Available online at www.sciencedirect.com
ScienceDirect
Remote Sensing of Environment 106 (2007) 1–16
Remote Sensing of Environment
www.elsevier.com/locate/rse

Spatial, spectral and temporal patterns of tropical forest cover change as observed with multiple scales of optical satellite data

Daniel J. Hayes [a,*], Warren B. Cohen [b]

[a] Department of Forest Science, 321 Richardson Hall, Oregon State University, Corvallis, OR 97331, USA
[b] Forestry Sciences Laboratory, Pacific Northwest Research Station, USDA Forest Service, 3200 SW Jefferson Way, Corvallis, OR 97331, USA

Received 1 June 2006; received in revised form 13 July 2006; accepted 15 July 2006

Abstract

This article describes the development of a methodology for scaling observations of changes in tropical forest cover to large areas at high temporal frequency from coarse resolution satellite imagery. The approach for estimating proportional forest cover change as a continuous variable is based on a regression model that relates multispectral, multitemporal MODIS data, transformed to optimize the spectral detection of vegetation changes, to reference change data sets derived from a Landsat data record for a study site in Central America. A number of issues involved in model development are addressed here by exploring the spatial, spectral and temporal patterns of forest cover change as manifested in a time-series of multi-scale satellite imagery.

The analyses highlighted the distinct spectral change patterns from year-to-year in response to the possible land cover trajectories of forest clearing, regeneration and changes in climatic and land cover conditions. Spectral response in the MODIS Calibrated Radiances Swath data set followed more closely with the expected patterns of forest cover change than did the spectral response in the Gridded Surface Reflectance product. With forest cover change patterns relatively invariant to the spatial grain size of the analysis, the model results indicate that the best spectral metrics for detecting tropical forest clearing and regeneration are those that incorporate shortwave infrared information from the MODIS calibrated radiances data set at 500-m resolution, with errors ranging from 7.4 to 10.9% across the time periods of analysis.

Keywords: Satellite image change detection; Tropical land cover and land use change; MODIS

1. Introduction

Forest cover conditions are in constant flux from changes that may be subtle or abrupt, and can result from natural and/or anthropogenic forces. The capacity to monitor these on-going changes is critical to efforts toward sustainable resource management, habitat conservation, and balancing the global carbon budget and understanding its role in climate change. These changes will have characteristic spatial, spectral and temporal patterns that can be observed in multiple acquisitions of satellite data over time. As defined by Coppin et al. (2004), this variability can result from a number of sources, including the conversion from one land use type to another (e.g. forest to pasture) or the more subtle modification of land cover over time (e.g. the accumulation of biomass in growing forests). Furthermore, year-to-year changes in ecosystem condition due to climatic variation, for example, will result in significant spectral variability observed even in unaltered land cover types. Understanding the link between the conversion, modification and condition of land cover with the interannual spectral variability in a multitemporal satellite data set is required to: 1) separate the changes of interest from spurious change and unchanged areas and, 2) identify the type, extent and intensity of change.

1.1. Remote sensing observations

For over 30 years, scientists and practitioners have been using Landsat data to monitor vegetation and land cover at landscape and regional scales. The Landsat data set has notable advantages that account for it being so widely-used in ecological applications, including a long-running historical time-series, a spatial resolution appropriate to regional land cover and land use change (LCLUC) investigations, and a spectral coverage appropriate to studies of vegetation properties (Cohen & Goward, 2004). However, the low

* Corresponding author. Tel.: +1 541 758 7758; fax: +1 541 737 1393.
E-mail address: Daniel.Hayes@Oregonstate.edu (D.J. Hayes).

0034-4257/$ - see front matter
doi:10.1016/j.rse.2006.07.002

repeat frequency (with recurrent cloud coverage) and small area coverage per scene preclude its use for large-area monitoring in the tropics. The current sensor malfunctions with the Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and the unavailability of Landsat 5 Thematic Mapper (TM) data for some regions, lends further credence to considering other sources of data for monitoring LCLUC. With its improved radiometric properties, moderate spatial and spectral resolutions and twice-daily coverage, the freely-available suite of products generated by NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) hold the promise of becoming the most important data source for land cover characterization and operational monitoring at regional and global scales (Townshend & Justice, 2002).

1.2. Satellite image change detection

Relating the spectral changes among multitemporal satellite data sets to changes in surface features and biophysical variables has long been an important application of remote sensing technology (Coppin et al., 2004; Singh, 1989). The most common approaches used in change detection have included the "post-classification" comparison of two dates of land cover classifications and the simple two-date, univariate differencing of single bands or band ratios. Multivariate techniques that incorporate more spectral information in the change detection algorithm often produce improved results. Examples include clustering of multitemporal composite images (e.g. Hayes & Sader, 2001) and the use of multivariate transformations such as the tasseled cap (e.g. Healey et al., 2005), principal components analysis (e.g. Fung & LeDrew, 1987; Muchoney & Haack, 1994), and the less-used canonical correlation analysis (e.g. Neilsen et al., 1998). One commonality amongst most of these approaches is the need to apply some often arbitrary or subjective thresholding or classification decision-rule that separates the change areas of interest from spurious or no change areas in a given study area (Fung & LeDrew, 1988).

Traditional classification techniques have typically relied on the mapping of land cover and change by associating each pixel in an image with a single, discrete category, under the assumption that each pixel represents a homogenous area on the ground. However, these techniques do not fully exploit the information about the variability of the features of interest inherent in the spectral signal of remotely sensed data. A variety of approaches have been used to estimate the "sub-pixel" composition of land cover and biophysical variables, including regression techniques (Cohen et al., 2003; Iverson et al., 1989; Zhu & Evans, 1992), fuzzy class membership functions (Foody & Cox, 1994), linear mixture modeling (Adams et al., 1995; DeFries et al., 2000), and artificial neural networks (Foody et al., 1997). These techniques become particularly important when the phenomena of interest operate at scales below the spatial resolution of the satellite data set.

The majority of these investigations have focused on the utility of such techniques to detect gradients in and/or decompose the spectral signal of a pixel for mapping land cover at one point in time. Because of the advantages involved with mapping sub-pixel cover, Hansen et al. (2002) speculate that the MODIS-based Vegetation Continuous Fields products can be used to measure changes in proportional forest cover over time, yet few studies have applied sub-pixel classifications in a change detection context. Prototype studies of change algorithms using 250 m MODIS radiance data suggest satisfactory results for extreme events such as large-scale flooding and burning, yet the detection of deforestation was less reliable where forest clearing is patchy (Morton et al., 2005; Zhan et al., 2002). Furthermore, it has been well established that estimating change from two separate dates of classifications (post-classification change detection) suffers from the propagation of often significant errors inherent in the classification at each date (Jensen, 1996). Direct spectral modeling of the sub-pixel proportion of change in a multitemporal, moderate resolution data set therefore offers the opportunity to take advantage of the improved information on pixel heterogeneity while reducing the potential for post-classification error propagation.

1.3. Objectives

The fundamental goal of this study is to develop a methodology for scaling observations of changes in tropical forest cover to large areas at high temporal frequency. Despite its advantages, the primary challenge of using MODIS data is to extract useful information on land cover change when the processes of interest operate at a scale below the spatial resolution of the sensor. Patches of clearing and regeneration in slash-and-burn agriculture systems, for example, often average 2 or 3 ha in size (Hayes et al., 2002; Sader, 1995), whereas a MODIS pixel, at minimum, covers more than 6 ha. The approach for estimating proportional change as a continuous variable is based on a regression model that relates multispectral, multitemporal MODIS data, transformed to optimize the spectral detection of vegetation changes, to reference change data sets derived from the Landsat analysis.

In general, there are a number of issues involved in building a forest cover change model, which are addressed here by exploring the spatial, spectral and temporal patterns of land cover change in a tropical forest study area, and examining how these patterns are manifested in the multitemporal data sets. An advantage exploited in this study is the direct comparison of accurate and detailed land cover change information derived from Landsat data with a MODIS data set, for a time-period when concurrent data from the two sensors exists for the study area of interest. Specifically, these analyses are used to address three main objectives of the modeling effort:

1) Explore the spatial patterns of both the land cover change and spectral data sets to examine the effect of grain size on spatial pattern and to compare the image quality of the various spectral data used in this study;
2) Investigate the potential spectral trajectories of major land cover types as they change over time in order to develop an understanding of the expected relationship of forest cover change to year-to-year changes in spectral response; and
3) Compare the results of using different spectral indices from different data sets at varying spatial resolution for predicting proportional change in forest cover over time.

2. Methods

2.1. Study area and data

Meeting the research objectives of this study requires detailed, landscape-scale land cover maps and analyses of LCLUC. These data have been derived from high spatial and temporal resolution databases developed for a study area located in northern Guatemala, corresponding to Landsat Worldwide Reference System Path 20/Row 48 (Fig. 1). This study area covers much of the Maya Biosphere Reserve, a complex of delineated management units including five national parks, four biological reserves, a multiple use zone, and various community-level forest concessions. The Reserve is characterized by mainly lowland forest and expansive freshwater wetlands, part of the largest contiguous tropical moist forest remaining in Central America (Nations et al., 1998). However, since the mid-1980s, the land within this study area has been subjected to widespread LCLUC as a result of human migration and expansion of the agricultural frontier (Sader et al., 1997). The resulting land cover pattern consists of a mosaic of mature forest, pasture and croplands, and a range of ages of secondary forests (Hayes et al., 2002).

The analyses described in this study (Fig. 2) span a three-year time period (2000 to 2003) in which MODIS data could be acquired concurrently with Landsat 5 and pre-scan line corrector (SLC)-off Landsat 7 data. Landsat 7 SLC-off data was acquired in 2004 and 2005 to extend the analysis of the spectral trajectories of known land cover and land cover change areas. Among the myriad data and products available from the MODIS sensor, of primary interest to this study were the Calibrated Radiances (MOD02) and Surface Reflectance (MOD09) products. To minimize external effects from the timing of land cover changes, as well as from atmospheric and seasonal differences, single-date MODIS data were acquired on the same day as, or as close as possible to, the annual Landsat acquisitions (Table 1). Ancillary spatial data, previously developed land cover and change maps, and field visits were also used to guide the target and training site selection, classifications and other analyses described in this paper.

There are numerous other products available in the MODIS data stream that have potential relevance for mapping forest cover change, including the Enhanced Vegetation Index (EVI) and Bidirectional Reflectance Distribution Function (BRDF) products. However, evaluations of these data are not included in this exploratory analysis of forest cover change patterns for various

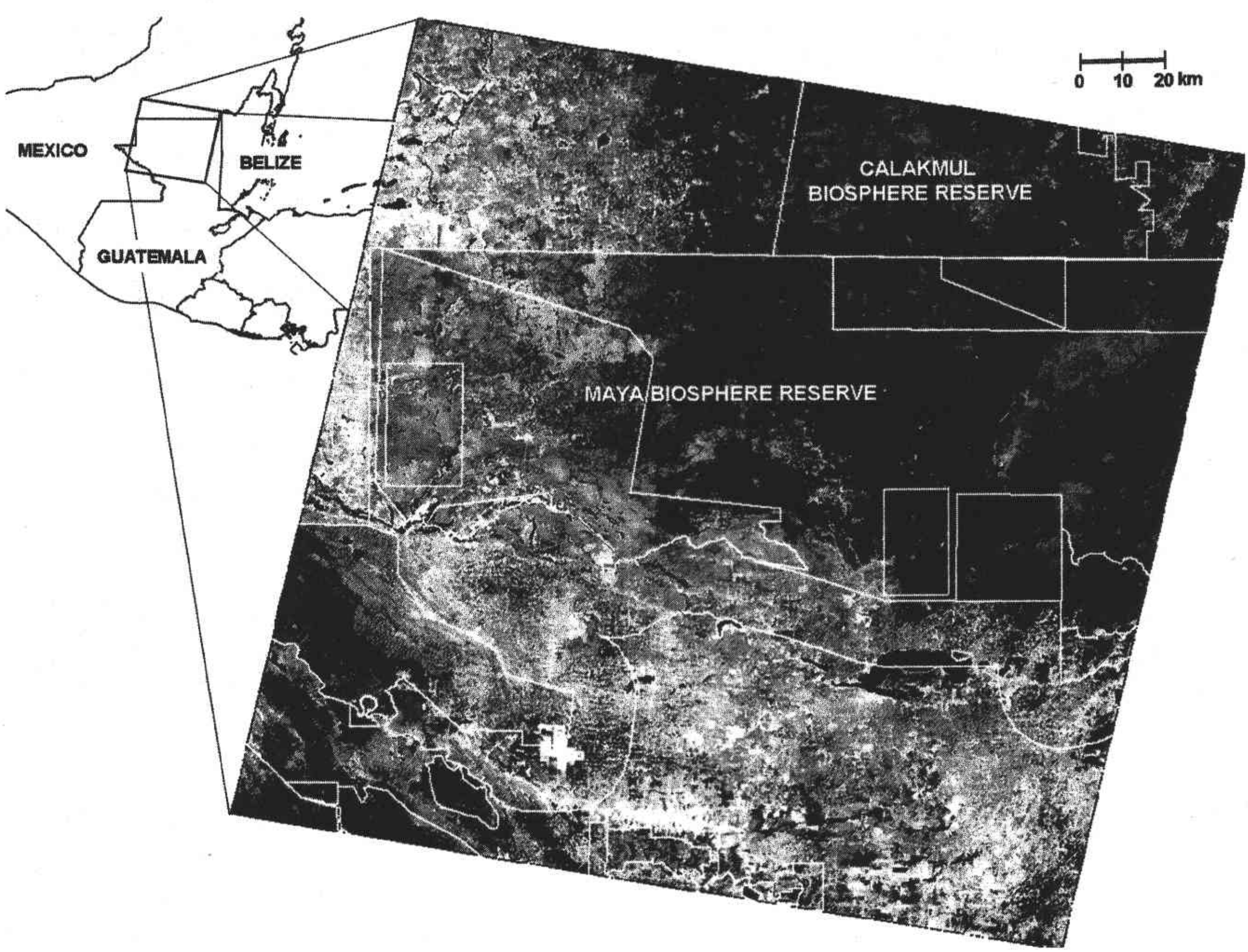

Fig. 1. Location of the study area (Landsat WRS Path 20/Row 48, 2000 ETM+ band 5 shown) in relation to the protected areas of the southern Yucatan Peninsula of Guatemala and Mexico.

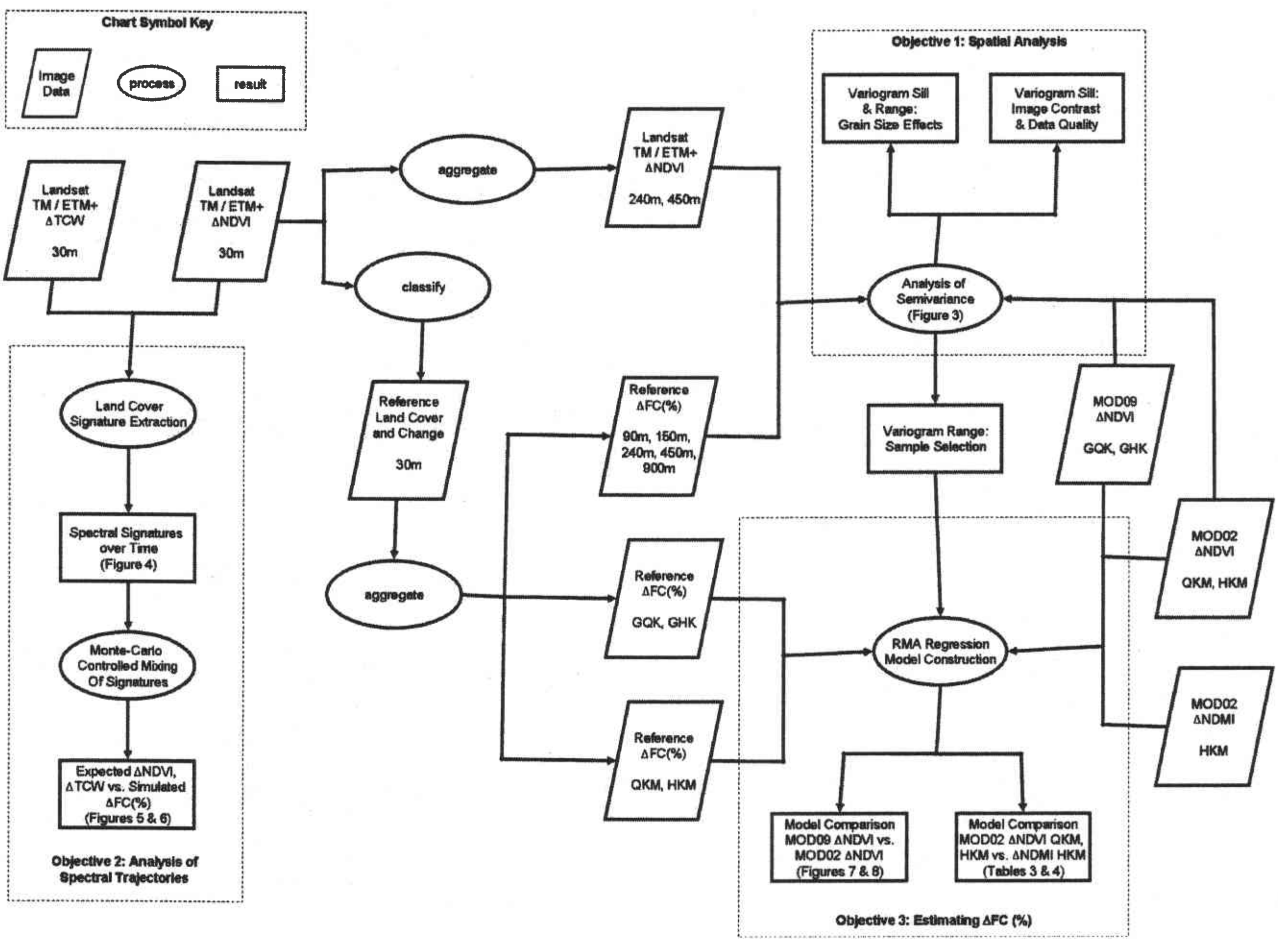

Fig. 2. Flow-chart illustrating the relationship of the various data sets, processes, results and interpretations to the objectives outlined by this study.

Table 1
Acquisition dates for the images used in this study

	Year					
Sensor	2000	2001	2002	2003	2004	2005
MODIS	27 Mar 2000	30 Mar 2001	12 Mar 2002	07 Mar 2003		
Landsat						
TM		22 Mar 2001				
ETM+	27 Mar 2000	-	17 Mar 2002	07 May 2003		
ETM+ (SLC off)					18 Jan 2004	11 May 2005

reasons. While the EVI has shown improved ability to distinguish vegetation parameters in high biomass areas over the traditional NDVI (Huete et al., 2002), the MODIS vegetation indices are based on MOD09 spectral data and is a product of a 16-day compositing algorithm. This study's objective of an evaluation of the spatial, spectral and temporal patterns of forest cover change while controlling for external effects led to the analysis of the original, single-date MODIS radiance and reflectance bands. The EVI was not included so as to reduce the complexity and redundancy (i.e. with the MOD09 data) in the analysis and to avoid any errors in the timing of land cover change that would be lost in the 16-day composites. Likewise, other higher-level products such as the BRDF are also based on MOD09 and multi-date composited, as well as being collected at spatial resolution (1-km) more coarse than appropriate for this study.

2.2. Satellite data pre-processing

The Landsat 7 ETM+ scene acquired on 27 March 2000 was used as the "base-line" for the geometric correction and radiometric normalization of the entire Landsat data set. The reference scene was previously geo-rectified to ground control points and resampled with a 30-m pixel resolution in the Sinusoidal Projection (WGS 84) to match the native grid of the MODIS data and products. The other five Landsat scenes were geo-rectified to the reference via the automated selection of image tie-points, as described in Kennedy and Cohen (2003), and first-order polynomial resampling. Overall root-mean-square error was less than half a pixel (<15 m) for each resampled image. Digital numbers were converted to reflectance values in the reference image using the COST model (Chavez, 1996). The subsequent Landsat scenes were radiometrically normalized, band-by-band, to the reference scene using linear regression techniques (Hall et al., 1991). The algorithm was based on the selection of end-point targets from the wet and dry non-vegetated extremes in each band at each date, as described for this study site in Hayes and Sader (2001).

Collection 4 Terra MODIS data and products were used in these analyses, which are distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at the U.S. Geological Survey (USGS) Center for Earth Resources Observation and Science (EROS) http://LPDAAC.usgs.gov. For MOD02 and MOD09 data, red (R) and near-infrared (NIR) spectral information (bands 1 and 2, respectively) are available at quarter-kilometer resolution. The half-kilometer data contains a total of 7 layers, including blue, green, and three shortwave infrared (SWIR) channels (bands 3–7), in addition to resampled bands 1 and 2. The MOD09 product has been converted to surface reflectance values according to the algorithm described by Vermote et al. (1997) and is available in gridded sinusoidal projection format. The MOD02 data consists of at-satellite calibrated radiance values, prior to any atmospheric correction algorithm. Another important difference between the two data sets is that the MOD02 data are available as the original swath observations, prior to being resampled to the pre-defined storage bins used in the higher-order gridded sinusoidal products, including MOD09. Tools to reproject and reformat MOD09 and MOD02 data are available from the USGS (http://edcdaac.usgs.gov/landdaac/tools/modis/) and were used here. The MODIS Reprojection Tool for gridded products (MRT) provided MOD09 output in the sinusoidal grid with cell sizes of 231.7 m and 463.3 m for the quarter-kilometer (MOD09GQK) and half-kilometer (MOD09GHK) data sets, respectively. The MODIS Reprojection Tool for swath data (MRT Swath) required latitude and longitude data for geolocation of MOD02, which is available as part of the MODIS data stream (MOD03) at 1-km resolution. MOD02 data was output from the MRT Swath in the sinusoidal projection to maintain consistency with the MOD09 data but retained the native cell resolutions of 250 m and 500 m for the quarter-kilometer (MOD02QKM) and half-kilometer (MOD02HKM) data sets, respectively.

2.3. Reference forest cover change classification

Researchers have monitored LCLUC in the Maya Biosphere Reserve using Landsat data from the past three decades (Hayes et al., 2002; Sader et al., 2001). For this study, the mapping of forest clearing and regeneration was extended to include annual assessments of LCLUC from 2000 to 2003 using Landsat TM and ETM+ acquisitions. After removing clouds from the imagery (by a combination of multivariate clustering and hand-editing), forest clearing, burning and regeneration were classified for each time period (Table 2) by visual interpretation and clustering of multi-date normalized difference vegetation index (NDVI) composite imagery, as described by Hayes and Sader (2001). The authors report an overall accuracy of 85% (Kappa = 0.83) for this method, using the visual interpretation technique for accuracy assessment of change classifications reported by Cohen et al. (1998). In addition to forest clearing for agricultural and pasture establishment, the study area experienced extensive loss of forest cover during this time period due to wildfire (CONAP, 2005).

Table 2
Area summaries (hectares) for the reference change classifications by time period

Time period	Category			
	Clouds	Clearing	Regeneration	Burn
2000 to 2001	221,708	33,711	19,848	0
2001 to 2002	515,014	18,667	11,136	0
2002 to 2003	1,599,166	21,303	8765	174,196
2000 to 2003	21,742	80,007	83,656	208,231

The Landsat-derived reference LCLUC was aggregated to a continuous variable representing the net proportional change in forest cover (ΔFC) to match the grid resolutions of the MODIS data and products. In this study, the definition of forest cover includes both mature forest cover as well as young forests ($\geq$1 year-old) regenerating on abandoned agricultural or fallow sites. To develop the model, a grid matching the location and size of the MODIS pixels was overlaid on the reference land cover change data and a count of the relative number of cleared, regrowth, and no change Landsat pixels were used to calculate the reference $\Delta FC_{i,t}$, which has a possible range of -100% (completely cleared) to $+100\%$ (completely regrown) for each grid cell (i) at each time period (t). Reference ΔFC data sets were calculated according to the grid cells matching the spectral data set being investigated so that separate ΔFC data were created using the MOD02 QKM and HKM cells as well as the MOD09 GQK and GHK grids.

2.4. Spatial analysis

The spatial properties of the observed land cover change and the spectral data were explored (objective 1) through an analysis of semivariance, a fundamental measurement of geostatistics that has been used in remote sensing to study the spatial structure of both ground-level and satellite data sets (see Curran & Atkinson, 1998). The semivariance (γ) of a spatial variable is calculated as half the average squared difference in the value of the variable between sample points. The spatial variation in a variable can be represented, then, by constructing a semivariogram ($\gamma(h)$), which is a plot of γ as a function of distance or lag, h. In remote sensing studies, semivariograms have been used to guide spatial sample selection (Atkinson, 1991), to indicate the amount of spatial heterogeneity in image data (Cohen et al., 1990), for modeling biophysical properties (Berterretche et al., 2005) and to provide an assessment of image quality (Tian et al., 2002).

Several key features of semivariograms allow interpretation of the spatial dependence and variability in remotely-sensed and ground data (Curran, 1988). In a nonrandom spatial data set, γ increases from an initial value at lag zero (the nugget) to an asymptotic value (the sill), after which γ remains relatively constant at larger h. The lag distance whereupon γ reaches the asymptote is referred to as the range, which defines the distance over which the variable is spatially autocorrelated. The sill provides an assessment of the spatial heterogeneity, or contrast, in the spectral data and, when compared to the semivariance in reference data, is an indicator of the satellite data quality (Tan et al., submitted for publication).

Semivariograms were constructed for both the reference change in forest cover (ΔFC) and the change in the normalized difference vegetation index (ΔNDVI) between two dates and at multiple grain sizes. Both reference ΔFC and ΔNDVI were summarized for the three-year time period between 2000 and 2003. The semivariogram analysis was conducted on a subset area in the center of the Landsat scene, covering approximately 75×75 km and representing a mixture of undisturbed mature forest interspersed with shifting cultivation agriculture concentrated in a few small communities within the Reserve's multiple use zone (Hayes et al., 2002). Omni-directional semivariance of the ΔFC and ΔNDVI subsets were calculated using the "kriging" routine in ArcGIS (ESRI, 2005) and plotted against lag in the variograms. The semivariograms of the reference ΔFC were used to interpret the distance beyond which there was no longer spatial correlation in the data, and this range was used as the minimum distance allowed between sample points selected for building the regression modeling of ΔFC, as explained below. The semivariance at the sill of the variogram was compared among ΔNDVI derived from the MOD02 and MOD09 at both quarter-kilometer and half-kilometer resolution, along with that of the Landsat data, degraded in pixel size from 30 m to 240 m (8×8 pixels) and 450 m (15×15) resolution.

2.5. Analysis of temporal spectral trajectories

In this study, the spectral variability associated with vegetation and land cover conversion, modification and condition was investigated (objective 2) through an analysis of 6 dates (2000 to 2005) of Landsat data (Table 1) for "training" areas of known land cover type. The set of 70 training areas represented 10 sites each of the following types: mature, high-stature forest; low-stature, seasonally inundated forest; savannas; wetlands; agricultural plots; pastures; and secondary forest regenerating on abandoned or fallow agricultural or pasture sites. Each area selected comprised at minimum a 10- by 10-pixel (9 ha) area and represented a single land cover category, unchanged in type over the 6 dates and known for 2005 from field visits and/or visual interpretation of Landsat color composite imagery (see Hayes & Sader, 2001). Care was taken to ensure that training sites did not include any data gaps in the 2004 and 2005 Landsat 7 SLC-off data.

The temporal spectral trajectories of each land cover type were investigated by examining their inter-annual response using two vegetation indices: the NDVI and the tasseled cap wetness (TCW) transformation. The NDVI, a widely-used spectral vegetation index that has been correlated to crown closure and other vegetation parameters (Tucker, 1979), measures the contrast between surface reflectance in the R and NIR wavelengths. NDVI was calculated for each date of Landsat normalized reflectance bands using the following formula:

$$\mathrm{NDVI} = \frac{(\mathrm{Band4}-\mathrm{Band3})}{(\mathrm{Band4}+\mathrm{Band3})} \tag{1}$$

The TCW is a transformation that exploits a more full spectral domain and emphasizes contrasts in the SWIR against the NIR and visible wavelengths (Crist & Cicone, 1984). The TCW index has been shown to be sensitive to the moisture content of vegetation as well as various forest canopy structural characteristics (Cohen & Spies, 1992). It was used here to explore the spectral behavior of the land cover trajectories in wavelengths other than the R and NIR, particularly with respect to the SWIR. The atmospherically-corrected Landsat reflectance data was converted to TCW at each date using coefficients derived by Crist (1985) for ground-based spectral data, as discussed by Cohen et al. (2003). For each spectral measure, the mean and standard deviation of the spectral response were summarized for each land cover type at each date and these trajectories were plotted over time.

2.6. Estimating forest cover change

The fundamental goal of this study was to build a model that predicts the change in forest cover between two dates as a continuous variable from a multitemporal spectral data set (i.e. MODIS). Models were constructed using a multiple linear regression approach in which the independent variables (*X*s) were derived from two dates of spectral data and used to model ΔFC, the dependent variable (*Y*), for the same time period. As an illustration, a model that predicts net ΔFC between two dates as a percentage of a given pixel from a linear combination of the digital numbers (DN) of a multitemporal, multi-spectral Landsat or MODIS data set can be depicted as:

$$\Delta\mathrm{FC}[\%] = \beta_0 + \sum_{b=1,t=1}^{n,2} (\beta_{b,t}\mathrm{DN}_{b,t}) + \varepsilon \qquad (2)$$

with DNs summed over n number of bands (b) at times (t) 1 and 2 and $\beta_{b,t}$ coefficients relate $\mathrm{DN}_{b,t}$ to ΔFC for the time period, where β_0 is the intercept, and ε is the error. The relationship of ΔFC to spectral change was investigated in this study by comparing (objective 3) models developed from a "control" data set based on the Landsat-derived spectral signatures of the training sites, as described in Section 2.5 above, and sample sets of actual spectral data from MOD02 and MOD09 imagery.

2.6.1. Simulations with Landsat spectral signatures

To investigate the year-to-year spectral behavior of the land cover states and changes among them, the Landsat-derived signature statistics for the training sites (Section 2.5) were used in controlled simulations of spectral change for each time period. This investigation employed a Monte-Carlo type analysis in which the mean and standard deviation in spectral response (NDVI and TCW) were used to create an inverse cumulative probability distribution for each land cover type at each date, assuming a normal distribution in spectral response. In the simulations, one hundred random variates of spectral response were drawn from each of these distributions. Each random variate from a given date was paired with a variate of the same spectral index from another date to calculate a change in spectral response, i.e. ΔNDVI or ΔTCW. The frequency distribution of variates used to calculate the change in spectral response between two dates of the same land cover type gave an indication of the year-to-year spectral changes associated with variation in vegetation condition.

The expected change in spectral response that would result from a land cover conversion between any two dates was investigated by comparing random variates of different land cover types at each date. For example, each of the one-hundred random variates from simulations of NDVI response in mature forest for the year 2000 were paired with a variate derived from the distribution of pastures in the year 2003. The resulting frequency distribution of ΔNDVI, then, indicated the expected range and variability in spectral change resulting from forest clearing between 2000 and 2003. Similarly, NDVI of random variates generated from the distribution for agricultural plots in 2000 was subtracted from the NDVI of young (1 to 3 years old), regenerating forest in 2003 to assess the expected spectral change for vegetation regrowth. Finally, the full range of proportional ΔFC and its relationship with spectral change was investigated by a controlled mixing of the spectral responses from the Monte-Carlo simulations, for different fractions of land cover types at each date.

2.6.2. Multitemporal MODIS spectral data

The trade-off in spectral and spatial resolution with the use of MODIS data was explored in the analysis of the data's utility in detecting changes in forest cover. The higher resolution (quarter-kilometer) data contains only two bands of spectral information (R and NIR), while gaining greater spectral information (such as in the SWIR wavelengths) in the half-kilometer data requires a reduction in the spatial grain size of analysis. The effect of spatial grain size on the modeling of ΔFC was explored by comparing models based on ΔNDVI from the quarter- and half-kilometer MOD02 and MOD09 data sets. For MODIS data sets, the NDVI at each date was calculated as:

$$\mathrm{NDVI} = \frac{(\mathrm{Band2} - \mathrm{Band1})}{(\mathrm{Band2} + \mathrm{Band1})} \qquad (3)$$

Because the NDVI calculation requires only R and NIR bands, it offers a common index for comparison across the different data sets and spatial resolutions. The effect of using SWIR information on the prediction of ΔFC was investigated at half-kilometer resolution where these wavelengths are recorded by MODIS.

Just as the NDVI provides a measure of the contrast between reflectance in the NIR and R wavelengths, other band ratios that contrast the SWIR wavelengths with the NIR band have been used with Landsat data to distinguish forest structure and moisture content (Fiorella & Ripple, 1993). These "infrared indices (II)" can be calculated as simple band ratios or normalized differences:

$$\mathrm{II} = \frac{(\mathrm{NIR} - \mathrm{SWIR})}{(\mathrm{NIR} + \mathrm{SWIR})} \qquad (4)$$

where, in Landsat data, NIR reflectance is recorded in band 4 and SWIR reflectance can be represented with band 5 or 7. Hardisky et al. (1983) found the normalized difference in Landsat bands 4 and 5 to be more sensitive to changes in plant biomass and drought stress than the NDVI. Recent studies by Wilson and Sader (2002) and Jin and Sader (2005) suggest that the Landsat 4–5 normalized difference index improves the accuracy of forest harvest detection over the traditional NDVI. These authors have termed this index the "normalized difference moisture index (NDMI)" because of its sensitivity to canopy moisture content and its high correlation with the TCW transformation (Jin & Sader, 2005). The NDMI was used in this study for its ease in calculation with MODIS data while also addressing the objective of testing the effect of adding SWIR information in the ability to model forest cover change. NDMI was calculated for both the MOD02HKM and MOD09GHK data sets by contrasting NIR band 2 with SWIR band 6.

2.6.3. Regression model construction

As described above (Section 2.3), the reference classifications were summarized according to the grid cells matching the data set being investigated, allowing direct comparison of ΔFC and MODIS spectral change for quarter- and half-kilometer resolution MOD02 and MOD09 data. The relationship was investigated for each data set and each time period using a stratified random sample of spatially independent grid cells to construct the regression model shown in Eq. (2). To select the sample, the continuous reference variable was first categorized into incremental classes of 10% ΔFC across the entire range (−100 to +100%), with an added class of 0% change. Equal numbers of samples were taken from each class according to the criteria that samples: 1) were far enough separated in space to be considered independent, as per interpretation of the range of autocorrelation in the variograms of the reference data, 2) corresponded to "good" quality data as interpreted from the MODIS quality assurance files and/or visual screening for the presence of water, clouds, smoke or other unwanted defects, and 3) in the added class of 0% change, incorporated the range of both unchanged forest and non-forest mixtures. Thus, a set of sample cells for model building was created for each data set (MOD02 and MOD09) at each resolution (quarter- and half-kilometer) for each time period in the analysis (annually for 2000 to 2001, 2001 to 2002, and 2002 to 2003; and the three-year period from 2000 to 2003).

Each sample data set to be modeled consisted of the radiance or reflectance values in each band for each of the two dates of MODIS data, as well as continuous variables on initial percent forest and non-forest cover and the reference ΔFC for the corresponding data set, cell size and time period. Based on the Landsat signature simulation exercise, it was apparent that ΔFC could not be modeled on spectral changes as one population, i.e. there were different relationships for forest regeneration versus forest clearing. As such, each data set was segmented into two separate model sets, one where ΔFC>0% (the "forest regeneration" set) and the other where ΔFC<0% (the "forest clearing" set). Where ΔFC=0%, sample cells with an initial forest cover less than 50% were included in the regeneration set and the remainder (initial forest cover greater than or equal to 50%) were modeled in the clearing set. For the clearing and regeneration model sets, ΔFC was modeled separately on two-date changes in MODIS spectral bands, band ratios, and other multivariate transformations described above.

The variables were modeled using multiple linear regression and ordinary least squares (OLS) for finding the coefficients of the best-fit line relating ΔFC to two-dates changes in the spectral variables. Cohen et al. (2003) demonstrated a regression approach based on the Reduced Major Axis (RMA) transformation that improves on OLS for the prediction of continuous variables with remotely sensed spectral data. In this study, all model coefficients and prediction results are given after an RMA transformation of the OLS predictions, as described in Cohen et al. (2003).

2.6.4. Regression model evaluation

The models developed were evaluated and compared based on several statistics describing the relationship between, and the predictive power of, the MODIS spectral data on ΔFC for each time period and each cell size, and for the clearing and regeneration model sets for both the spatially independent sample cells. The slope (β_1) and intercept (β_0) of the best-fit RMA line relating ΔFC to MODIS spectral variables were calculated separately for the clearing and regeneration model sets, along with their respective prediction vs. observed correlations (R). The "starting point" represents the point in spectral change space where each modeled relationship (clearing and regeneration) crosses the ΔFC=0% axis.

The two model sets were combined into one observed vs. predicted model over the complete range of ΔFC (−100% to +100%) for the pooled set of sample cells. The pooled prediction models were compared across data sets and time periods based on the coefficient of determination (adjusted-R^2) and the root mean square error (RMSE) of the cross-validation predictions (see Cohen et al., 2003) against observed values from both model sets. The model "offset", defined as the difference between the spectral response of the forest clearing and regeneration relationships when ΔFC=0% (i.e. the clearing start point subtracted from the regeneration start point), was calculated by subtracting the β_0 of the clearing model set from that of the regeneration set, in units of the spectral variable being modeled. The models developed from the sample set were tested by applying the RMA coefficients on the complete image data set, with the exception of the sample cells, and compared among data sets and time periods based on the complete predicted versus observed model correlation, RMSE, and the

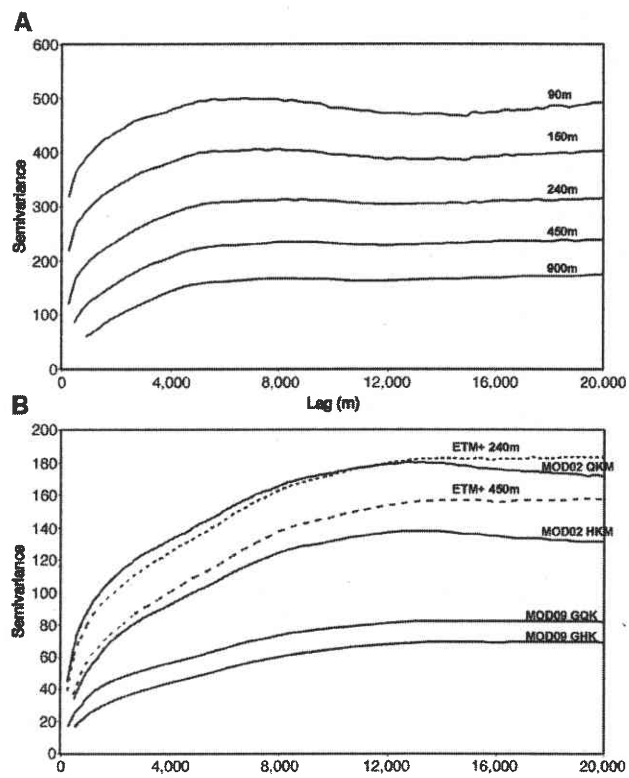

Fig. 3. Semivariograms calculated over a 75 km×75 km area in northern Guatemala comparing A) the reference ΔFC% (2000 to 2003) at varying spatial grain size and B) ΔNDVI (2000 to 2003) calculated from spatially degraded Landsat ETM+ data against ΔNDVI (2000 to 2003) from both MODIS Calibrated Radiances data (MOD02) and Surface Reflectance products (MOD09).

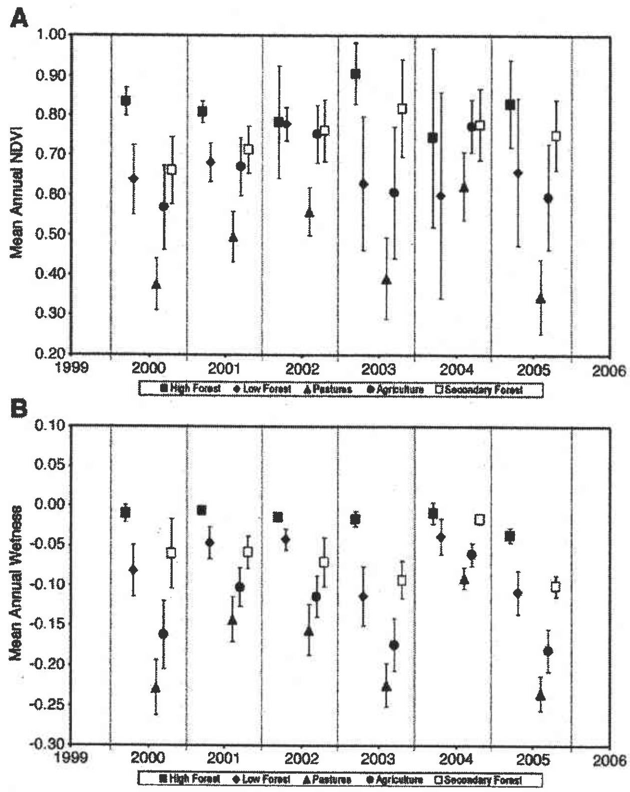

Fig. 4. Time-series plot of the spectral trajectories of forest and non-forest land cover types showing the mean NDVI (A) and tasseled cap wetness (B) between 2000 and 2005, calculated from normalized Landsat reflectance data for the selected training sites.

error as percentage of the range in observed values (see Cohen et al., 2003).

3. Results and discussion

3.1. Spatial patterns

The analysis of semivariance was used here for a number of reasons, including: 1) the establishment of a minimum distance rule for sample selection in model building; 2) to examine the effect of grain size on the spatial pattern of the property of interest, ΔFC; and 3) to assess the spatial contrast and image quality of the spectral change data amongst the different satellite data sets. Visual inspection of the variance in the reference ΔFC allowed for interpretation of the degree of spatial clustering of land cover change and therefore the distance over which the variable is spatially correlated. A range rule of 5000 m (the minimum distance between sample points used in regression modeling) was selected by visual interpretation of the semivariogram of the reference data (Fig. 3A). Establishing the minimum distance between any two sample points ensured that variables used in the modeling exercises were spatially independent and thereby did not violate that key assumption of linear regression modeling.

Semivariograms (Fig. 3) reveal the affect of varying spatial resolution and the different spectral data sets on the spatial structure of the observed reference ΔFC and the corresponding ΔNDVI from the image data. Within each semivariance plot, the range of autocorrelation remained constant across the different spatial resolutions, demonstrating that the observed changes in forest cover and NDVI were relatively invariant across the scales used in this analysis. Visual comparison between the semivariogram of the reference ΔFC data (Fig. 3A) and that of the image data (Fig. 3B) suggests that the range of autocorrelation in ΔFC is much lower than the approximately 12,000 m in ΔNVDI. Multiple factors may contribute to the difference in spatial pattern between the variables, including the interannual variability in spectral response due to changes in climatic and vegetation condition.

Both the reference ΔFC and spectral difference semivariograms revealed the predictable pattern of decreasing spatial heterogeneity, represented as the height of the sill of the semivariogram, as data are degraded from fine to coarse resolution. Comparing the different data sets in the semivariograms of the image data shows the spatial pattern of ΔNDVI from the MOD02 to follow that of the degraded Landsat data, with similar degrees of heterogeneity at a given spatial grain size. The low semivariance at the sill of the MOD09 data indicates a significant loss of heterogeneity, or spectral contrast in space, as compared to the ΔNDVI from the MOD02 and Landsat data at equivalent spatial resolution. This result is consistent with the findings of Tan et al. (submitted for publication), who also described the degradation of image quality in higher-level MODIS products as compared to reference data. They demonstrated that this phenomenon could be largely attributed to geolocation offsets, or "pixel shift", associated with the pre-defined sinusoidal grid cells within which MODIS products, including the MOD09, are resampled. Our results indicate that MODIS spectral data collected in the original swath format at the 250 and 500 m native cell resolutions are more representative of the spectral response of the reference Landsat data. Although not formally tested here, these results also raise the

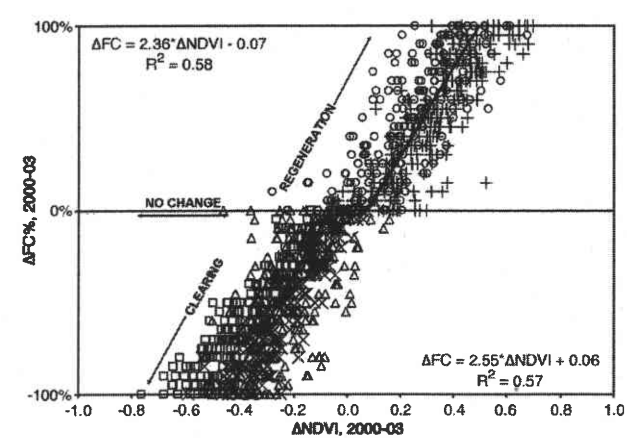

Fig. 5. Expected relationship of ΔFC to ΔNDVI between 2000 and 2003 developed by mixing the Landsat-derived spectral responses from the Monte-Carlo simulations, for different fractions of land cover types at each date. The possible land cover change trajectories shown on the plot are: mature high forest cleared for pasture (□); mature high forest cleared for agricultural plots (X); low stature forest cleared for pasture (Δ); agricultural plots regenerated to young forest (O); and pasture regenerated to young forest (+). The RMA best-fit line, regression model equation and adjusted-R^2 are given for each model set (clearing and regeneration).

possibility of a "smoothing" of the data by the use of more coarse resolution spectral information in the surface reflectance algorithms used to produce the MOD09 product (Vermote et al., 1997).

3.2. Spectral trajectories

The analysis of the spectral signatures of different land cover types illustrates the expected behavior of different spectral indices as they change over time as a result of the potential trajectories of forest clearing, regeneration and no change. A time-series plot of the spectral trajectories (Fig. 4) shows the variability in spectral response to year-to-year changes in land cover condition in the absence of forest clearing and regeneration. The data points show the annual mean and standard deviation in both NDVI (Fig. 4A) and TCW (Fig. 4B) by land cover type, as calculated from the training sites. Each data point in Fig. 4 represents the mean and standard deviation in NDVI response used to create the normal distributions (for each land cover type, each year) from which random variates were generated in the Monte-Carlo analysis (Section 3.3.1 below) of the effect of proportional ΔFC on spectral changes.

When plotting the normalized reflectance values from Landsat data over time, even relatively homogeneous training areas of known, unchanged land cover types show significant interannual variability in spectral response, both in NDVI and TCW. The low forest and agriculture types tend to have greater within-class variability (higher standard deviations) than the more homogeneous high forest and pasture types. The analysis demonstrates the effect of changes in climatic and vegetation condition from year to year. While the spectral response in mature forest is more consistent from year to year, especially with the TCW index, there remains higher variability in agriculture and pasture due to various factors including atmospheric conditions, soil moisture, vegetation phenology and the presence or absence of crops.

In lieu of the interannual variability in spectral response, the analysis does reveal predictable patterns of spectral change from one date to the next that give rise to different spectral "starting points" along the different land cover change trajectories. For example, tracking the mean NDVI response of the training areas shows that, while unchanged land cover types may have a different spectral response each year, in general forest types move

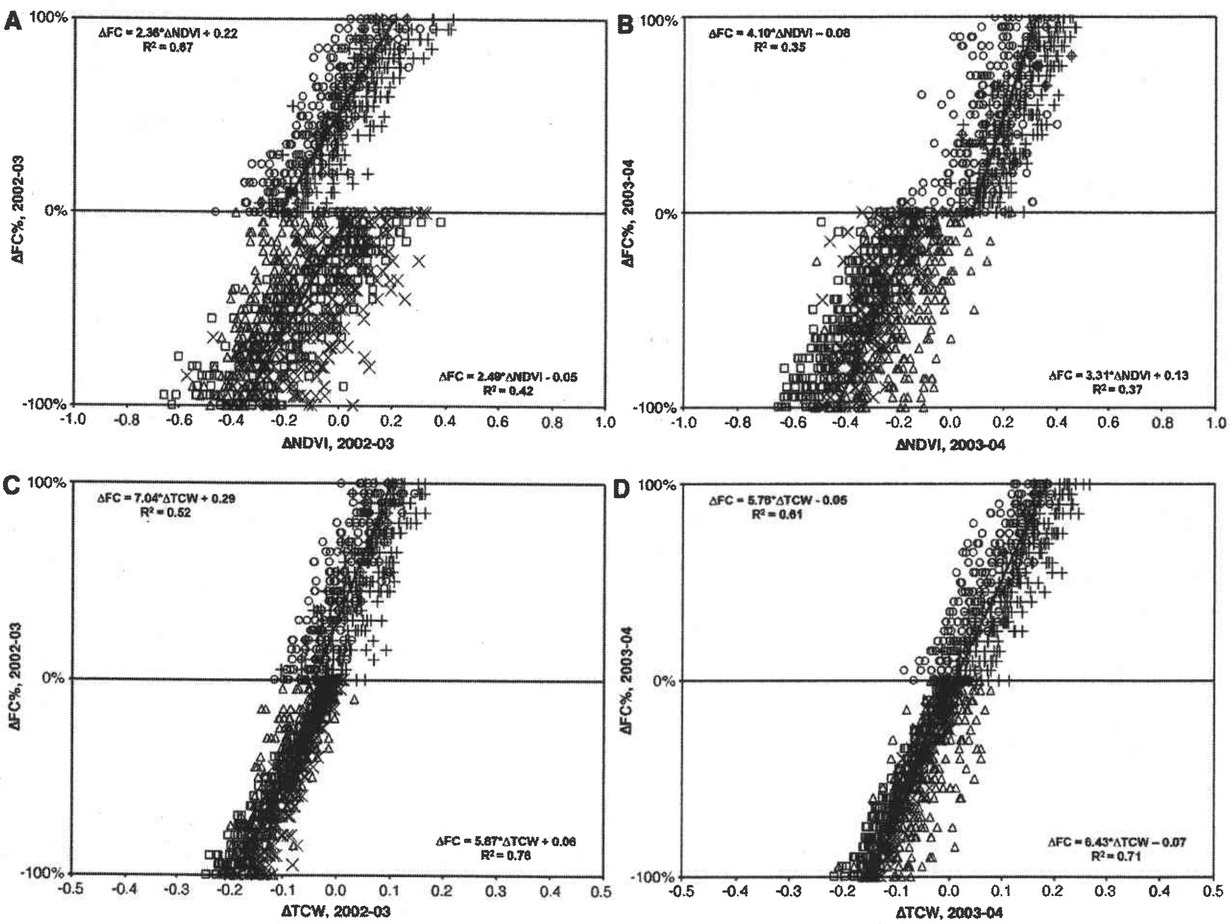

Fig. 6. Variability in the relationship of ΔFC to ΔNDVI and ΔTCW over two one-year time periods developed by mixing the Landsat-derived spectral responses from the Monte-Carlo simulations, for different fractions of land cover types at each date. The relationship is shown for ΔNDVI between 2002 and 2003 (A) and 2003 and 2004 (B), as well as for ΔTCW between 2002 and 2003 (C) and 2003 and 2004 (D). The possible land cover change trajectories shown on the plot are: mature high forest cleared for pasture (□); mature high forest cleared for agricultural plots (X); low stature forest cleared for pasture (Δ); agricultural plots regenerated to young forest (O); and pasture regenerated to young forest (+). The RMA best-fit line, regression model equation and adjusted-R^2 are given for each model set (clearing and regeneration).

in the opposite direction in spectral space than non-forest types. These patterns can be exploited in the modeling effort by identifying the different starting points of the possible spectral trajectories. Furthermore, the analyses confirm the presence of a different spectral relationship for forest conversion to pasture versus that of the regeneration of young forest on abandoned or fallow agriculture plots. As such, the different land cover trajectories (clearing and regeneration) need to be identified and modeled separately. This will bear out in different relationships in ΔFC for different time periods, including the slope of each relationship and the starting point along the ΔFC=0% axis, according to the year-to-year variability in climatic conditions, particularly seasonal precipitation levels.

3.3. Estimating forest cover change

3.3.1. Expected spectral changes with ΔFC

A plot of simulated proportions of land cover change against expected spectral changes in NDVI from the Monte-Carlo analysis of the Landsat signatures of different land cover types (Fig. 5) promotes interpretation of the expected relationship of ΔFC to ΔNDVI. The points below the ΔFC=0% axis represent different proportions of forest clearing, i.e. high forest or low forest at the first date converted to some proportion (0 to 100%) of pasture or agriculture at the second date. Similarly, the points above the ΔFC=0% axis represent proportions of forest regeneration from abandoned or fallow agriculture or pasture. At ΔFC=0%, the spread of points across the range of ΔNDVI represents spectral changes associated with the interannual variation in land cover condition, described above in the temporal analysis of spectral trajectories.

The RMA best-of-fit line for the relationship is shown on the plot for both forest clearing and regeneration. The two relationships have similar values for the slope that relates ΔNDVI to ΔFC (2.55 for forest clearing and 2.36 for regeneration). The plot also reveals the "offset" phenomenon that suggests that forest clearing and regeneration have different starting points, as measured along the ΔFC=0% axis. This offset is an artifact of the year-to-year difference in the mean ΔNDVI between unchanged forest types and unchanged non-forest types. It is related to the pattern, described above, where forest and non-forest move in opposite directions in spectral space in response to changes in land cover condition. In Fig. 5, the forest clearing best-fit line intersects the ΔFC=0% axis at a ΔNDVI of −0.06, and at +0.07 for forest regeneration, resulting in an offset (in ΔNDVI units) of −0.14.

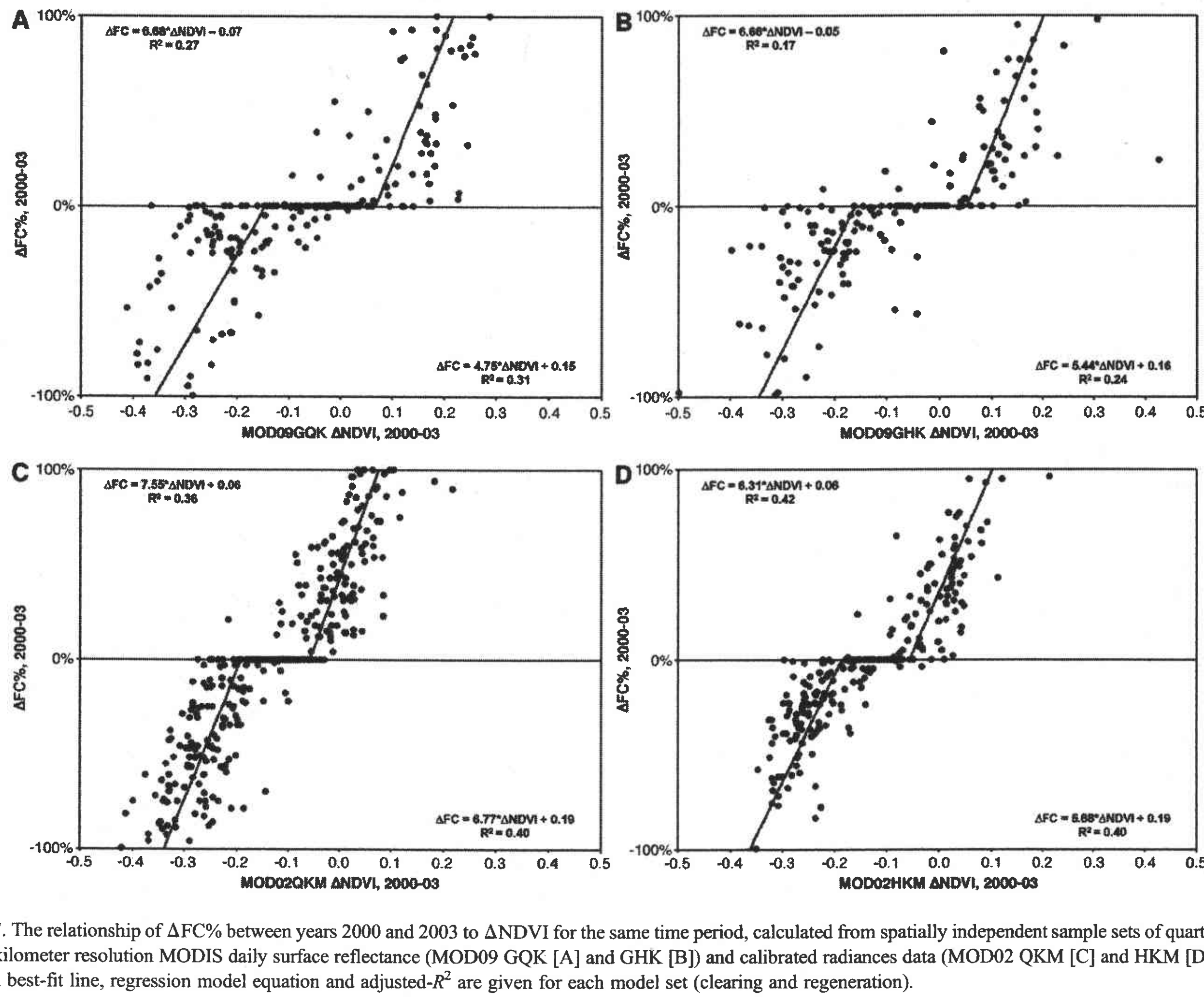

Fig. 7. The relationship of ΔFC% between years 2000 and 2003 to ΔNDVI for the same time period, calculated from spatially independent sample sets of quarter- and half-kilometer resolution MODIS daily surface reflectance (MOD09 GQK [A] and GHK [B]) and calibrated radiances data (MOD02 QKM [C] and HKM [D]). The RMA best-fit line, regression model equation and adjusted-R^2 are given for each model set (clearing and regeneration).

Plots of simulated ΔFC against ΔNDVI and ΔTCW for two one-year time periods (Fig. 6) provide an example of the interannual variability that can be encountered in this relationship. Fig. 6a and b show considerable variation in spectral response to changes in land cover condition, i.e. the variation in ΔNDVI where no land cover changes have occurred. In particular, the spread of points along the ΔFC=0% axis is wide enough to drown changes in the spectral signal associated with significant proportions of land cover clearing or regeneration. The slopes of the relationship of ΔFC to ΔNDVI can vary between time periods, with the best-fit lines for forest clearing and regeneration between 2003 and 2004 (3.31 and 4.10) being much steeper than those between 2002 and 2003 (2.49 and 2.36). The plots also demonstrate the variability in the nature of the offset between different time periods, i.e. the ΔNDVI for unchanged forest can be either greater or less than that of unchanged non-forest, depending on year-to-year changes in vegetation condition. By comparison, the response of the ΔTCW metric to ΔFC (Fig. 6c and d) shows less variability to changes in condition than does the ΔNDVI. The result is an improved ability to distinguish and model land cover clearing and regeneration, with more consistent starting points for the change trajectories and higher adjusted-R^2 values for the ΔFC on ΔTCW relationship.

The simulation of spectral changes by a controlled mixing of the spectral signatures over the range of change trajectories provides insight into the expected relationship of varying proportions of forest cover change to the different vegetation indices. This analysis also suggests the utility of including additional spectral information in the modeling of ΔFC, specifically from the SWIR wavelengths. The TCW index, which here emphasizes the contrast in Landsat bands 5 and 7, is less affected by changes in land cover condition but more sensitive to the changes of interest (clearing and regeneration) than the NDVI. The result is less residual variation overall between ΔFC and ΔTCW than ΔNDVI across the time periods analyzed in this study. Furthermore, the spectral starting points of the clearing and regeneration trajectories are closer and more consistent in the ΔFC on ΔTCW relationship, indicating the greater utility in modeling with this index over the ΔNDVI.

3.3.2. Estimating ΔFC with coarse-resolution MODIS spectral data

With an understanding of the expected relationship of ΔFC to changes in spectral response based on the Landsat-derived signatures, the various MODIS data, products, cell sizes and spectral bands were evaluated for their ability to estimate proportional changes in forest cover. The relationship of ΔFC to ΔNDVI calculated from two dates (2000 and 2003) of MODIS spectral data is shown in Fig. 7. The four plots compare this relationship among two different MODIS data sets (MOD02 and MOD09) and at two spatial resolutions (quarter- and half-kilometer).

Table 3
RMA Regression model statistics for each data set, variable, model set and time period

Time period	Data set	Variable	Model set	n	β_0	β_1	R	Start pt.	"Offset"	R^2	RMSE
2000 to 2001											
	MODO2 QKM	ΔNDVI	ΔFC≥0	145	−0.28	6.32	0.46	0.04	0.06	0.79	0.17
			ΔFC≤0	198	0.07	5.86	0.68	−0.01			
	MODO2 HKM	ΔNDVI	ΔFC≥0	114	−0.16	5.23	0.54	0.03	0.04	0.72	0.20
			ΔFC≤0	131	0.04	4.90	0.57	−0.01			
		ΔNDMI		114	−0.03	3.80	0.58	0.01	0.00	0.73	0.17
				131	−0.04	3.01	0.59	0.01			
2001 to 2002											
	MODO2 QKM	ΔNDVI	ΔFC≥0	155	−0.21	5.40	0.51	0.04	0.07	0.72	0.19
			ΔFC≤0	167	0.21	6.09	0.50	−0.03			
	MODO2 HKM	ΔNDVI	ΔFC≥0	101	−0.32	3.99	0.54	0.08	0.09	0.62	0.24
			ΔFC≤0	110	0.03	2.92	0.53	−0.01			
		ΔNDMI	ΔFC≥0	101	0.16	3.31	0.64	−0.05	0.07	0.64	0.17
			ΔFC≤0	110	0.21	1.75	0.52	−0.12			
2002 to 2003											
	MODO2 QKM	ΔNDVI	ΔFC≥0	117	0.36	2.55	0.42	−0.14	−0.12	0.66	0.18
			ΔFC≤0	210	0.05	2.18	0.72	−0.02			
	MODO2 HKM	ΔNDVI	ΔFC≥0	89	0.64	3.33	0.48	−0.19	−0.18	0.66	0.21
			ΔFC≤0	149	0.05	3.28	0.64	−0.12			
		ΔNDMI	ΔFC≥0	89	0.33	2.51	0.54	−0.13	−0.10	0.67	0.18
			ΔFC≤0	149	0.08	2.55	0.64	−0.03			
2000 to 2003											
	MODO2 QKM	ΔNDVI	ΔFC≥0	175	0.45	4.79	0.79	−0.09	0.06	0.81	0.17
			ΔFC≤0	179	0.68	4.36	0.75	−0.16			
	MODO2 HKM	ΔNDVI	ΔFC≥0	101	0.35	4.30	0.70	−0.08	0.07	0.77	0.16
			ΔFC≤0	141	0.52	3.42	0.69	−0.15			
		ΔNDMI	ΔFC≥0	101	−0.18	3.37	0.78	0.05	0.10	0.81	0.14
			ΔFC≤0	141	0.11	2.67	0.78	−0.04			

ΔFC≥0 refers to the forest regeneration model set; ΔFC≤0 is the forest clearing model set.
Model R^2 was calculated from the predicted versus observed values of the combined (regeneration and clearing) model sets.
Model RMSE was calculated by combining the cross-validation predictions against observed values from both model sets.

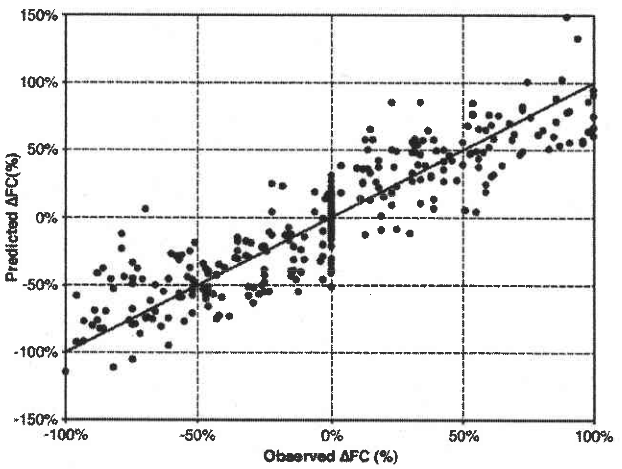

Fig. 8. Predicted versus observed ΔFC% between years 2000 and 2003. The predictions in this example were generated by combining the predictions of the separate clearing and regeneration model sets, which were modeled on ΔNDVI (2000 to 2003) calculated from the MOD02 QKM data set.

The analysis here provided further evidence of the improved performance in modeling ΔFC with MOD02 Calibrated Radiances data over the MOD09 Surface Reflectance product. The MOD09 shows considerably more scatter in ΔNDVI across all levels of ΔFC than does the MOD02 data, which is reflected in the lower adjusted-R^2 values of MOD09 for forest clearing and regeneration, and the greater offsets between the "starting points" of unchanged land cover types. The MOD02 data do not show a significant loss of predictability in the relationship of ΔFC to ΔNDVI between the QKM and HKM pixel resolutions. The greater scatter in ΔNDVI calculated from MOD09 may be a function of the pixel shift problem and/or the surface reflectance algorithm, as described above. A better relationship to the reference ΔFC was found using ΔNDVI calculated from MOD02 data, with good correlation at QKM resolution and a small decrease in adjusted-R^2 when modeling forest clearing and regeneration with HKM data.

Regression model statistics from the spatially independent sample data demonstrate the effect of the different grain size and MOD02 spectral indices on modeling forest clearing, regeneration, and overall ΔFC (Table 3). The overall ΔFC models were evaluated by combining the separate clearing and regeneration model sets into one predicted versus observed model (Fig. 8). The amount of variation in overall ΔFC explained by the spectral data, as measured by the coefficient of determination (R^2) of the combined model, shows a generally decreasing trend from the QKM to HKM data sets across all time periods. The QKM ΔNDVI model did perform significantly better than that at HKM resolution in the first two time periods (2000 to 2001; 2001 to 2002), which had a lower overall amount of forest cover change than the other two time periods in the analysis (Table 2). Modeling ΔFC on only the ΔNDVI, then, confirms the importance of higher spatial resolution in detecting small amounts of scattered change. On the other hand, the ΔNDMI model outperformed the ΔNDVI at the same resolution (HKM), and generally had lower, but similar, adjusted-R^2 results compared to the ΔNDVI at QKM resolution. The ΔNDMI metric resulted in high correlation with both the clearing and regeneration model sets, showed the least variability between unchanged forest and non-forest cover (i.e. had the smallest "offsets"), and had high overall R^2 for all time periods.

These results suggest that, while spatial resolution is important, modeling ΔFC is relatively invariant to cell size (between 250 m and 500 m) and that including spectral information from the SWIR wavelengths improves model performance over using only R and NIR (i.e. the NDVI). The importance of including SWIR information in the modeling forest cover change was shown using the TCW index calculated from Landsat reflectance data. While the transformation coefficients for tasseled cap indices are available for MODIS data (Lobser & Cohen, submitted for publication), they were developed using the global 1-km resolution BRDF product but were not applied in this study to the 250 and 500 m MOD02 data. The NDMI was used instead because of its ease in calculation from MOD02 data and because it has been demonstrated as a good substitute for TCW with data sets for which tasseled cap coefficients are not available (Jin & Sader, 2005). Using the NDMI index satisfied the objective of testing the utility of SWIR information in modeling ΔFC, whereas the particular index itself was not as important.

3.3.3. Model testing

The model coefficients (β_0 and β_1) shown in Table 3 represented the best-fit RMA lines for each model set, spectral index, cell size and time period. These coefficients were used to transform the spectral data to predictions of ΔFC for the remaining pixels in each data set (i.e. pixels not used in the sample). Model prediction results were compared and evaluated (Table 4) based on their correlation (R) with the observed values of ΔFC, the root mean square error (RMSE) of the residuals and this error as a percent of the range of observed values (%Error). Whole image modeling of forest cover change demonstrated the promise of using MODIS data to monitor proportional changes in forest cover across the range of conversion and modification (ΔFC from −100 to +100%). Cell-by-cell prediction of ΔFC over the entire study area shows errors ranging from 7.4 to 10.9% across the time periods of analysis.

Table 4
Whole image model testing results for each data set, variable and time period

Time period	Data set	Variable	n	R	RMSE	%Error
2000 to 2001						
	MODO2 QKM	ΔNDVI	383,185	0.47	0.18	9.07%
	MODO2 HKM	ΔNDVI	86,303	0.50	0.19	9.42%
		ΔNDMI	86,303	0.58	0.17	8.74%
2001 to 2002						
	MODO2 QKM	ΔNDVI	291,578	0.38	0.20	10.05%
	MODO2 HKM	ΔNDVI	56,855	0.37	0.22	10.93%
		ΔNDMI	56,855	0.45	0.20	9.88%
2002 to 2003						
	MODO2 QKM	ΔNDVI	400,294	0.53	0.18	8.97%
	MODO2 HKM	ΔNDVI	94,910	0.54	0.19	9.34%
		ΔNDMI	94,910	0.54	0.17	8.61%
2000 to 2003						
	MODO2 QKM	ΔNDVI	436,837	0.61	0.18	9.00%
	MODO2 HKM	ΔNDVI	103,968	0.58	0.18	8.80%
		ΔNDMI	103,968	0.57	0.15	7.39%

Comparing the ΔNDVI models between cell sizes, the % Error is generally similar within each time period, with only a slight, if any, increase in error when moving from QKM to HKM resolution. With the ability to predict change relatively invariant between 250 m and 500 m, the model test results also support the value of using spectral information from the SWIR wavelengths at 500 m over using only NDVI at higher resolution (250 m). Within each time period, the lowest %Error amongst the models is consistently found with the ΔNDMI data sets. Model correlation (predicted vs. observed) is similar for ΔNDVI at QKM and HKM resolution, and increases when modeling with the ΔNDMI metric at HKM cell size. Across indices and cell sizes, model R is highest in the three-year time period (2000–2003) where there was a longer amount of time for changes to accumulate, resulting in the greatest area affected by change amongst all of the time periods (Table 2).

4. Summary and conclusions

This study provides the foundation for understanding the spatial, spectral and temporal patterns of land cover change in a dynamic tropical forest ecosystem. The analyses demonstrate many of the issues involved with scaling observations of forest cover change within a monitoring system based on a high temporal frequency, coarse spatial resolution spectral data set. The direct comparison of change patterns observed in the Landsat data set with the spectral response of concurrent dates of MODIS data offered several key findings, including:

- The different general land cover change trajectories (forest clearing, regeneration and change in condition) resulted in distinct spectral response patterns;
- Spectral response in the MOD02 Calibrated Radiances Swath data set followed more closely with the expected patterns of forest cover change than did the spectral response in the MOD09 Surface Reflectance Gridded data set;
- The ability to detect forest cover change patterns was relatively invariant to the spatial grain size of the analysis (250 m and 500 m); and
- Including additional spectral information, particularly from the SWIR wavelengths, improved the ability to model changes in forest cover.

The results indicate that the best metrics for detecting tropical forest clearing and regeneration are those based on vegetation indices which incorporate the SWIR bands available in the MOD02 data set at 500-m resolution. Franklin et al. (2001), Cohen et al. (2002), Wilson and Sader (2002), and Healey et al. (2005) found similar results in studies using Landsat-derived, SWIR-based indices for detecting forest cover changes in northern temperate and boreal forests. Furthermore, models based on linear regression techniques showed promise in predicting the full range of changes in forest cover as a continuous variable with good accuracy.

The analyses and results from this study also highlight important issues for further investigation in the effort to develop a regional forest cover monitoring protocol based on coarse resolution MODIS data. As with most satellite image change detection studies, there is a clear need to improve the capability for separating the changes of interest from background variation in the spectral response. Techniques such as the thresholding of image difference histograms or multivariate clustering techniques warrant investigation for their ability to remove unchanged land cover types prior to regression modeling. Other multivariate data transformations may prove to better isolate changes due to forest clearing and regeneration from atmospheric influences and changes in land cover condition.

This study demonstrated the improved ability to model forest cover change with MOD02 data over MOD09 product. However, these at-satellite calibrated radiance data (MOD02) are more likely to be affected year-to-year variation in atmospheric, seasonal and aerosol conditions than atmospherically-corrected reflectance data sets. While the absolute correction of these data to reflectance is not critical to modeling spectral patterns over time (Schroeder et al., 2006), the normalization of these time-series data sets and the consistency of the model parameters will be essential in the effort to expand these models in an operational mode over time and space. Other MODIS products such as the EVI and BRDF data sets employ algorithms that may result in more consistent patterns over time and allow better discrimination of vegetation parameters (Diner et al., 2005; Huete et al., 2002; Lobser & Cohen, submitted for publication; Mildrexler et al., in press). Given that the patterns observed in this analysis were relatively invariant to scale (250 m and 500 m), these products are being evaluated for their utility in estimating forest cover change over time at more coarse resolution (1 km), which may reduce the geolocation offset problem of the MODIS sinusoidal grid (Tan et al., submitted for publication) and be more practical for forest cover monitoring over larger areas.

Acknowledgements

This research was funded by a NASA Earth System Science Fellowship (NGT5-30526). Image and field data supports were provided by ProPetén in Guatemala and the NASA/USAID SERVIR Data Center in Panamá. We greatly appreciate the indispensable assistance in research development from W. David Ritts of Oregon State University and Thomas K. Maiersperger at the USGS EROS Data Center. We also acknowledge the help of Dr. Steven A. Sader of the University of Maine for the many discussions on, and reviews of, this research project. Anonymous reviewers provided valuable comments which ultimately greatly improved the organization, clarity and quality of the manuscript.

References

Adams, J. B., Sabol, D., Kapos, V., Filho, R. A., Roberts, D. A., Smith, M. O., et al. (1995). Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, *52*, 137–154.

Atkinson, P. M. (1991). Optimal ground-based sampling for remote sensing investigations: Estimating the regional mean. *International Journal of Remote Sensing*, *12*, 559–567.

Berterretche, M., Hudak, A. T., Cohen, W. B., Maiersperger, T. K., Gower, S. T., & Dungan, J. (2005). Comparison of regression and geostatistical methods for mapping Leaf Area Index (LAI) with Landsat ETM+ data over a boreal forest. *Remote Sensing of Environment*, *96*, 49–61.

Chavez, P. S., Jr (1996). Image-based atmospheric corrections — Revisited and improved. *Photogrammetric Engineering and Remote Sensing*, *62*(9), 1025–1036.

Cohen, W. B., & Goward, S. N. (2004). Landsat's role in ecological applications of remote sensing. *Bioscience*, *54*(6), 535–545.

Cohen, W. B., Fiorella, M., Gray, G., Helmer, E., & Anderson, K. (1998). An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery. *Photogrammetric Engineering and Remote Sensing*, *64*, 293–300.

Cohen, W. B., Maiersperger, T. K., Gower, S. T., & Turner, D. P. (2003). An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing of Environment*, *84*, 561–571.

Cohen, W. B., & Spies, T. A. (1992). Estimating structural attributes of Douglas-fir/western hemlock forest stands from LANDSAT and SPOT imagery. *Remote Sensing of Environment*, *41*, 1–17.

Cohen, W. B., Spies, T. A., Alig, R. J., Oetter, D. R., Maiersperger, T. K., & Fiorella, M. (2002). Characterizing 23 years (1972–1995) of stand replacement disturbance in western Oregon forests with Landsat imagery. *Ecosystems*, *5*, 122–137.

Cohen, W. B., Spies, T. A., & Bradshaw, G. A. (1990). Semivariograms of digital imagery for analysis of conifer canopy structure. *Remote Sensing of Environment*, *34*, 167–178.

CONAP-WCS. (2005). Perspectivas esperadas, temporada de incendios 2005. Centro de Monitoreo y Evaluación de CONAP-WCS.

Coppin, P., Jonckheere, I., Nackaerts, K., Muysand, B., & Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: A review. *International Journal of Remote Sensing*, *25*(9), 1565–1596.

Crist, E. P. (1985). A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, *17*, 301–306.

Crist, E. P., & Cicone, R. C. (1984). A physically-based transformation of Thematic Mapper data — The TM tasseled cap. *IEEE Transactions on Geoscience and Remote Sensing*, *22*, 256–263.

Curran, P. J. (1988). The semiovariogram in remote sensing. *Remote Sensing of Environment*, *24*, 493–507.

Curran, P. J., & Atkinson, P. M. (1998). Remote sensing and geostatistics. *Progress in Physical Geography*, *22*(1), 61–78.

DeFries, R. S., Hansen, M. C., & Townshend, J. R. G. (2000). Global continuous fields of vegetation characteristics: A linear mixture model applied to multi-year 8 km AVHRR data. *International Journal of Remote Sensing*, *21*, 1389–1414.

Diner, D. J., Braswell, B. H., Davies, R., Gobron, N., Hu, J., Jin, Y., et al. (2005). The value of multiangle measurements for retrieving structurally and radiatively consistent properties of clouds, aerosols, and surfaces. *Remote Sensing of Environment*, *97*, 495–518.

Fiorella, M., & Ripple, W. J. (1993). Determining successional stage of temperate coniferous forests with Landsat satellite data. *Photogrammetric Engineering and Remote Sensing*, *59*, 239–246.

Foody, G. M., & Cox, D. P. (1994). Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing*, *15*, 619–631.

Foody, G. M., Lucas, R. M., Curran, P. J., & Honzak, M. (1997). Non-linear mixture modeling without end-members using an artificial neural network. *International Journal of Remote Sensing*, *18*, 937–953.

Franklin, S. E., Lavigne, M. B., Moskal, L. M., Wulder, M. B., & McCaffrey, T. M. (2001). Interpretation of forest harvest conditions in New Brunswick using Landsat TM Enhanced Wetness Difference Imagery (EWDI). *Canadian Journal of Remote Sensing*, *27*(2), 118–128.

Fung, T., & LeDrew, E. (1987). Application of principal component analysis to change detection. *Photogrammetric Engineering and Remote Sensing*, *53* (12), 1649–1658.

Fung, T., & LeDrew, E. (1988). The determination of optimal threshold levels for change detection using various accuracy indices. *Photogrammetric Engineering and Remote Sensing*, *54*(10), 1449–1454.

Hall, F. G., Strebel, D. E., Nickeson, J. E., & Goetz, S. J. (1991). Radiometric rectification: Toward a common radiometric response among multidate, multisensor images. *Remote Sensing of Environment*, *35*, 11–27.

Hansen, M., DeFries, R., Townshend, J. R. G., Sohlberg, R., Dimiceli, C., & Carroll, M. (2002). Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data. *Remote Sensing of Environment*, *83*, 303–319.

Hardisky, M. A., Klemas, V., & Smart, R. M. (1983). The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of *Spartina alterniflora* canopies. *Photogrammetric Engineering and Remote Sensing*, *49*, 77–83.

Hayes, D. J., & Sader, S. A. (2001). Change detection techniques for monitoring forest clearing and regrowth in a tropical moist forest. *Photogrammetric Engineering and Remote Sensing*, *67*(9), 1067–1075.

Hayes, D. J., Sader, S. A., & Schwartz, N. B. (2002). Developing a forest conversion history database to explore the temporal and spatial characteristics of land cover change in the Maya Biosphere Reserve, Guatemala. *Landscape Ecology*, *17*, 299–314.

Healey, S. P., Cohen, W. B., Zhiqiang, Y., & Krankina, O. (2005). Comparison of tasseled cap-based Landsat data structures for forest disturbance detection. *Remote Sensing of Environment*, *97*, 301–310.

Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., & Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195–213.

Iverson, L. R., Cook, E. A., & Graham, R. L. (1989). A technique for extrapolating and validating forest cover across large regions: calibrating AVHRR data with TM data. *International Journal of Remote Sensing*, *10*, 1805–1812.

Jensen, J. R. (1996). *Introductory digital image processing*, Second edition Upper Saddle River, NJ: Prentice-Hall 316 pp.

Jin, S., & Sader, S. A. (2005). Comparison of time-series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sensing of Environment*, *94*(3), 364–372.

Kennedy, R. E., & Cohen, W. B. (2003). Automated designation of tie-points for image-to-image coregistration. *International Journal of Remote Sensing*, *24* (17), 3467–3490.

Lobser, S. E. and Cohen, W. B. (Submitted for publication). The MODIS tasseled cap: Land cover characteristics expressed through transformed MODIS data. *International Journal of Remote Sensing*, June 2006.

Mildrexler, D. J., Zhao, M., Heinsch, F. A., and Running, S. W. (in press). A new satellite based methodology for continental scale disturbance detection. *Ecological Applications*.

Morton, D. C., DeFries, R. S., Shimabukuro, Y. E., Anderson, L. O., Del Bon Espírito-Santo, F., Hansen, M., et al. (2005). Rapid assessment of annual deforestation in the Brazilian Amazon using MODIS data. *Earth Interactions*, *9*, 1–22.

Muchoney, D. M., & Haack, B. N. (1994). Change detection for monitoring forest defoliation. *Photogrammetric Engineering and Remote Sensing*, *60*, 1243–1251.

Nations, J. D., Primack, R. B., & Bray, D. (1998). Introduction: The Maya forest. In R. B. Primack, D. Bray, H. A. Galleti, & I. Ponciano (Eds.), *Timber, tourists, and temples.* Washington, D.C.: Island Press.

Neilsen, A. A., Conradsen, K., & Simpson, J. J. (1998). Multivariate Alteration Detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sensing of Environment*, *64*, 1–19.

Sader, S. A. (1995). Spatial characteristics of forest clearing and regrowth as detected by Landsat Thematic Mapper Imagery. *Photogrammetric Engineering and Remote Sensing*, *61*, 1145–1151.

Sader, S. A., Hayes, D. J., Hepinstall, J. A., Coan, M., & Soza, C. (2001). Forest change monitoring of a remote biosphere reserve. *International Journal of Remote Sensing*, *22*, 1937–1950.

Sader, S. A., Reining, C., Sever, T., & Soza, C. (1997). Human migration and agricultural expansion: A threat to the Maya tropical forests. *Journal of Forestry*, *95*(12), 27–32.

Schroeder, T. A., Cohen, W. B., Song, C., Canty, M. J., & Yang, Z. (2006). Radiometric correction of multitemporal Landsat data for characterization of early successional forest patterns in western Oregon. *Remote Sensing of Environment*, *103*, 16–26.

Singh, A. (1989). Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, *10*, 989–1003.

Tan, B., Woodcock, C. E., Hu, J., Zhang, P., Ozdogan, M., Huang, D., Yang, W., Knyazikhin, Y., & Myneni, R. B. (submitted for publication). The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions. *Remote Sensing of Environment*, November 2005.

Tian, Y., Woodcock, C. E., Wang, Y., Privette, J. L., Shabanov, N. V., Zhou, L., et al. (2002). Multiscale analysis and validation of the MODIS LAI product. I. Uncertainty assessment. *Remote Sensing of Environment*, *83*, 414–430.

Townshend, J. R. G., & Justice, C. O. (2002). Towards operational monitoring of terrestrial systems by moderate-resolution remote sensing. *Remote Sensing of Environment*, *83*, 351–359.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*, 127–150.

Vermote, E. F., El Saleous, N. Z., Justice, C. O., Kaufman, Y. J., Privette, J., Remer, L., et al. (1997). Atmospheric correction of visible to middle infrared EOS-MODIS data over land surface, background, operational algorithm and validation. *Journal of Geophysical Research*, *102*(14), 17131–17141.

Wilson, E. H., & Sader, S. A. (2002). Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sensing of Environment*, *80*, 385–396.

Zhan, X., Sohlberg, R., Townshend, J. R. G., DiMiceli, C., Carroll, M., Eastman, J. C., et al. (2002). Detection of land cover changes using MODIS 250 m data. *Remote Sensing of Environment*, *83*, 337–351.

Zhu, Z. L., & Evans, D. L. (1992). Mapping midsouth forest distributions. *Journal of Forestry*, *90*, 27–30.