Supporting Information

Redo et al. 10.1073/pnas.1201664109

SI Materials and Methods

Classification of Satellite Imagery to Map Land-Use and Land-Cover. To map land-use and land-cover (LULC) classes, we used the MODIS MOD13Q1 (16-d L3 Global 250 m) product (1, 2). The product is a 16-d composite of the highest-quality pixels from daily images and includes the Enhanced Vegetation Index (EVI), red, near infrared (NIR), and midinfrared (MIR) reflectance and pixel reliability (3). There are 23 samples available per year, with data available from 2001 to the present. For each pixel, we calculated the following statistics: mean, SD, minimum, maximum, and range for EVI, and red NIR and MIR reflectance values from calendar years 2001–2010. These statistics were calculated for all 12 mo (annual), two 6-mo periods, and three 4-mo periods. The MOD13Q1 pixel reliability layer was used to remove all unreliable samples (value = 3) before calculating statistics.

Reference data for classifier training and accuracy assessment were collected with human interpretation of high-resolution imagery in Google Earth using a Web-based tool called VIEW-IT (Virtual Interpretation of Earth Web-Interface Tool) (2). The VIEW-IT tool uses a GE plug-in to allow users to visually estimate percent cover of LULC within a sample grid defined by a 250-m MODIS pixel overlaid on high-resolution GE satellite imagery. GE provides high-resolution, geo-referenced imagery from datasources, such as DigitalGlobe, GeoEye-1, IKONOS, EarthSat, and TerraMetric, with spatial resolutions often as fine as submeter to 4 m. Thus, GE images are similar in detail to aerial photographs, which are a common source for accuracy assessment (4) and allow a form of visual accuracy assessment. In Central America, most imagery is from DigitalGlobe's QuickBird satellite, with resolutions as fine as 2.4 m to 2.8 m. Samples were placed only in areas with high-resolution QuickBird imagery, with locations selected both randomly and manually within patch types for the corresponding land-cover classes (2). All samples were more than 1,000-m apart to avoid spatial autocorrelation, which resulted in a total of 4,560 individual training points for all of Central America.

After point placement, at least two independent users visually estimated percent cover of a particular LULC class to the nearest 10% and recorded the year of GE imagery. All training samples corresponded to eight distinct classes for all years between 2001 and 2010, following definitions used in previous studies (1, 2). Our maps were created with three separate classifiers trained on reference data from the regional biomes (i.e., moist forest, dry forest, and conifer forest) of Central America (5), with borders defined by municipalities assigned to one of the three biomes based on majority cover. For mapping purposes, municipalities belonging to Mangroves and Desert and Xeric Shrubland biomes (i.e., less the 1% of the total area) were subsumed into the most dominant surrounding biome. Predictor variables were MODIS-derived 4-, 6-, and 12-mo statistics extracted for the year corresponding to the GE image year (range 2002–2010) for each reference sample. For each biome map, a Random Forests (RF) classifier was generated with 1,999 trees.

RFs have proved useful in classifying and detecting land-cover and land-use elsewhere (1, 6–11). RFs are considered the cutting edge of land-use and land-cover classification (7, 8). The RFs are decision-tree algorithms that use bootstrap samples with replacement to grow a large set of classification trees (12). Pixels are assigned to the classes that receive the most votes from the user-specified number of classification trees. These classifiers offer a number of advantages over more traditional classifiers, but the most important is that they do not overfit the data (11).

This classification procedure resulted in land-change maps with eight classes for each year from 2001 to 2010. For the purposes of this study, only the woody vegetation (trees and shrubs that cover greater than or equal to 80% of the pixel) and agricultural/herbaceous (annual crop, grasslands, and pastures where cover is greater than 80%) classes were used. We focused mainly on trends in the woody vegetation class because it represents change associated with natural vegetation, such as deforestation or reforestation, which has important implications for species habitat use, carbon dynamics, and forest transition (FT) theory. Map classification accuracy assessment was based on the information obtained from reference data that were not used in training an individual tree in the RF (1, 2, 12). The average overall accuracy for the three biome maps that covered Central America was 85.1%. Average producer's accuracy for the two classes under investigation—woody vegetation and agricultural/herbaceous vegetation—was slightly higher at 86.7% and 85.1%, respectively. User's accuracy was similar for the two classes at 86.7% (woody) and 84.2% (agricultural/herbaceous vegetation).

The analysis of the trends in forest (i.e., woody vegetation) and agriculture/herbaceous vegetation area was done at the municipality scale. Within Central America, there are a total of 1,188 municipalities (Belize = 6; Costa Rica = 82; El Salvador = 264; Guatemala = 332; Honduras = 293; Nicaragua = 142; and Panama = 66). For each municipality, a linear regression was calculated for either woody vegetation or agriculture/herbaceous vegetation area (dependent variable) against time (independent variable, 10 y). If more than 1% of the total municipality area had pixels mapped as "No Data" for a given year, then the class area for that year was removed from the regression. Regression models were only fit for municipalities that had 3 or more years of valid class area data. Absolute areas of woody vegetation and agriculture/herbaceous vegetation were reported for 2001 and 2010 using estimates from the linear regression model developed for each municipality.

Relating Land-Cover Change with Socioeconomic Variables. To measure the strength of linear dependence between country level socioeconomic variables and woody vegetation change, we used Pearson product-moment correlation coefficients between the country level measures and the aggregated net forest cover change during the 10-y period. We show three measures of relative forest cover change by dividing the area of forest cover change by country area: (*i*) relative total forest change; (*ii*) moist forest cover; and (*iii*) forest cover in dry + coniferous forests. These variables were regressed against the following socioeconomic variables derived from the Central Intelligence Agency World Factbook (13) using ordinary least squares regression: country population in 2010 (POP); population change between 1990 and 2010 (PCH9010); remittances per capita in 2009 (REM_PC); total foreign investment in 2010 (FOR_INV); foreign investment per capita in 2010 (FINV_PC); human development index (HDI) in 2010 that combines income, education, and health subcomponents; gross domestic product (GDP) per capita in 2010 (GDP_PC); international migration rate in 2010 (MIGRA); percentage of urban population in 2010 (%URBAN); urbanization rate in 2010–2015 (URRATE); infant mortality in 2010 (INFMOR); percentage of agricultural GDP over total GDP in 2006 (AGRIC); and percent of population below poverty line in 2010 (POVERT).

The variables included in this analysis are not intended to be all inclusive of all correlates of woody vegetation change in Central America; the rationale behind their inclusion is that these variables partially serve as expressions of drivers or their proxies of socioeconomic development and are available for all countries in Central America. These variables tend to be highly correlated because they are all related, to some degree, to changes in human well being. The comparison of their correlation coefficients was meant to assess the relative explanatory power of more integrative variables, such as HDI, in comparison with some of its particular components (e.g., infant mortality, GDP, poverty) and other variables that are not explicitly included as measures of development, but are usually related to it (e.g., percent of agriculture, urbanization, remittances).

Given that HDI was the socioeconomic variable showing the highest overall correlation with forest change (see *Results*), and its direct significance for FT theory (Fig. 1), we explored the association between HDI and four different measures of forest change: (*i*) total relative forest change (net woody cover change over country area); (*ii*) relative moist forest change [moist forest change over the country area in the moist forest biome at the beginning of the study (i.e., 2001]; (*iii*) relative dry + conifer forest change (change in these two types of forests over 2001 cover); and (*iv*) relative forest instability (absolute summation of net change in moist forest plus dry + conifer forests). The relative forest instability (forest instability index) was used to capture the level of country-scale level of forest redistribution.

1. Clark ML, Aide TM, Grau HR, Riner G (2010) A scalable approach to mapping annual land-cover at 250-m using MODIS time-series data: A case study in the Dry Chaco ecoregion of South America. *Remote Sens Environ* 114:2816–2832.
2. Clark ML, Aide TM (2011) Virtual interpretation of Earth Web-Interface Tool (VIEW-IT) for collecting land-use/land-cover reference data. *Remote Sens* 3:601–620.
3. Huete AR, et al. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens Environ* 83:195–213.
4. Jensen JR (2007) *Remote Sensing of the Environment: An Earth Resource Perspective* (Prentice-Hall, New Jersey).
5. Olson DM, et al. (2001) Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience* 51:933–938.
6. Redo DR, Aide TM, Clark ML (2012) The relative importance of socio–economic and environmental variables in explaining land change in Bolivia, 2001–2010. *Ann Assoc Am Geogr*, in press.
7. Redo DR, Aide TM, Clark ML (2012) Impacts of internal and external policies on land change in Uruguay, 2001 to 2009. *Environ Conserv* 39(2):122–131.
8. Hüttich C, et al. (2011) Assessing effects of temporal compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems: Implications for global monitoring. *Remote Sens Environ* 115:2445–2459.
9. Ghimire B, Rogan J, Miller J (2010) Contextual land-cover classification: Incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sensing Letters* 1:45–54.
10. Rogan J, et al. (2008) Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sens Environ* 112:2272–2283.
11. Pal M (2005) Random forest classifier for remote sensing classification. *Int J Remote Sens* 26:217–222.
12. Breiman L (2001) Random forests. *Mach Learn* 45:5–32.
13. CIA (2011) World Factbook. Available at https://www.cia.gov/library/publications/the-world-factbook/. Accessed May 12, 2012.

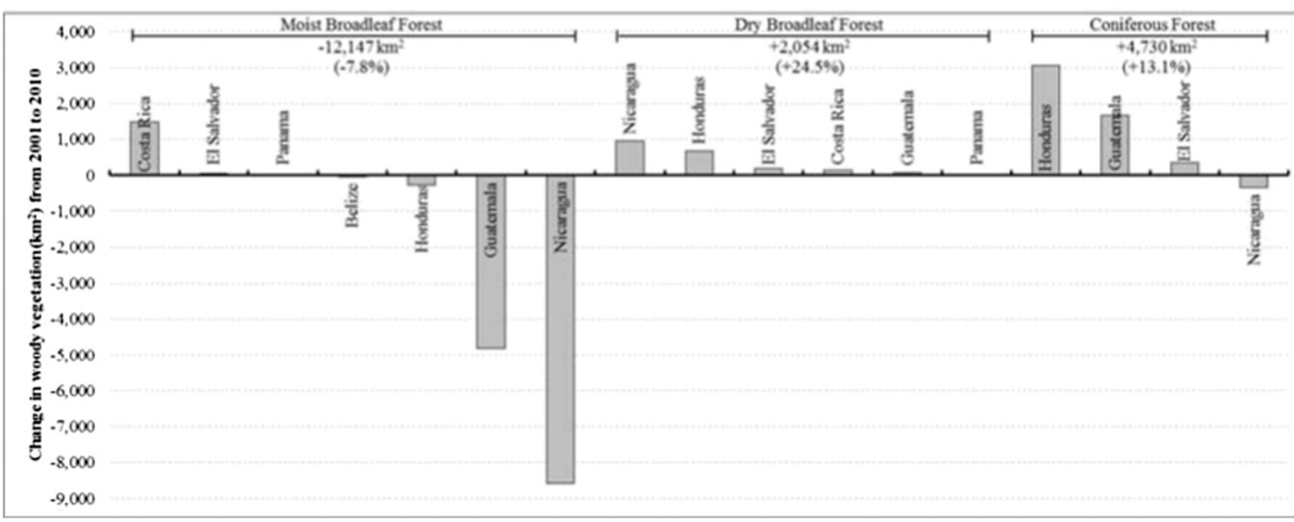

Fig. S1. Estimated change in area of woody vegetation (km^2) from 2001 to 2010 by country and by biome. Note that total forest change does not equal the total listed in main body (5,376 km^2) because of the exclusion of mangroves and deserts. Value labels indicate the total amount of woody change that occurred from 2001 to 2010 (*Upper*) and the percentage of forest lost/gained in each biome that was present in 2001 (*Lower*).

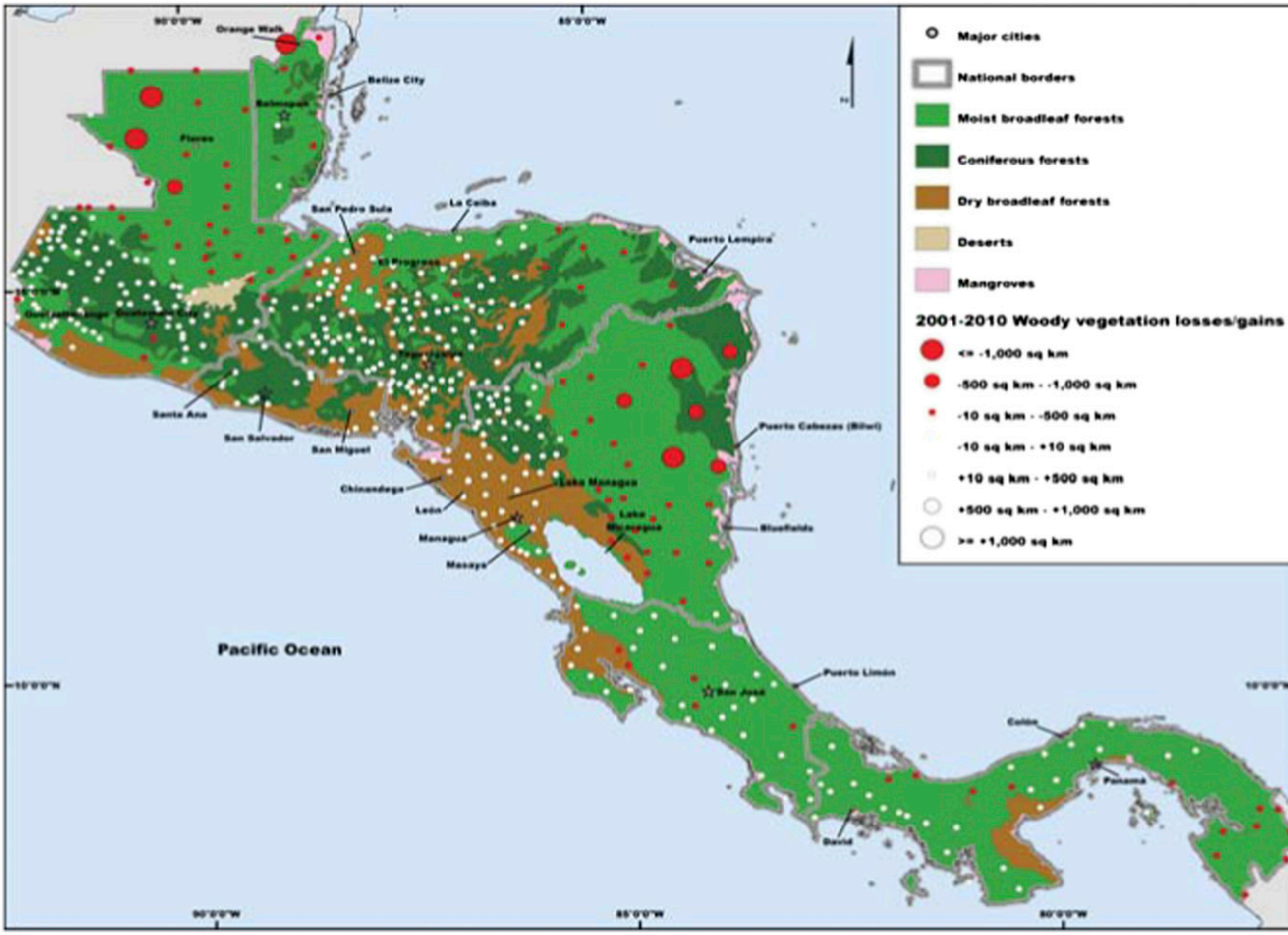

Fig. S2. Central America map showing biomes and woody vegetation gains and losses (km^2) for 2001–2010 by municipality.

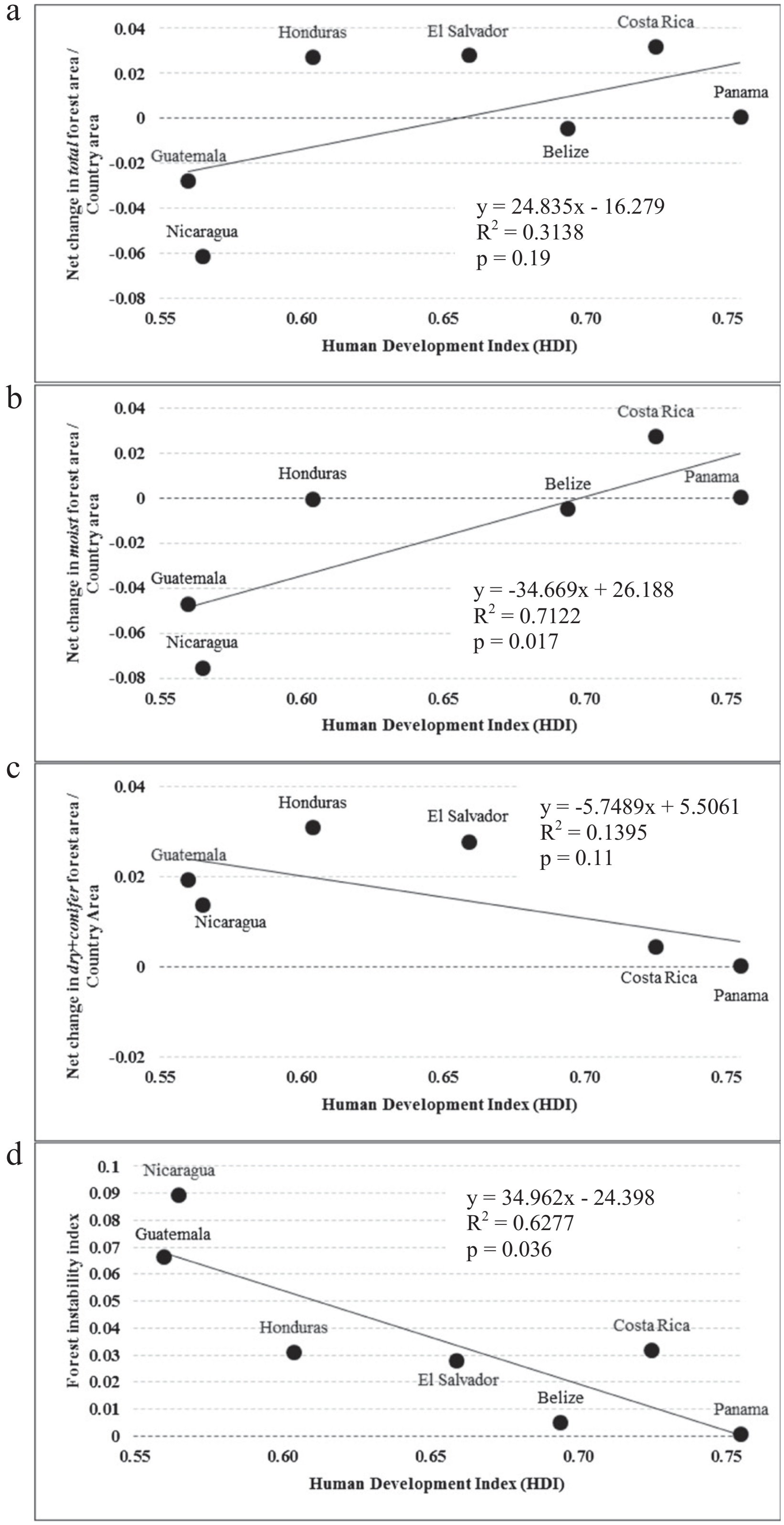

Fig. S3. Associations of HDI with: (*A*) net forest change of all forest types/country area; (*B*) net moist forest change/country area; (*C*) net conifer + dry forest change/country area; and (*D*) relative forest instability (absolute summation of moist forest and dry + conifer forest net change). Note that El Salvador is missing from *B* because moist forest accounts for only 4.8% of total area. In *C*, Belize has been eliminated because it has neither dry nor coniferous forest.

Table S1. Location of deforestation hotspots

Biome	Country	Department/province/district	Municipalities	Change (km^2)
Moist broadleaf forests	Nicaragua	Atlántico Norte	Rosita	−2,314
			Siuna	−867
			Bonanza	−431
			Waspam	−426
		Atlántico Sur	La Cruz de Rio Grande	−1,752
			Laguna de Perlas	−660
			Bluefields	−410
			El Rama	−320
			Kukra Hill	−261
		Jinotega	El Cua	−367
		Río San Juan	El Castillo	−195
	Guatemala	Petén	San Andres	−1,460
			La Libertad	−1,087
			Sayaxche	−528
			Dolores	−270
			San Luis	−266
		El Quiché	Ixcan	−179
		Alta Verapaz	Coban	−125
	Honduras	Olancho	Catacamas	−460
			Dulce Nombre de Culmí	−274
		Colón	Iriona	−409
			Limón	−63
		Gracias a Dios	Brus Laguna	−110
	Belize	Orange Walk		−155
		Stann Creek		−53
Coniferous forests	Nicaragua	Atlántico Norte	Prinzapolka	−530
			Puerto Cabezas	−591

Table S2. Location of reforestation hotspots

Biome	Country	Department/province/district	Municipalities	Change (km^2)
Moist broadleaf forests	Costa Rica	Puntarenas	Golfito	192
			Osa	149
			Buenos Aires	127
			Corredores	123
		Alajuela	San Ramón	110
			San Carlos	100
Coniferous forests	Honduras	Olancho	Gualaco	286
		El Paraíso	Danli	215
		Francisco Morazán	Distrito Central	142
			Guaimaca	78
			Lepaterique	72
	El Salvador	Santa Ana	Santa Ana	32
			Metapan	28
		Cabañas	Sensuntepeque	20
			Victoria	15
	Guatemala	Quiché	Chicaman	102
			Nebaj	49
			Uspantan	48
		Baja Verapaz	Cubulco	60
Dry broadleaf forests	Honduras	Olancho	Juticalpa	101
		Yoro	El Progreso	39
		Choluteca	Pespire	39
		Ocotepeque	San Marcos	33
	Nicaragua	León	El Sauce	101
			La Paz Centro	65
			Malpaisillo	52
			Achuapa	46
		Managua	Mateare	93
		Rivas	San Juan Del Sur	71
	Costa Rica	Guanacaste	Santa Cruz	69
			La Cruz	49
			Liberia	45
			Carrillo	35
	El Salvador	San Miguel	Chirilagua	21
		La Union	Poloros	14
			San Alejo	14
			Conchagua	12
	Panama	Coclé	Anton	12