less than 10%.</td>
</tr>
<tr>
<td><strong>Croplands</strong>: lands with &gt;80% of the landscape covered in crop-producing fields. Note that perennial woody crops will be classified as the appropriate forest or shrubs land cover type.</td>
<td><strong>Croplands</strong>: lands covered with temporary crops followed by harvest and a bare soil period (e.g. single and multiple cropping systems). Note that perennial woody crops will be classified as the appropriate forest or shrubs land cover type.</td>
</tr>
<tr>
<td><strong>Barren</strong>: lands of exposed soil, sand, rocks, snow or ice which never have more than 10% vegetated cover during any time of the year.</td>
<td><strong>Barren</strong>: lands of exposed soil, sand, rocks or snow and never has more than 10% vegetated cover during any time of the year.</td>
</tr>
<tr>
<td><strong>Urban and Built-up</strong>: land covered by buildings and other man-made structures. Note that this class will not be mapped from the AVHRR imagery but will be developed from the populated places layer that is part of the Digital Chart of the World (Danko 1992).</td>
<td><strong>Urban and Built-up</strong>: land covered by buildings and other man-made structures. Note that this class will not be mapped from the AVHRR imagery but will be developed from the populated places layer that is part of the Digital Chart of the World (Danko 1992).</td>
</tr>
<tr>
<td><strong>Water bodies</strong>: oceans, seas, lakes, reservoirs, and rivers. Can be either fresh or salt water.</td>
<td><strong>Water bodies</strong>: oceans, seas, lakes, reservoirs, and rivers. Can be either fresh or salt water.</td>
</tr>
<tr>
<td><strong>Covers not in common with UMd</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Permanent Wetlands</strong>: Lands with a permanent mixture of water and herbaceous or woody vegetation that cover extensive areas. The vegetation can be present in either salt, brackish, or fresh water.</td>
<td></td>
</tr>
<tr>
<td><strong>Cropland/Natural Vegetation Mosaics</strong>: lands with a mosaic of croplands, forest, shrublands, and grasslands in which no one component comprises more than 60% of the landscape.</td>
<td></td>
</tr>
<tr>
<td><strong>Snow and Ice</strong>: lands under snow and/or ice cover throughout the year.</td>
<td></td>
</tr>
</tbody>
</table>
produced a number of nodes associated with snow cover. Snow cover, especially relating to the distribution and number of training pixels within and without the snow area, can produce undesired results. By binning all metrics on only the eight greenest months, snow effects are largely limited to those places with perpetual snow and ice cover and very high-latitudes, while still retaining most of the seasonal variability associated with vegetation phenology. The eight greenest months are not necessarily consecutive, but represent the eight months with the clearest view of green vegetation. In this manner, globally applicable, timing-insensitive metrics with minimized cloud presence are created. The metrics used included maximum, minimum, mean and amplitudes for all bands associated with the eight greenest months. Individual band values associated with peak greenness were also derived.

Using only eight months of data means that for areas like the tropics and much of the temperate zone, four months of useful data were thrown away. To try and recapture some of this information, metrics were binned on the four warmest months, as measured by Channel 4, and two additional metrics per band and for NDVI were calculated. These were means associated with the four warmest months and individual values occurring at maximum Channel 4 temperature. The four warmest months were found to be associated with the dry season, or senescent phase of much tropical vegetation. By compositing on these values, data not used in the eight greenest months can be included for some areas without introducing snow values at high latitudes and elevations. The metrics derived from the 1992–1993 year for bands 1–5 and NDVI are listed in table 2.

3.4. Procedure

The classification procedure followed that of the 8 km product except for the use of a cascading two-class hierarchy of trees in implementing the classification tree algorithm. An initial attempt at using the previous methodology revealed the inability to create a single simplified tree such as the 8 km tree which described the globe in 57 nodes. One potential reason for this is the increased heterogeneity of the Earth’s surface at 1 km which precludes the creation of a single, simple global tree. Secondly, as mentioned previously, 1 km data are more cloud-contaminated than the 8 km data due to different recording procedures, geo-registering and compositing techniques. This results in a more complex tree structure than that derived for the relatively cleaner 8 km data. Thus, the decision was taken to create an imposed tree hierarchy, shown in figure 1, somewhat similar to that used by Running et al. (1995). In this manner, only two classes are depicted within any single tree, allowing for a simplified, structured approach and easier interpretation of the results. The original training pixels were run through the successive trees and a preliminary result was obtained. Pruning based on visual interpretation of this preliminary global map was then applied where nodes were accepted or rejected based on their global geographic distributions. From this, a subset of pixels associated with approved nodes was created. These pixels were then re-run through the tree structure and a final tree was derived. These trees were again pruned based on visual interpretation to produce the final map.

For both the UMdl 1 km and 8 km maps, an automated classification algorithm is employed to depict land cover, but obvious errors in the product make it necessary to apply an interpretative step. Figure 2 outlines the steps in the procedure along with some of the sources of error within this product which mandate that an interpretative pruning be applied. Cloud contamination, bad scan lines, missing data, geometric
Table 2. Metrics employed in the production of the University of Maryland 1 km product using AVHRR data from April 1992 to March 1993.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>maximum NDVI value</td>
</tr>
<tr>
<td>2</td>
<td>minimum NDVI value of 8 greenest months</td>
</tr>
<tr>
<td>3</td>
<td>mean NDVI value of 8 greenest months</td>
</tr>
<tr>
<td>4</td>
<td>amplitude of NDVI over 8 greenest months</td>
</tr>
<tr>
<td>5</td>
<td>mean NDVI value of 4 warmest months</td>
</tr>
<tr>
<td>6</td>
<td>NDVI value of warmest month</td>
</tr>
<tr>
<td>7</td>
<td>maximum channel 1 value of 8 greenest months</td>
</tr>
<tr>
<td>8</td>
<td>minimum channel 1 value of 8 greenest months</td>
</tr>
<tr>
<td>9</td>
<td>mean channel 1 value of 8 greenest months</td>
</tr>
<tr>
<td>10</td>
<td>amplitude of channel 1 over 8 greenest months</td>
</tr>
<tr>
<td>11</td>
<td>channel 1 value from month of maximum NDVI</td>
</tr>
<tr>
<td>12</td>
<td>mean channel 1 value of 4 warmest months</td>
</tr>
<tr>
<td>13</td>
<td>channel 1 value of warmest month</td>
</tr>
<tr>
<td>14</td>
<td>maximum channel 2 value of 8 greenest months</td>
</tr>
<tr>
<td>15</td>
<td>minimum channel 2 value of 8 greenest months</td>
</tr>
<tr>
<td>16</td>
<td>mean channel 2 value of 8 greenest months</td>
</tr>
<tr>
<td>17</td>
<td>amplitude of channel 2 over 8 greenest months</td>
</tr>
<tr>
<td>18</td>
<td>channel 2 value from month of maximum NDVI</td>
</tr>
<tr>
<td>19</td>
<td>mean channel 2 value of 4 warmest months</td>
</tr>
<tr>
<td>20</td>
<td>channel 2 value of warmest month</td>
</tr>
<tr>
<td>21</td>
<td>maximum channel 3 value of 8 greenest months</td>
</tr>
<tr>
<td>22</td>
<td>minimum channel 3 value of 8 greenest months</td>
</tr>
<tr>
<td>23</td>
<td>mean channel 3 value of 8 greenest months</td>
</tr>
<tr>
<td>24</td>
<td>amplitude of channel 3 over 8 greenest months</td>
</tr>
<tr>
<td>25</td>
<td>channel 3 value from month of maximum NDVI</td>
</tr>
<tr>
<td>26</td>
<td>mean channel 3 value of 4 warmest months</td>
</tr>
<tr>
<td>27</td>
<td>channel 3 value of warmest month</td>
</tr>
<tr>
<td>28</td>
<td>maximum channel 4 value of 8 greenest months</td>
</tr>
<tr>
<td>29</td>
<td>minimum channel 4 value of 8 greenest months</td>
</tr>
<tr>
<td>30</td>
<td>mean channel 4 value of 8 greenest months</td>
</tr>
<tr>
<td>31</td>
<td>amplitude of channel 4 over 8 greenest months</td>
</tr>
<tr>
<td>32</td>
<td>channel 4 value from month of maximum NDVI</td>
</tr>
<tr>
<td>33</td>
<td>mean channel 4 value of 4 warmest months</td>
</tr>
<tr>
<td>34</td>
<td>channel 4 value of warmest month</td>
</tr>
<tr>
<td>35</td>
<td>maximum channel 5 value of 8 greenest months</td>
</tr>
<tr>
<td>36</td>
<td>minimum channel 5 value of 8 greenest months</td>
</tr>
<tr>
<td>37</td>
<td>mean channel 5 value of 8 greenest months</td>
</tr>
<tr>
<td>38</td>
<td>amplitude of channel 5 over 8 greenest months</td>
</tr>
<tr>
<td>39</td>
<td>channel 5 value from month of maximum NDVI</td>
</tr>
<tr>
<td>40</td>
<td>mean channel 5 value of 4 warmest months</td>
</tr>
<tr>
<td>41</td>
<td>channel 5 value of warmest month</td>
</tr>
</tbody>
</table>

misregistration and incorrect ancillary data can lead to the production of anomalous nodes which cannot be eliminated except by a subjective step. For this study, pruning was performed through manually clipping nodes, relabelling nodes and finding alternative splits for nodes which yielded depictions of land cover not consistent with known global distributions. Although the derivation of the trees cannot be duplicated, the application of these trees will result in the same map product. In this sense, the procedure is reproducible given the metrics and the trees, but the product reproduced represents a mix of objectively and subjectively derived relationships.
Efforts are underway to improve the ability to derive land cover products more objectively by applying more sophisticated algorithms for growing and pruning trees (Quinlan 1993). Upon application of the final trees, a regional relabelling is performed on pixels which do not agree with known global geographic distributions.

4. Results

After the initial run of the training data through the tree hierarchy and pruning procedure, a subset of 26,208 pixels was taken to create the final classification tree structure. All trees were created using the 7,205 by 3,122 pixel subset of the 1 km data. The final trees were then run on the entire 1 km data set. No sieve was applied.

Among the most frequently used metrics for all trees were minimum annual red reflectance and maximum annual NDVI. Figure 3 shows subsets of boreal, temperate and tropical areas for these metrics with the same spectral enhancements. These subset windows will be used to illustrate certain qualities of the product as the discussion of the tree hierarchy is developed.

4.1. Hierarchical classification tree

4.1.1. Vegetated/bare ground tree

The first tree was a vegetated/non-vegetated tree used to classify bare ground (figure 4). For all trees depicted, only the dominant nodes accounting for over 5% of the respective class land areas are shown. A simple maximum annual NDVI threshold provides greatest initial discrimination for the vegetated/non-vegetated tree. In figure 5(a), the land areas associated with the dominant nodes for the tree are highlighted. Areas such as volcanic features in the Sahara Desert are not easily discriminated with a single split (shown in black as lesser nodes), and extra nodes
were required that largely accounted for these barren ground subtypes. Barren areas with a maximum NDVI greater than 0.155 are associated with a misplaced AVHRR 1 km data set swath in the Atacama desert. The rest of the black area is associated with areas having a peak NDVI less than 0.155. Outcrops in the Sahara and other places are confused with sparse grasslands in China, and additional splits are needed to depict them. Volcanic outcrops of the Sahara have been analysed before in efforts to account for background soil reflectances which make certain areas spectrally similar to sparsely vegetated ones (Huete and Tucker 1991). The results here reflect
Figure 3. Subsets for (a) boreal Russia, (b) temperate United States and (c) tropical West Africa. On the left-hand side, red is maximum annual NDVI and cyan is minimum annual red reflectance. All three images had the same linear stretch applied to both bands. On the right-hand side: 1 = needleleaf evergreen forest; 2 = broadleaf evergreen forest; 3 = needleleaf deciduous forest; 4 = broadleaf deciduous forest; 5 = mixed forest; 6 = woodland; 7 = wooded grassland; 8 = closed shrubland; 9 = open shrubland; 10 = grassland; 11 = cropland; 12 = bare ground; 14 = urban/built-up areas. The scale bar represents 300 km.
The difficulty in separating these areas using NDVI alone, and different spectral information through the use of additional splits is needed to characterize them.

4.1.2. Tall/short vegetation tree

Use of this tree was intended to separate woodlands and forests with nearly closed tall canopies from open parklands, shrublands, croplands and herbaceous covers. Figure 6 shows the tree's structure and figure 5(b) where the dominant nodes map globally. In general, low minimum annual visible red values (< 5.3% reflectance) successfully discriminated woodlands and forests. Exceptions to this relationship included non-woody areas with water present, such as inundated grasslands, areas of rice production and other wetland formations. Commission errors of woody classes are associated with some non-woody areas for these types of land cover. Many tropical inundated grasslands were confused with needleleaf evergreen stands in the visible and near-infrared, and temperature values were used to separate the two. This, however, did not resolve the characterization of similar wetlands at higher latitudes.

Some croplands, particularly in the midwest of the United States, also had very dark minimum red values and this created confusion between croplands and woodlands. Part of this problem may be due to bad data in the 1 km data set. For example, figure 3 (b) shows the minimum annual Channel 1 metric for the Midwest USA. The line in the image is the Mollweide/Sinusoidal boundary of the Goode projection. The dark northern portion creates a node in the woody/non-woody tree specific to itself, indicating a possible problem with the data set. Figure 5(b) shows this area in black as one of the lesser nodes not easily discriminated along with the rest of the woody and non-woody types. This type of problem is hard to isolate in a global approach and can create undesirable results.

In general, along forest/non-forest boundaries it is apparent that the extent of
Figure 5. (a) Results of the vegetated/non-vegetated classification tree. Red = largest node for vegetated class; cyan = largest node for non-vegetated class; black = lesser nodes in the tree. (b) Results of tall/short vegetation tree. Red = largest node for tall vegetation, orange = second largest, yellow = third largest; cyan = largest node for short vegetation; black = lesser nodes in the tree; grey = bare ground class from previous tree. (c) Results of forest/woodland tree. Red = largest node for forest, orange = second largest, yellow = third largest; cyan = largest node for woodland, green = second largest; black = lesser nodes; dark grey = bare ground class; light grey = short vegetation from previous tree. Refer to figures 4, 6 and 7 to view node paths.
Figure 6. Tall/short vegetation classification tree. Only those nodes which account for 5% or more of the respective class totals are shown. Paths to lesser nodes are shown with ~. The text in the ellipses gives the metrics used for the split, with the left-hand side less than the value indicated and the right-hand side greater. Metrics beginning with t, such as tmmaxch5 are metrics derived from the four warmest months. Other metrics, such as minch1, are derived from the eight greenest months. For example, tmmaxch5 stands for the mean Channel 5 value of the four warmest months. Minch1 stands for the minimum Channel 1 value of the eight greenest months. Values for Channels 1 and 2 are in per cent reflectance. Values for Channels 3, 4 and 5 are in degrees Kelvin.

forest stands is exaggerated. For example, clearings of grasslands within the Amazon basin are reduced in size as if a buffer of forest were added to the boundary. This is the result of binning on maximum NDVI, retaining data from low scan angles and georegistration inaccuracies. When combining these effects, the greener class along any class boundary is usually overemphasized. This loss of heterogeneity and bias of dominant classes has been discussed by others (Cushnie 1987, Moody and Woodcock 1994) and users should be aware of this problem.

One noticeable area lacking in woodlands in this product is west Africa. While woodlands are mapped, they are not present in this product as much as ancillary data indicate they should be (http://www.geog.umd.edu/landcover/global-cover/global-resources.html). Figure 3(c) shows a two-band composite of an area in west Africa. In general, west African woodlands are considerably brighter than other woody areas, such as the miombo woodlands of southern Africa, or the Gran Chaco of Argentina, possibly reflecting a longer history of anthropogenic disturbance in the region. Low minimum annual red reflectance for tall woody areas represents the combined effects of canopy shadowing and chlorophyll absorption. For most of west Africa, high minimum annual red reflectance values limit the areal extent of forests
Global and regional land cover characterization

and woodlands. Figure 3(c) also shows the problem of persistent cloud cover along the coast of West Africa in affecting the minimum Channel 1 metric. The classification of grasslands along the western coast of the image is suspected of being largely the cause of cloud presence in all monthly composites. This area, according to ancillary information, has considerably more forested land.

It is of interest that the first woody/non-woody split of minimum annual red reflectance was nearly the same as that derived for our 8 km land cover classification using 1984 Pathfinder data (5.35% compared to 5.38%). The potential reproducibility of tree splits between data sets and over time has implications not only for land cover mapping, but also for change detection. A reproducible tree structure over time would allow for the depiction of spectral migration and change.

4.1.3. Forest/woodland tree

Woodlands were also distinguished from forests based on minimum visible red values. Figure 7 shows the forest, woodland classification tree while figure 5(c) shows the spatial extent of the dominant nodes. Some tropical woodlands were as dark

```
area totals
forest: 22,010,513 sq. km.
woodland: 7,436,875 sq. km.
```

```plaintext
minch1<> 4.55
```

```plaintext
tmeanch5<> 297.36
```

```plaintext
meanndvi<> 0.615
```

```plaintext
maxndvi<> 0.705
```

```plaintext
minch1<> 3.85
```

```plaintext
maxndvi<> 0.545
```

```plaintext
minch1<> 4.35
```

```plaintext
maxndvi<> 0.505
```

```plaintext
forest 78.0%
```

```plaintext
woodland 56.2%
```

```plaintext
forest 29.4%
```

```plaintext
forest 6.9%
```

```plaintext
woodland 10.8%
```

Figure 7. Forest/woodland classification tree. Only those nodes which account for 5% or more of the respective class totals are shown. Paths to lesser nodes are shown with ~. The text in the ellipses gives the metrics used for the split, with the left-hand side less than the value indicated and the right-hand side greater. Metrics beginning with t, such as tmeanch3 are metrics derived from the four warmest months. Other metrics, such as minch1, are derived from the eight greenest months. For example, tmeanch3 stands for the mean Channel 3 value of the four warmest months. Meanch1 stands for the mean Channel 1 value of the eight greenest months. Values for Channels 1 and 2 are in per cent reflectance. Values for Channels 3, 4 and 5 are in degrees Kelvin.
and green as tropical forests, and temperature bands such as the mean of the four Channel 5 values associated with the warmest four months of the year were used to stratify these areas. This metric helped to create a fairly clear forest/woodland boundary in central Africa, but may have increased confusion between some more seasonal tropical forests and adjacent woodlands. Ancillary data from Asia referring to deciduous forests and African sources for miombo woodlands are not reconciled in the metrics used here and errors between tropical seasonal forest and woodlands are suspected to exist in areas such as Asia and west Africa. The depiction of woodlands revealed the heterogeneity of areas such as the boreal forest, and within fragmented forest/agricultural mosaics such as the south-eastern United States. The global approach of applying a set of universal splits revealed a lack of true forest in a number of areas such as the Atlas Mountains and the hills of east-central India in contradiction to ancillary sources which depict extensive forest stands.

4.1.4. **Remaining classification trees**

Descriptions of the remaining trees can be found in the metadata of the 1 km product at the website cited above. Important aspects include the following.

1. **Mixed/pure leaf type forest classification tree**:  
   - Mixed forests defined within 45–60 degrees latitude are separated largely by the mean Channel 4 of the eight greenest months metric and maximum NDVI;  
   - Maximum NDVI is repeatedly used;  
   - Many high-latitude broadleaf forests are classified as mixed forest due to lower maximum NDVI values;  
   - Broadleaf forests mixed with non-forest covers are often labelled mixed forest, creating a buffering effect around core broadleaf areas (see figure 3(a)).

2. **Broadleaf/needleleaf forest classification tree**:  
   - Minimum Channel 3 separates tropical broadleaf from extratropical needleleaf forest;  
   - Maximum NDVI separates temperate broadleaf from needleleaf forest;  
   - Lush evergreen needleleaf areas, such as that found in the Pacific Northwest, and Araucaria forests in South America, are confused with broadleaf evergreen forest;  
   - Eucalyptus forests in Australia and elsewhere map as needleleaf evergreen forests.

3. **Needleleaf evergreen/deciduous forest classification tree**:  
   - Amplitude of NDVI largely separates these leaf types;  
   - Needleleaf deciduous forests confused with mixed forests.

4. **Broadleaf evergreen/deciduous forest classification tree**:  
   - Minimum Channel 3 separates temperate from tropical broadleaf forests;  
   - Mean of the four Channel 5 values associated with the four warmest months separates tropical deciduous from tropical broadleaf forests;  
   - Amplitude of NDVI separates temperate deciduous forests from other evergreen forests not separated by the Channel 3 split.
(5) Sparse trees (wooded grassland)/croplands, grass or shrubs classification tree:
- high intraclass spectral variability, largest wooded grassland node accounts for only 17% of class total;
- minimum red reflectance, infrared reflectance at peak greenness and minimum Channel 3 are the metrics of the first three splits;
- most difficult class to map since it represents a wide range of partial woody covers, for example, scrub savannas, boreal transitional woodlands and crop/forest mosaics.

(6) Croplands/shrubs and grass classification tree:
- NDVI and near-infrared metrics are the first two used in this tree;
- mechanized agriculture is in general agreement with ancillary data;
- agriculture in developing nations poorly depicted, as is all low biomass agriculture, due to the difficulty in separating cropping from natural background phenologies and errors of omission and commission exist for many areas.

(7) Grass/shrubs classification tree:
- mean red reflectance and Channel 5 and NDVI means for the warmest four months are used in the first three splits;
- pastures within temperate cropping areas are not depicted;
- semi-arid and very high-latitude grasslands not well depicted.

(8) Open/closed shrubs classification tree:
- temperature and NDVI metrics separate shrub classes;
- confusion between these two classes and grasslands suspected to exist as background soil reflectances make discrimination difficult (Huete and Tucker 1991).

5. Regional relabelling

The last step was a regionally based reassignment of obvious inaccuracies which was performed to remove clearly erroneous results. This modification of the product changed only 0.67% of the total land area, and represents areas spectrally inseparable within the present training data signatures. The following is a summary of classes which were remapped for this purpose: above boreal zone agriculture mapped to grassland; exterior to Siberia deciduous needleleaf forest mapped to mixed forest; needleleaf evergreen forest in Australia mapped to broadleaf evergreen forest; needleleaf forest in humid tropical basins such as the Amazon and the Congo mapped to woodland; shrub classes on volcanic outcrops in the Sahara mapped to bare ground; evergreen broadleaf forest in temperate latitudes mapped to evergreen needleleaf forest; extensive agriculture on the Tibetan plateau mapped to grassland; broadleaf evergreen forest in the miombo belt of southern Africa mapped to woodland; and two reassignments of classes due to misplaced AVHRR swaths. Table 3 shows the number of pixels per class changed in this manner. Needleleaf deciduous forest is by far the most affected class. Most of the pixels changed for this class represent mixed forest pixels. One possible explanation for the inability of this approach to cleanly delineate needleleaf deciduous forest may be the possible presence of broadleaf forest within the needleleaf deciduous training sites. When viewing maximum NDVI values for the deciduous versus evergreen needleleaf forests, deciduous forest has a median value of 0.64 compared with an evergreen value of 0.61.
Table 3. Remapping of likely misclassifications based on application of regional rules. Shown are remapped area and remapped area as a per cent of total class area as predicted by the classification tree.

<table>
<thead>
<tr>
<th>Class</th>
<th>Remapped area (km$^2$)</th>
<th>% remapped/class total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Needleleaf Evergreen Forest</td>
<td>309747</td>
<td>5.54</td>
</tr>
<tr>
<td>Broadleaf Evergreen Forest</td>
<td>121271</td>
<td>1.08</td>
</tr>
<tr>
<td>Needleleaf Deciduous Forest</td>
<td>174047</td>
<td>23.41</td>
</tr>
<tr>
<td>Broadleaf Deciduous Forest</td>
<td>51238</td>
<td>2.84</td>
</tr>
<tr>
<td>Mixed Forest</td>
<td>53988</td>
<td>1.62</td>
</tr>
<tr>
<td>Woodland</td>
<td>27240</td>
<td>0.16</td>
</tr>
<tr>
<td>Wooded Grassland</td>
<td>31993</td>
<td>0.14</td>
</tr>
<tr>
<td>Closed Shrublands</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Open Shrublands</td>
<td>37400</td>
<td>0.21</td>
</tr>
<tr>
<td>Grassland</td>
<td>224</td>
<td>0.00</td>
</tr>
<tr>
<td>Cropland</td>
<td>150404</td>
<td>1.33</td>
</tr>
<tr>
<td>Bare Ground</td>
<td>11747</td>
<td>0.04</td>
</tr>
<tr>
<td>Urban and Built-up</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Global Totals</td>
<td>969299</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Table 4. Total class areas for the University of Maryland 1 km product.

<table>
<thead>
<tr>
<th>Class</th>
<th>Area (km$^2$)</th>
<th>Per cent area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Needleleaf Evergreen Forest</td>
<td>5277925</td>
<td>3.67</td>
</tr>
<tr>
<td>Broadleaf Evergreen Forest</td>
<td>11138639</td>
<td>7.74</td>
</tr>
<tr>
<td>Needleleaf Deciduous Forest</td>
<td>569299</td>
<td>0.40</td>
</tr>
<tr>
<td>Broadleaf Deciduous Forest</td>
<td>1752105</td>
<td>1.22</td>
</tr>
<tr>
<td>Mixed Forest</td>
<td>3272545</td>
<td>2.28</td>
</tr>
<tr>
<td>Woodland</td>
<td>16533042</td>
<td>11.50</td>
</tr>
<tr>
<td>Wooded Grassland</td>
<td>22653618</td>
<td>15.75</td>
</tr>
<tr>
<td>Closed Shrublands</td>
<td>7436875</td>
<td>5.17</td>
</tr>
<tr>
<td>Open Shrublands</td>
<td>17938741</td>
<td>12.47</td>
</tr>
<tr>
<td>Grassland</td>
<td>12382238</td>
<td>8.61</td>
</tr>
<tr>
<td>Cropland</td>
<td>11126625</td>
<td>7.74</td>
</tr>
<tr>
<td>Bare Ground</td>
<td>33483362</td>
<td>23.28</td>
</tr>
<tr>
<td>Urban and Built-up</td>
<td>260092</td>
<td>0.18</td>
</tr>
<tr>
<td>Total</td>
<td>143825106</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Higher peak greenness may be an indicator of the presence of broadleaf types within this class.

6. Final global land cover classification

The areal extent of each class is shown in table 4 and the final product in figure 8. The comparison of these totals to the 1984 8 km product can be seen in figure 9. Note the marked increase in wooded grassland in the 1 km product and the corresponding decrease in more forests and woodlands. Beyond the addition of more intermediate woody training sites, the nature of the resampling in producing the GAC and 8 km Pathfinder data would also make an 8 km product more forested. Resampling of GAC data has been studied and shown to be biased towards the greenest covers and reducing spatial heterogeneity (Justice et al. 1989). This would tend to smooth mosaic areas and reduce the expression of intermediate woody
Figure 8. Final classified product: 1 = needleleaf evergreen forest; 2 = broadleaf evergreen forest; 3 = needleleaf deciduous forest; 4 = broadleaf deciduous forest; 5 = mixed forest; 6 = woodland; 7 = wooded grassland; 8 = closed shrubland; 9 = open shrubland; 10 = grassland; 11 = cropland; 12 = bare ground; 14 = urban and built-up.
covers. The most greatly reduced forest in terms of area is the needleleaf evergreen class. This is the result of the addition of more open canopy needleleaf training areas delineated in order to reduce the dominance of this class in certain regions of the 8 km product. The shrub classes appear to disagree because the 8 km product includes a class for mosses and lichens (tundra), and this area is mapped mostly as shrublands in the 1 km map.

7. Reproduction of training data

All of the training data were used in the production of the trees. Since no independent data yet exist for validating global data sets, it is useful to examine the results of the product compared to the training pixels. The majority of the confusion within the training data relates to physiognomically similar class types and classes representing mixed assemblages. There is relatively little confusion between core classes representing dominant vegetation forms and forest types. For example, the training agreement when viewing the confusion between only the classes of evergreen needleleaf and broadleaf forests, deciduous needleleaf and broadleaf forests, shrubland, grassland, cropland and bare ground is 88% (positively identified pixels of these classes/positively identified pixels of these classes + errors only across these classes). By adding the mixed forest class, woodlands, wooded grasslands, and open and closed shrublands, significant additional confusion results. The total training accuracy for all classes is reduced to 69% (total positively identified pixels for all classes/total number of training pixels). The accuracies for all classes are shown in figure 10. The best training accuracies are for the most homogeneous cover types such as bare ground (99% = positively identified bare ground pixels/total number of bare ground training pixels), open shrubs (84%) and broadleaf evergreen forest (80%). However, nearly one-third (32%) of mislabelled pixels are errors of commission or omission for the woodland class. Over one-half (55%) of training errors are associated with either woodland or wooded grassland pixels. Only one-third (33%) of all errors represent confusion between the aforementioned core classes.

The mixed class types are particularly problematic due to the reliance on ancillary data sets in delineating training sites. An example of forest/woodland confusion can be taken from a set of pixels interpreted from an MSS scene covering part of eastern
India. All of the pixels from the Chota Nagpur Plateau area were characterized as broadleaf deciduous forest but were classified as woodland in the final product due to the fact that within the metrics they are spectrally most similar to other woodland training sites. The original interpretation of the MSS scene was based on ancillary map data depicting the area as forest, but pixels from this scene have a mean minimum annual red reflectance of 5.1% compared to 3.8% for other global forest training pixels. Thus, within the tree structures, these pixels were labelled as woodland and not forest. A single misclassified scene such as this one can greatly reduce training accuracies. The pixels from this Indian scene represent 17% of all deciduous broadleaf forest training pixels and are counted as training errors.

The training accuracy numbers are less than that achieved in our 8 km effort, which had an overall training accuracy, for a set-aside data set, of 81%. This reflects the increased heterogeneity of the Earth’s surface at 1 km, the greater presence of noise and other data problems within the 1 km data set, and a proportional increase of mixed classes within the 1 km training pixels.
8. Evaluation

A global validation data set against which to measure the accuracy of this land cover product does not exist, although an effort is underway to develop such a database at 1 km (Belward 1996). Many researchers have stressed the need for statistically rigorous validation efforts for maps being used for scientific investigations and policy decisions (Stehman and Czaplewski 1998). However, the validation of this 1 km data product is beyond available resources at this time. In order to evaluate this map, a few comparisons were made with other existing regional data sets which employed high-resolution data sources. Although these cannot be used as validation data, they do help characterize the map by yielding a measure of concurrency between products that were derived entirely independently from one another.

8.1. Environmental Protection Agency Region 3 characterization

The United States Environmental Protection Agency (EPA) has begun an effort to classify the ten Federal Standard Regions as part of an effort called the Multi-Resolution Land Characteristics Consortium National Land Cover Data Base (MRLC) (Vogelman et al. 1998). To date only region three, consisting of Delaware, the District of Columbia, Maryland, Pennsylvania, Virginia and West Virginia, has been finished. It is a 30 m map product derived from Landsat Thematic Mapper (TM) data and characterizes some classes in common with the UMd product. By reprojecting and resampling the classes to their proportional representation at 1 km, an evaluation of the agreement between the two can be undertaken.

Much of this area is taken up by the Appalachian chain of mountains, which are largely forested, while the valley bottoms and the eastern coastal plain consist of a mosaic of remnant forests and agriculture. By taking the most frequently occurring land cover at the 30 m spatial resolution within geo-registered 1 km squares, a map of the EPA product aggregated to 1 km was produced. Both land cover products, as aggregated to the UMd scheme, are shown in figures 11(a) and 11(b). This study area is dominated by broadleaf deciduous forest, and many valleys present at fine resolutions are not captured at 1 km resolution. The overall agreement per pixel is 65%. When examining homogeneous, or core area, 1 km pixels which consist of greater than 90% one cover type within the high-resolution map layer, the agreement increases to 81%. The corresponding forest/non-forest numbers are 83% and 92%. In all comparisons, forest includes all five forest types and non-forest includes the remaining eight classes.

The per cent agreement numbers should not be confused with accuracies, but are reported only in order to aid in the visual interpretation of the graphics and to reflect a measure of general thematic agreement. Although viewing core areas can yield overly optimistic results (Hammond and Verbyla 1996), it is worth doing here for a number of reasons. First, some mosaicked areas for the MRLC data do not have a single cover which is dominant at 1 km. For EPA Region 3, these areas are most likely to be represented in the 1 km map as a partial tree cover class such as woodlands or wooded grassland. These intermediate classes were not included in the MRLC classification scheme and a straightforward comparison cannot be made. Also, the use of AVHRR data reduces the heterogeneity present in mosaicked areas while this spatial complexity remains even in the resampled MRLC data. Evaluating core areas yields a measure of thematic agreement while minimizing problems associated with this inherent incompatibility.

Grasslands in the MRLC map are confused with croplands and wooded grassland
in the UMd map. Pastures occurring within intensive agricultural areas were not trained on for the 1 km data set and this could lead to errors of omission for the grassland class. Areal comparisons of all regional data sets with the UMd map are shown in figure 12.
8.2. European Coordination of Information on the Environment data

Digital maps for much of western Europe are now available in the European Coordination of Information on the Environment (CORINE) data set produced by the European Topic Center on Land Cover (CEC 1993). Germany was used for comparison as a large country for which the classes aggregated reasonably well to the UMd scheme. A graphic comparison of the UMd and CORINE products for Germany can be seen in figures 11(c) and 11(d). The agreement between the maps at 1 km, using a resampling of dominant cover type for the CORINE data into the 1 km grid, is 65% for all pixels and 83% for those 1 km grid cells with greater than 90% of one land cover type, showing that the core areas for the respective cover types have good agreement. Forest/non-forest comparisons agree at a level of 81% for the entire country and 92% when viewing only the 90%-pure CORINE pixels. The agreement of needleleaf forests increases from 56% to 80% going from all pixels to just the 90%-pure ones, while there is great confusion between mixed and broadleaf forests. Grasslands are poorly depicted in the north, revealing once again the limited ability to depict pasture within areas of intensive cropping. However, one-quarter of the CORINE grassland is labelled wooded grassland, mostly within areas of forest/grasslands mosaics in the south and west of the country.

8.3. NASA Landsat Humid Tropical Deforestation Pathfinder Project data at the University of Maryland

The UMd Pathfinder (Townshend et al. 1995) data sets for Colombia, Peru, Bolivia and the Democratic Republic of the Congo were examined in order to test the success of the UMd product in mapping tropical forest boundaries. The NASA Landsat Humid Tropical Deforestation Project depicts only a forest and non-forest layer, where forest represents humid tropical closed canopy forest and all other land covers are grouped as non-forest. The results for the South American data are shown in figure 13(a). The UMd product misses small clearings within the forest boundary and some montane forests, possibly due to the presence of clouds. One area of
interest is south-east Bolivia where extensive tropical deciduous woodlands and forests are found. The UMd product has significant areas of this forest to the south trending into the Gran Chaco, where the Pathfinder product shows none as these woodlands/forest are no longer a humid type formation. The agreements of 89.0%, 91.8% and 82.0% for Colombia, Peru and Bolivia respectively show the general success of the UMd product for this area in classifying tropical forest.

The Landsat Pathfinder data for the Democratic Republic of the Congo (Townshend et al. 1995) was examined in a similar fashion. Figure 13(b) shows this comparison. The overall agreement is 85.7%. The largest source of error is found within the forest boundary and in the eastern highlands. As stated earlier, depicting
areas with persistence haze was found to be problematic. Training data from Equatorial Guinea and the Republic of the Congo created splits associated with cloud cover. However, these splits created problems in other areas of the globe, where cloudy signals were being mapped as broadleaf evergreen forest. As such, these splits were dropped from the trees, as the signals are not representative of a characteristic land cover. The result is a spotty depiction of forest cover within the central forest with problems increasing towards the Atlantic coast and Gabon. Within the context of central Africa, these cloudy splits could be used for delineating likely forest in the central basin, along the Atlantic coast, and in the eastern highlands abutting the western rift valley.

8.4. Evaluation summary

A number of conclusions can be drawn based on the comparisons made between the regional databases and the UMd product. The basic distinction between forest and non-forest shows good agreement with other sources, ranging from 81% to 92%. One area of possible improvement for the UMd map is the mapping of pastures within heavily agricultural areas. Future iterations of this product must include better training for this cover subtype. Atmospheric degradation of the remote sensing signal in central Africa is difficult to handle in the global context and suggests the possible value of fusing other data sources such as radar in these areas.

Landscape heterogeneity found in high-resolution data sets is reduced in the 1 km multi-temporal UMd product. Favouring the dominant classes when using coarser resolution data, especially the greener classes due to multi-temporal NDVI compositing, is in agreement with other findings (Moody and Woodcock 1994) and possible ways to reduce the loss of information for coarse-scale maps are needed (Moody 1998). Using high-resolution data as a surrogate for ground truth may be a cost-effective way to characterize errors present in coarse-scale maps (Kloditz et al. 1998).

9. Comparison with Food and Agriculture Organization forest statistics

United Nations Food and Agricultural Organization (FAO) country forest statistics were compared with the UMd country totals for the three levels of canopy closure mapped in the classification system. The FAO statistics are provided by individual countries and date from different years. An adjustment function was then used to estimate 1995 totals for all countries (FAO 1997). For developed countries, the FAO definition of forest describes areas with a minimum of 20% tree crown cover and for developing countries, a minimum of 10% tree crown cover. Although the definitions of forest for FAO and the UMd data differ slightly, some conclusions can be made by comparing the global 1 km UMd product to statistics generated by individual countries.

Figure 14 shows plots of FAO forest versus UMd aggregate classes of forest (> 60% canopy cover), forest plus woodland (> 40%), and forest plus woodland plus

Figure 14. Plot of FAO total forest cover country statistics versus predicted woody cover of UMd product in millions of square kilometres. (a) UMd forest (> 60% tree canopy cover) versus FAO forest. (b) UMd forest plus woodland (> 40% tree canopy cover) versus FAO forest. (c) UMd forest plus woodland plus wooded grassland (> 10% tree canopy cover) versus FAO forest. (d) best agreeing of the three UMd canopy closure figures versus FAO forest.
Global and regional land cover characterization
wooded grassland (> 10%). Overall the UMd aggregate class which best agrees with the FAO numbers globally is forest plus woodland, or areas with 40% and greater canopy closure. The UMd global totals for this aggregate are 11% higher than the FAO numbers, compared with being 80% higher when using all three woody classes combined and 37% lower using forest classed pixels alone.

However, it is possible that different definitions of forest are being applied in different countries. Others examining the FAO data set have found discrepancies due to different forest definitions, data sources and data processing (Mayaux et al. 1998). For example, many regional subsets of countries such as those of the humid tropical Pan-Amazon, including Venezuela, Columbia, Peru and Bolivia, have FAO numbers which best agree with the forest numbers from the UMd map. On the other hand, many semi-arid nations of central Asia best agree with the aggregate of all woody classes, though their areas are too small to show in figure 14. By taking the one UMd class aggregate which is closest to the FAO number (figure 14(d)), the overall agreement between FAO global forest estimates and UMd estimates is reduced to a 7.0% lower figure for the UMd map compared to the FAO.

The continent with the highest disagreement between the two sources is Africa. By breaking the African countries into regional groupings, some disagreements can be reconciled when viewing the different canopy threshold aggregates. For a subset of humid tropical countries including Liberia, Sierra Leone, Cameroon, Equatorial Guinea, the Republic of the Congo, the Democratic Republic of the Congo and Rwanda, the total disagreement, when comparing FAO forest (> 10% crown cover) to UMd forest (> 60% canopy cover), is less than 1%: 1 562 270 km$^2$ to 1 577 816 km$^2$. For these countries the FAO statistics do not appear to reflect the FAO definition of forest for developing nations, but are closer to measuring the UMd-defined forest subset of that definition. The disagreement jumps to 48% and then to 93% overestimations of forest cover when adding woodland and then wooded grassland totals to the UMd forest numbers. In general, it appears that countries with significant proportions of tall, dense forest are more likely to map this type of formation as the forested land, and not other less dense formations which still meet the FAO definition of 10% tree crown cover. On the other hand, countries such as Namibia, Botswana, Senegal, Mali and Niger appear to include sparser stands of trees such as that from the UMd wooded grassland class. For these countries, forest plus woodland UMd totals underestimate the FAO total by 97%, but by adding wooded grassland, the total disagreement is reduced to a 24% overestimation. This variable standard of woodiness can be discerned through the use of the global UMd land cover characterization.

The generation of internally consistent global classifications using remotely sensed data could be of help to users of data sets such as the one generated by FAO. The ground-based forest statistics compiled by the FAO represent a combination of various sources which could be brought into greater harmony through the use of remote sensing. For modellers and researchers interested in the human impacts of global change, an internally consistent approach to mapping land cover should prove valuable in standardizing the depiction of natural resources across regions, continents and the globe.

10. Conclusions

Performing global classifications of remotely sensed data provides for an internally consistent product which allows for the comparison of land cover between
regions and continents. In this study, a 1 km global land cover classification conforming largely to the IGBP class definitions has been made. A set of classification trees were created to map land use cover using AVHRR 1 km data from 1992–1993. The minimum annual red reflectance metric proved very useful in delineating woody areas, while peak annual greenness was useful in describing leaf type. Temperature metrics were also used in discriminating tropical woodlands from forest, drought deciduous broadleaf forest from evergreen broadleaf forest, and in stratifying the tropics from temperate and boreal zones. Temperature metrics also helped in separating shrublands from grasses and agriculture. Near-infrared metrics were found to be helpful in separating crops from grass and shrub covers, and tropical inundated grasslands from woodlands. Many of the splits lend themselves to ready biophysical interpretations and point to the possibility of using the same tree for separate years in order to test the method’s repeatability and for eventual use in detecting land cover change.

The classification trees also revealed the relatively few steps it takes to characterize most of the globe. However, many of the trees featured subtrees of considerable complexity, possibly related to the quality of the data. Future efforts using sensors such as MODIS will reveal the possibility of creating decision trees where a handful of splits should successfully describe the entire globe. For the data used here, this was not possible.

All compositing methods available should be assessed for their utility in mapping land cover at coarse resolutions. NDVI compositing has been shown to be biased towards high view zenith angles in the forward scatter direction, preferentially binning on BRDF affected pixels. This has the effect of introducing geometrically distorted pixels, as well as making the derivation of true at-nadir reflectances difficult. The geometric distortion of pixels due to compositing methods was evident in the comparisons with high-resolution derived map products. Preliminary examinations of the UMD map to other AVHRR-derived maps using single-date imagery (Zhu and Evans 1991, Mayaux et al. 1997) also show an increased blurring of the landscape due to the multi-temporal signals and maximum NDVI compositing. These characteristics of multi-temporal compositing imply that the map would more appropriately be made at a resolution greater than 1 km, as the footprints of many pixels are actually considerably larger than the at-nadir size of 1.1 km. Also, red and especially near-infrared values from the 1 km data set are suspected of being greatly affected by BRDF effects and this limits the utility of these bands for certain areas and land covers. The resulting noisy time series for Channels 1 and 2 complicate the use of these bands in classification. More robust approaches relying on one or more different compositing criteria should be used (Cihlar et al. 1994, Lambin and Ehrlich 1996), along with corrective procedures (Cihlar et al. 1997), in order to avoid the typical problems associated with maximum NDVI compositing. Additionally, adding corrections for water vapour and aerosols will help create less noisy time series (Ouaidrari et al. 1997) and should allow for the creation of simpler classification trees.

Acknowledgments

Funding from the National Aeronautics and Space Administration supported this research under Grants NAG56004, NAS596060 and NAG56364.

We are grateful to the providers of regional data sets: the University of Maryland NASA Pathfinder Humid Tropical Deforestation Project, the Environmental Protection Agency, Robert Enesund of the European Topic Center for Land Cover
and Zhu Zhiliang of the EROS Data Center. We also would like to thank Jeff Eidenshink of the EROS Data Center for assistance with the AVHRR global data set. We also thank the following for their comments on earlier versions of this paper: Otto Bauer, Jesus Duran, Scott Goetz, Satya Kalluri, Nadine Laporte-Goetz, Hassan Ouaidrari, Stephen Prince, David Roy, Marc Steininger and Compton Tucker. However, responsibility of this product is solely that of the authors. We also extend our thanks to the three anonymous reviewers.

References


